BEBR

FACULTY WORKING
PAPER NO. 1389

Analysing Third-World Urbanization: A Theoretical Model with Empirical Evidence

Jan K. Brueckner

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 1389

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

September 1987

Analysing Third-World Urbanization: A Theoretical Model with Empirical Evidence

Jan K. Brueckner, Professor
Department of Economics

Abstract

This paper proposes and tests a simple model of third-world urbanization. The theoretical framework results from imbedding the urban economist's monocentric city model in an economy experiencing rural-urban migration. When urban and rural real incomes are set equal to guarantee migration equilibrium, an equilibrium city size is determined by the model. This city size depends on a variety of variables describing the urban and rural sectors of the economy. To test the model, urbanization measures and urban growth rates are regressed on these variables using cross section data from a small number of third world countries.

Analysing Third-World Urbanization: A Theoretical Model with Empirical Evidence

by

Jan K. Brueckner*

1. Introduction

Economic development in the third world is being accompanied by explosive urban growth. United Nations data summarized by Rogers (1982) show that while annual urban growth rates in developed countries ranged between 1.5 and 2.4 percent from 1950 to 1990 (projected), third-world cities grew at rates between 3.9 and 4.7 percent over this period. This growth has more than doubled the urban share of the third-world population, which rose from around 17 percent in 1950 to a projected 36 percent in 1990. The urban share in developed countries, by contrast, is projected to rise from 53 to 75 percent over this period. Rapid third-world urbanization has also created very large cities. The U.N. data show that while developed countries claimed 11 of the world's 15 largest cities in 1950, the top 15 will include only 3 developed-country cities by the year 2000 (these are Tokyo, Los Angeles, and New York). Moreover, of the 414 cities expected to house a million or more people in the year 2000, a majority of 264 will be located in third-world countries.

Urban growth has two sources: rural-urban migration and natural increase of the urban population. Although high birth rates make the latter source an important factor in third-world city growth, rural-urban migration plays a more important role than in developed countries. Such migration has been the subject of intense study by economists, demographers, and other researchers. The main lesson of empirical work in this area has been that migration in third-world countries appears to be economically rational, with rural-urban migrants lured to cities by the prospect of better living standards (see Fields (1982) and Schultz (1982) for recent contributions). An important theoretical insight underlying this research is that since the impetus to rural-urban migration is expected income gain (in a probabilistic sense), high urban unemployment need not deter such migration if wages in the modern urban sector are appreciably higher than agricultural wages. This insight, which explained a puzzling aspect of the migration process, originated in the work of Todaro (1969) and Harris and Todaro (1970).

While interest in rural-urban migration has been long-standing, attempts by economists to construct comprehensive, migration-based models of third-world urbanization have been more recent. The watershed work in this area is that of Kelley and Williamson, which culminated in the 1984 monograph entitled What Drives Third World City Growth? This book describes the structure of a rich and complex computable general equilibrium (CGE) model built around a Harris-Todaro migration mechanism.[1] Simulations of the model accurately reproduce the recent history of third-world city growth and yield provocative predictions about future urbanization. Building on Kelley and Williamson's work, Becker and Mills (1986) and Becker, Mills and Williamson (1986) constructed a similar CGE model of Indian urbanization.[2]

Although the performance of the CGE models is impressive, their complexity strains the economic intuition of the average reader and rules out standard empirical testing (validation of the models relies instead on simulation exercises). In fact, simple theoretical models of third-world urbanization that are amenable to empirical testing are curiously lacking in the literature. Research in this area in effect appears to have skipped an entire generation of potential models in arriving at the current state of the art. The purpose of the present paper is to help fill this gap by proposing and testing an elementary model of third-world urbanization. The theoretical framework results from imbedding the urban economist's monocentric-city model in an economy experiencing rural-urban migration. When urban and rural real incomes are set equal to guarantee migration equilibrium, an implied equilibrium city size is determined by the model. This city size depends on a variety of variables describing the urban and rural sectors of the economy (important variables are urban and rural income levels). The theoretical predictions of the model, which are developed in section 2 of the paper, are tested in section 4 through cross-section regressions relating urbanization measures for third-world countries to the explanatory variables identified in the theoretical analysis (the data are described in Section 3).

It is important to realize at the outset that the partial equilibrium nature of the model limits its ability to address fundamental questions. Because some explanatory variables are ultimately endogenous but are not determined within the model, the analysis cannot, for example, identify the ultimate sources of urbanization in the way that a general equilibrium framework can.[3] In spite of this, the paper provides useful information by answering the following more limited question: do urbanization levels and (endogenous) explanatory variables such as the urban-rural income differential vary across countries in a way that is consistent with the hypothesis that real incomes are equalized between city and countryside? The answer to this question is clearly important since the real-income-equalization hypothesis lies at the heart of most recent research on third-world urbanization.

2. The Theoretical Model

The analysis imbeds the standard urban model developed by Alonso (1964), Mills (1967), Muth (1969), and Wheaton (1974) in an economy with rural-urban migration. All consumers in the economy are assumed to have identical preferences for housing (q) and nonhousing consumption (c). Urban residents are employed in the modern sector, where they earn an income of y per period. Rural residents are employed in a traditional agricultural sector and earn an income of $y_A < y$. For the moment, unemployment is assumed not to exist in either sector.

Agricultural land earns a rent of r_A. Under the assumption (which is relaxed below) that housing is produced with land alone, the price of housing faced by rural residents is simply r_A. Urban land (housing) prices are determined according to the standard model. In this model, locational equilibrium requires that consumers living in different parts of the city reach the same utility level in spite of differences in commuting costs to the central workplace. Letting t denote commuting cost per mile, x denote radial distance from residence to the center of the city, and r denote urban land rent, the budget constraint of a representative consumer is $c + rq + tx = y$. With $v(c,q)$ denoting the utility function, the conditions that guarantee that utility is the same regardless of residential location are

$$v_q(y-tx-rq,q)/v_c(y-tx-rq,q) = r \qquad (1)$$

$$v(y-tx-rq,q) = u, \qquad (2)$$

where subscripts denote partial derivatives and u is the uniform urban utility level. Equation (1) indicates that q is chosen optimally, and (2) requires that the resulting utility level equals u. These equations determine land rent and land consumption as functions of x, y, t, and the utility level u, which is ultimately an endogenous variable. These dependencies can be written

$$r = r(x,y,t,u) \tag{3}$$

$$q = q(x,y,t,u) \tag{4}$$

Well-known results are $r_x < 0$ and $q_x > 0$, which indicate that land rent falls with distance to the center of the city to compensate consumers for lengthier commutes and that land consumption rises in response.

The overall equilibrium of the city is determined by the requirements that the urban land rent equal the agricultural rent r_A at the urban boundary m and that the urban population fit inside the boundary. These conditions can be written

$$r(m,y,t,u) = r_A \tag{5}$$

$$\int_0^m (2\pi x/q(x,y,t,u))dx = P \tag{6}$$

Note in (6) that 1/q equals population density and thus that $(2\pi x/q)dx$ is the population of a narrow ring of radius x. Equations (5) and (6) determine m and u as functions of underlying parameters:

$$m = m(P,y,t,r_A) \tag{7}$$

$$u = u(P,y,t,r_A). \tag{8}$$

Analysis by Wheaton (1974) established that $m_P > 0$, $m_y > 0$, $m_t < 0$, and $m_{rA} < 0$, showing that the distance to the urban boundary is an increasing function of population P and income y and a decreasing function of commuting cost per mile t and agricultural rent r_A. These results are crucial in the ensuing analysis.

The key step in the analysis of third-world urbanization is to combine (7) with the condition for rural-urban migration equilibrium. Since all urban residents achieve the same utility level, this condition can be developed by focusing on a resident living at the urban boundary. Such an individual pays r_A for his land by (5) and has disposable income net of commuting cost of $y - tm$. Since a rural resident also pays r_A for land and faces the same price (unity) for the nonhousing good c as the urban resident, the two individuals will be equally well off when net incomes are equalized, or when $y_A = y - tm$.[4] Recalling (7), this condition can be written

$$y_A = y - tm(P,y,t,r_A). \qquad (9)$$

Equation (9) is the critical relationship in the model. The equation implicitly defines the urban population size P that equates rural and urban real incomes for given values y, y_A, r_A, and t, yielding $P = P(y,y_A,r_A,t)$.[4] The partial derivatives of P with respect to these variables are found by differentiation of (9), which yields

$$P_y = (1 - tm_y)/tm_P > 0 \qquad (10)$$

$$P_{yA} = -1/tm_P < 0 \qquad (11)$$

$$P_{rA} = -m_{rA}/m_P > 0 \qquad (12)$$

$$P_t = -(m + tm_t)/tm_P < 0 \qquad (13)$$

The inequalities in (10)-(13) state that the urban population is an increasing function of the urban income level y and the agricultural rent level r_A and a decreasing function of the agricultural income level y_A and the commuting cost parameter t. Before considering the intuition behind these results, note that (11) and (12) follow directly from the facts (noted above) that m is increasing in P and decreasing in r_A. The inequalities in (10) and (13) follow because disposable income at the urban boundary is increasing in y and decreasing in t, or $(1-tm_y) > 0$ and $(m+tm_t) > 0$. These last two facts can be established by noting that the urban utility level from (8) increases with y and decreases with t (see Wheaton (1974)). Since the urban boundary resident faces a fixed prices and experiences these utility changes, it follows that disposable income at the boundary must rise and fall with y and t, as claimed.

The intuitive explanation for the results in (10)-(13) is straighforward. First, an increase in the urban income level y raises the urban standard of living relative to that in rural areas.[5] This creates an impetus for migration, which increases the urban population. By raising urban land prices, this population increase depresses real income in the city, dampening the incentive to migrate. Eventually, P rises far enough to reduce the urban living standard back to its original level, restoring equilibrium. When agricultural income y_A rises, the reverse process unfolds. A higher agricultural living standard lures urban residents to the countryside, and the resulting decline in P lowers urban land prices and raises real income in the city. When P has fallen enough to equate urban and rural living standards, equilibrium is restored. Similarly, an increase in the commuting cost parameter t reduces the real income of city dwellers and leads to an equilibrating migration flow to the countryside.[6] Finally, when the agricultural land rent r_A rises, real incomes decline for both urban and rural residents. However, since nominal income stays constant in the countryside while the disposable income of the resident at the urban boundary rises,[7] it follows that real income falls less in the city than in the countryside. The result is migration toward the city, which proceeds until living standards are equalized.

While some readers might question the quantitative significance of r_A's impact on the equilibrium, evidence from Brueckner and Fansler's 1983 study of the determinants of urban spatial sizes shows that agricultural land values do exert a significant negative impact on the spatial areas of small to medium-size U.S. cities, as equation (7) above would predict (this result controls for income and population size). This effect, which has an associated elasticity of -.23, suggests that the impact of r_A in the present framework can be quantitatively important.

Although the above discussion treats y, y_A, and r_A as parametric, these variables will in fact be influenced by the allocation of population between the city and the countryside. A simple marginal productivity argument, in fact, would predict that y would decline and that y_A and r_A would rise and fall respectively as labor shifts from agricultural to urban employment. This possibility affects the preceding analysis only in that it changes the interpretation of the equilibrium relationship (9). This relationship must now be viewed as one equation in a larger simultaneous system that jointly determines equilibrium values of P, y, y_A, and r_A. As noted in the introduction, the model's focus on just one equation from this system means that it is not able to identify the ultimate sources of urbanization, as would be posssible in a general equilibrium framework. Whatever the sources of urban growth, however, the migration equilibrium condition (9) is still relevant, and its predictions (as reflected in (10)-

(13)) can be tested as long as the explanatory variables are properly viewed as endogenous.

Another observation is that since the model determines a unique P, it appears to be inconsistent with the existence of a range of city sizes. To make the model realistic in this regard, all that is needed is to introduce a range of y values reflecting differences in the composition of employment across cities. Variation in y then leads to a range of equilibrium city sizes under the model, with residents of each city enjoying the same standard of living.

To ease empirical implementation of the model, it is useful to impose another assumption that results in a convenient simplification of equation (9). This assumption is that the utility function $v(c,q)$ is of the Cobb-Douglas variety. The appendix proves that under this assumption, the function (7) relating the urban boundary m to parameters is homogeneous of degree zero in its last three arguments. This means that the identity $m(P,y,t,r_A) \equiv m(P,1,t/y,r_A/y)$ holds. Substitution in (9) then yields $y_A = y - tm(P,1,t/y,r_A/y)$, and dividing through by y gives

$$Y = 1 - Tm(P,1,T,R), \tag{14}$$

where

$$Y = y_A/y \tag{15}$$

$$T = t/y \tag{16}$$

$$R = r_A/y. \tag{17}$$

Equation (14) shows that in the Cobb-Douglas case, the equilibrium population P depends only on the ratios Y, T, R and not on the levels of the underlying variables. Differentiation of (14) yields

$$P_Y = -1/Tm_P < 0 \tag{18}$$

$$P_T = -(m + Tm_t)/Tm_P < 0. \tag{19}$$

$$P_R = -m_{rA}/m_P > 0. \tag{20}$$

These results show that an increase in either y_A/y or t/y lowers P and that an increase in r_A/y raises P. While the effects of Y and T are intuitively clear given the positive effect of y and the negative effects of y_A and t on P, the positive impact of R is not as obvious given that increases in r_A and y both raise P. A clear advantage of this modified formulation from an empirical point of view is that rather than being denominated in the currency units of a given country, the explanatory variables are now unit-free, having been normalized by the urban income level.

From an empirical perspective, it could be argued that it is unrealistic to expect third-world economies to conform to the strict predictions of an equilibrium model. A preferable approach might be to view such an economy as slowly adjusting to the equilibrium implied by the above model. As in a standard stock-adjustment model, the speed of adjustment could be assumed to depend on the difference between the equilibrium urban population P(Y,T,R) from (14) and the current population P_O. Letting P* denote the time derivative of P, this assumption yields

$$P^* = f[P(Y,T,R) - P_O], \tag{21}$$

where f is a function satisfying $f' > 0$ and $f(0) = 0$. Using (18)-(20), it follows from (21) that $P_Y^* < 0$, $P_T^* < 0$, and $P_R^* > 0$. In other words, the rate of urban growth in this formulation is a decreasing function of Y and T and an increasing function of R. Also, (21) shows that an increase in current population P_O lowers P*.

While the above analysis has been based on the assumption that land is the only input into housing production, the results are essentially unchanged when a more realistic housing production process that uses both land and capital as inputs is introduced. This claim relies on Brueckner's (1987) demonstration that Wheaton's (1974) results signing the partial derivatives of m and u in (7) and (8) also apply to an urban economy where capital is used in housing production (Wheaton considered the land-only case). This equivalence means that the results in (10)-(13) giving the signs of P's partial derivatives are valid in the more realistic model. However, in order to carry out the normalization in (14) in the new model, it is necessary for the housing production function as well as the consumer utility function to be Cobb-Douglas (the proof of this fact is available on request). As long as these assumptions (which are used frequently in urban economics and elsewhere) hold, the convenient ratio form of the model applies in a realistic production setting.

As noted earlier, the model assumes that there is no unemployment in either the rural or the urban sector of the economy. One way of incorporating unemployment would be to replace y and y_A in the model by expected incomes gy and $g_A y_A$, where g and g_A are one minus the unemployment rates in rural and urban areas (this assumption follows Harris and Todaro (1970)). While this is an attractive modification on theoretical grounds, it cannot be implemented empirically since data on sectoral unemployment rates are not available in third-world countries. A related problem is that the model does not include the value of social services such as health care and education that are more readily available to urban than rural residents. Once again, the presence of these amenities, which raises living standards in cities, cannot be measured empirically in a satisfactory way.

Before proceeding to empirical implementation of the model, it is useful to contrast the current framework with the structure of the CGE models. First, the present model's equilibration mechanism, where urban population adjusts to equate urban and rural standards of living, is also present in the CGE models. By capturing general equilibrium feedbacks, however, these models solve for the urban and rural income levels and agricultural rent, which are not determined within the present partial equilibrium framework. Although the CGE models are rich in detail, the present model is in fact more detailed in one respect since the urban area has an explicit spatial structure. This permits the spatial size of the city to be determined endogenously through equalization of urban and rural land rents. The CGE models, by contrast, assume a fixed urban land area, which means that spatial growth of the city plays no role in the equilibration process.

3. Data

Given that agricultural land rent was estimated indirectly (as explained below), the principal difficulty in data collection was finding suitable rural and urban income measures. Acceptable measures were available for a small number of countries in two different data sources, resulting in the selection of two separate but overlapping country samples. Before discussing the income data and identifying the two samples, however, it will be convenient to discuss measurement of the other variables.

Three distinct urbanization measures were used as dependent variables in the regressions. The first is population of the country's largest city in 1975. Given that the reported populations of third-world cities vary widely across data sources, it seemed advisable to use a reliable secondary source for this information. Urbanization data in the World Bank's World Development Report allowed indirect computation of the largest-city population, as follows. The country population for 1975 was multiplied by the 1975 percentage of the population urbanized, with the result multiplied by the percentage of the urban population in the country's largest city. The resulting variable is called LGCITP75. The second dependent variable is the 1975 percentage of the population urbanized, or UP75. Since the analysis of section 2 is relevant to the determination of absolute city sizes, this urbanization measure is, strictly speaking, an improper choice for a dependent variable. In other words, since a country with a large fraction of its population urbanized is not necessarily a country with large cities, the model predictions may not be relevant to a regression with UP75 as dependent variable. This objection was not heeded, however, on the grounds that UP75 is a logical measure of the extent of urbanization in a country. To implement the disequilibrium version of the model, urban growth rates (again from the World Development Report) were tabulated. UG6070 and UG7080 represent the average annual growth rates for the urbanized population over the 1960-1970 and 1970-1980 decades respectively. The 1960-1980 average growth rate, denoted UG6080, is the average of these variables.

Since cross section data on agricultural land values are unavailable, an indirect approach was used to construct a measure of r_A. First, assume that agricultural output at the farm level in third-world countries is determined according to the Cobb-Douglas production function $Z = \theta S^{\tau} L^{\mu}$, where Z is output and S and L are inputs of land and labor respectively. Then, letting p be the price of the country's agricultural output, the first-order condition for choice of S is $\tau pZ/S = r_A$. This condition says that agricultural rent is proportional to the value of output per acre. Exploiting this relationship, gross domestic product in agriculture from the World Bank's World Tables was divided by hectares of arable land in each country (the latter figure, which excludes pasture and forest, is from the FAO Production Yearbook of the U.N. Food and Agriculture Organization). The resulting quantity is proportional to r_A under the above assumptions. Note that while this procedure realistically allows output prices to vary from country to country, it does assume that a single agricultural production function applies to all countries and crops. Without this assumption, τ will be country-specific and the r_A estimates will not be comparable in cross section.

In constructing a measure of the t variable, it must be recognized that commuting cost has both a direct monetary component and a time cost component.[8] While the monetary cost can be measured by using public transit fare data, as explained below, time cost is more difficult to capture. Fortunately, however, time cost can be ignored given a few plausible assumptions. First, suppose that commuting time is valued at some fraction σ of the urban wage rate, with σ the same in all countries. Furthermore, suppose that the speed of travel is the same in all cities, with α equal to the time required to commute one mile.[9] Then t_H, the time cost component of t, will be equal to $\alpha\sigma y$. Letting t_M denote t's monetary component, the variable $T = t/y$ can then be written $(t_M+t_H)/y = t_M/y + \alpha\sigma$. Therefore, under the above assumptions, cross-sectional variation in T is solely a result of variation in t_M/y, which means that monetary costs alone need be measured. While constancy of σ across countries seems plausible, the assumption of a uniform speed of travel may be criticized on the grounds that congestion levels will be higher (and commuting slower) in large cities. As a first approximation, however, the assumption seems defensible.

Public transit fare data from the 1979 International Statistical Handbook of Urban Public Transit were used to construct the t_M variable. Minimum and maximum bus fares (corresponding to shortest and longest trips) were tabulated for the largest city in each country. Although weighted averages of these fares (the average of the minimum and the maximum or perhaps the minimum itself) could be used directly to represent t_M, this procedure probably yields an incorrect measure of commuting cost per mile given that absolute fares will be higher in large cities because of the greater distance travelled within each fare zone. This suggests that fares should be normalized in some manner by the population of the city. A possible normalization is suggested by the results of Brueckner and Fansler (1983), who showed that urban spatial area is nearly proportional to population, other things equal.[10] Since city area is equal to πm^2 in the model of Section 2, this result implies that the distance to the edge of the city is proportional to the square root of population ($m = kP^{1/2}$). Assuming that the average of the minimum and maximum fares (AVGFR) corresponds to a trip of length m/2, it follows[11] that the cost per mile for such a trip will be proportional to $AVGFR/P^{1/2}$. Accordingly, $AVGFR/LGCITP75^{1/2}$ was used as a proxy for the t_M variable. Note that since largest-city values are used in computing this proxy for t_M, it is implicitly assumed that largest-city fares are similar on a cost-per-mile basis to fares in other cities of the country.

The use of two different sources for rural and urban income data led to the construction of two different samples of countries. The first data source is Jain (1975), which provides extensive income distribution data for most countries of the world and reports rural and urban income levels for a handful of third-world countries. The unavailability of urbanization or transit data for some of the latter countries reduced the number of usable observations to thirteen, as shown in Table 2. For each country, the reported yearly income figures for the year closest to 1970 were tabulated and converted to 1970 levels using the country's consumer price index.[12] The Y variable was computed by taking the ratio of the resulting rural and urban incomes, and the R variable was computed by dividing the r_A estimate described above (agricultural GDP per hectare of arable land, computed for the year 1970) by the urban income value. The T variable was computed by multiplying AVGFR/LGCITP75$^{1/2}$ by 288 to convert to an approximate yearly basis and then dividing by the urban income value.[13] Table 2 shows summary statistics for Y, R, and T as well as the urbanization variables UP75 and UG6080 for the sample.

Income estimates in the second sample of countries rely on wage data from the U.N. International Labor Organization's Yearbook of Labor Statistics. Agricultural income is measured by the 1975 monthly wage in agriculture, and urban income is represented alternatively by the monthly wage in manufacturing and the monthly wage in construction, both for 1975.[14] While it is clearly inaccurate to assume that urban incomes are identical to wages in either of these sectors, the agriculture-manufacturing or the agriculture-construction wage differential may still give an acceptable approximation to the rural-urban income differential in a third-world country. The 25 countries where manufacturing wages and the other variables were available are shown in Table 3. The construction-wage sample is a 17-country subset of this sample (the countries are indicated by asterisks).

In addition to the measurement issue discussed above, there are various comparability problems associated with the Yearbook data. First, wages are reported as a rate per pay period for some countries and as earnings per pay period for others, with the latter definition including overtime pay. There is no obvious way of adjusting for this reporting difference. Another problem is that reported agricultural wages for some countries include the value of both cash payments and payments in kind while for other countries, wages correspond to cash payments only. Again, there is no obvious remedy for this problem. Furthermore, the reported pay period differs from country to country, ranging from hour to day to week to month. To convert to monthly equivalents, it was assumed that workers work 48 hours per week and 24 days per month. The first assumption appears reasonable in light of sketchy hours data contained in the Yearbook, which show hours per week falling between 40 and 50 in the sample countries.[15]

The Y, R, and T variables were computed as in the other sample. Y was set equal to the ratio of the agricultural wage and the urban income proxy (either the manufacturing or the construction wage). R was computed by dividing the 1975 estimate of r_A by the urban income proxy, and T was set equal to 24 times AVGFR/LGCITP75$^{1/2}$ divided by the urban income proxy (recall the assumption of 24 workdays per month). Table 3 shows summary statistics for Y, R, T, UP75, and UG6080 for both the manufacturing-wage and the construction-wage samples.

As can be seen from Tables 2 and 3, variation in the Y and R variables is larger in the wage-based samples than in the income-based sample. For example, while the maximum R in the manufacturing-wage sample is 88 times as large as the minimum value, R shows only a six-fold increase in the income-based sample. Furthermore, the maximum Y is larger than unity in both wage-based samples, indicating that reported agricultural wages are higher than manufacturing or construction wages in some countries (the maximum Y is by constrast less than one in the income-based sample). These comparisons suggest that the wage data may contain a substantial noise component due to measurement problems within each country. Regression results from the wage-based samples should therefore be viewed with caution.

4. Empirical Results

a. The income-based sample

Regression results for the income-based sample are presented in Table 4. In view of the potential simultaneity problem discussed in Section 2, both ordinary least squares and two-stage least squares estimates are presented. The OLS results, which reflect linear regressions, are discussed first.[16]

The results of regressing LGCITP75 on Y, R, and T are shown at the top of the Table. Although the R^2 for the equation is a paltry .0085 and the t-ratios (shown in parentheses) are low, the signs of the estimated coefficients are exactly as predicted by the model. A high rural-urban income ratio depresses the population of the largest city, as does a high ratio of commuting cost to urban income. Conversely, a high ratio of agricultural rent to urban income raises the population of the largest city. Although the model predicts that the population of a country should have no effect on the size of its cities, other things equal, the 1975 country population (P75) was added as an explanatory variable to see if the fit of the equation could be improved. The third line of Table 4 shows that this modification dramatically raises the R^2 of the equation without changing the signs of the Y and R coefficients (their t-ratios do improve somewhat). The T coefficient, however, changes sign in the modified equation.

The next section of the Table shows the regression results when UP75 (percent of the population urbanized) is the dependent variable. While low, the R^2 for the equation is an acceptable .13 and the signs of the coefficients are again exactly as predicted by the model. High values of Y and T depress the urbanized population while a high value of R raises it. In addition, although the Y coefficient is still not significant, its t-ratio is now greater than one in absolute value.[17]

The last three sections of the Table show the results of regressions that use urban growth rates as dependent variables. The discussion in Section 2 showed that in a disequilibrium model, the impacts of Y, R, and T on urban growth are in the same direction as their impacts on city size in an equilibrium setting. In addition, the current urban population should enter the growth equation with a negative coefficient. In the estimated equations, the percent of the population urbanized at or near the base year of the growth period plays the role of current urban population. UP75 is used for the 1970-1980 regression, and the analogous variable UP60 is used for the 1960-1980 and 1960-1970 regressions. While it seems desirable to use Y, R, and T values corresponding to either the beginning or the midpoint of the growth period, the 1960-1970 regression violates this principle since these variables are measured at the end of the period.

The OLS growth regressions exhibit respectable R^2's, uniformly negative Y coefficients, and uniformly positive R coefficients, in conformance with the predictions of the model. Interestingly, the R coefficients are significant at the 5% level or nearly so (the critical value is 2.306), and the t-ratio for the Y coefficient in the last regression is appreciably larger than previous values. The performance of the other variables is mixed. The T coefficients consistently show the wrong sign (positive), and the UP coefficients are unstable in sign. The t-ratios on these coefficients are low, however.

To get a feel for the quantitative meaning of the results, consider first the implied elasticities from the UP75 equation. At the sample means, the elasticity of UP75 with respect to Y equals -.60. This means that if the urban-rural income ratio were to increase by 10 percent (depressing Y by 10 percent), a six percent increase in the urban share of the population would result, restoring equality between urban and rural living standards (UP75 would rise from .345 to .366). Similarly, the T elasticity in the UP75 equation equals -.09, which means that a fifty percent increase in T would depress the urban share of the population by 4.5 percent (from .345 to .339). This effect is of some policy relevance since such an increase in T could be engineered by a fifty percent increase in public transit fares. The relatively small impact on UP75 of such a large jump in fares indicates that public transit pricing policy may not be an effective tool for controlling city sizes in third-world countries. Finally, the T effect in the growth equations is in the wrong direction, but the Y elasticity of -.24 in the UG6080 equation indicates that a 10 percent increase in the urban-rural income ratio would raise the urban growth rate by 2.4 percent (from 4.4 percent to 4.5 percent).[18]

In spite of some unfavorable results, the OLS regressions are fairly encouraging. The coefficient of the key Y variable consistently has the correct negative sign, indicating the high urban-rural income ratios (low

Y's) are associated with high levels of urbanization and rapid urban growth. The R variable also performs as expected, with high ratios of agricultural rent to urban income associated with extensive urbanization and rapid growth. While the frequently poor t-ratios in the regressions could be used to dismiss the results, it should be borne in mind that the small sample size (13 observations) militates against the emergence of significant coefficients.

Since it can be argued that Y and R are jointly determined along with the urbanization variables, OLS may be an improper estimation procedure. To address this concern, two-stage least squares estimates are also presented in Table 4. For the LGCITP75 and UP75 regressions, the exogenous variables in the reduced form were T, 1970 per capita GNP in U.S. dollars (PCGNP), the annual growth rate of the population from 1960 to 1970 (PG6070), the average percentage of gross domestic product originating in agriculture between 1960 and 1970 (GDPAG), and the population density on agricultural land (AGDEN, which equals rural population divided by arable land). In the growth equations, the relevant UP variable was viewed as exogenous along with those listed above. The reduced-form results, which are not reported, are not especially informative.[19]

Inspection of Table 4 shows that the 2SLS estimates for the LGCITP75 and UP75 equations are qualitatively similar to the OLS estimates. The 2SLS growth equations, however, show some key sign reversals relative to the OLS equations. In particular, the Y variable now shows a positive rather than a negative coefficient in each of the growth equations, and previously-negative UP coefficients are now positive. While these results are discouraging, there is good reason to discount them. The problem is that the reduced-form Y equation has a fairly poor fit, showing an R^2 of either .28 or .23 depending on which UP variable is used. Since this leads to a poor correspondence between actual and fitted values of Y, the second-stage results cannot be taken too seriously. This problem results in part from the need to rely on ad hoc structural equations, which reflects the absence of a complete model of the economy.

b. The wage-based samples

The regression results for the wage-based samples are shown in Tables 5 and 6. These results are less favorable to the model than those from the income-based sample. First, the Y coefficients are positive in the LGCITP75 and UP75 regressions, regardless of whether the manufacturing or construction wage is used as the urban income proxy. Moreover, the t-ratios on these coefficients are often large, with the coefficient in the 2SLS UP75 equation of Table 6 nearly significant at the 5 percent level.[20] The R coefficients also do not conform to predictions, being sometimes positive and sometimes negative in the LGCITP75 and UP75 regressions. Only the T coefficients consistently show the expected negative sign in these equations. Several coefficients are in fact significant, showing t-ratios near three in absolute value.

In constrast, the Y variable performs as predicted in the growth equations, with its coefficient negative in all the regressions except the last one in Table 6.[21] Moreover, the t-ratios on these coefficients are frequently large in absolute value (the coefficient in the OLS UG6080 regression of Table 6 is in fact significant). While the T variables continue to show consistently negative coefficients and appreciable t-ratios, the R coefficients in the growth equations are again of inconsistent sign. Finally, the UP variables in these equations perform better than in the income-based sample. Their coefficients are consistently negative and almost always significant.[22]

Given the various shortcomings of the wage data (especially the noise issue discussed above), it is difficult to appraise the results shown in Tables 5 and 6. On the one hand, the results are quite unfavorable to the equilibrium version of the model, which underlies the LGCITP75 and UG75 regressions. However, in spite of the poor performance of the R variable, the growth regressions are reasonably encouraging. While it might be tempting to discount the equilibrium regressions because of the unreliability of the data, the same verdict would then be in order for the more satisfactory growth regressions. If the results are to be taken seriously, however, the only way to reconcile them with the model is to assume that urbanization levels in the sample are far from their equilibrium values but are adjusting to equilibrium in the manner described in Section 2. This would simultaneously explain the poor LGCITP75 and UP75 results and the more successful growth regressions. Whether this scenario is accurate is an open question.

If the growth equations can be taken seriously, they contain some useful quantitative information. For example, the elasticity of UG6080 with respect to Y in the OLS manufacturing-wage equation equals -.19 at sample means, indicating that a 10 percent increase in the urban-rural income ratio raises the urban growth rate by 2 percent (from 4.9 percent to 5.0 percent). The T elasticity of -.10 indicates that a fifty percent increase in transit fares would slow urban growth only by 5 percent (from 4.9 to 4.65 percent). This reinforces the earlier conclusion that transit pricing policy will not be very effective at restraining urbanization in third-world countries.

A final point is that the negative Y coefficients in many of the growth regressions shown in Tables 4-6 are consistent with the predictions of an ad hoc model that may seem attractive to some readers. Without specifying details of the economy, this model predicts that urban growth rates will be high wherever urban incomes are high relative to rural incomes. At a minimum, the empirical results of this paper can be viewed as evidence in favor of such a model.

5. Conclusion

This paper has developed and tested a simple model of urbanization in third-world countries. The theoretical framework imbeds the standard monocentric-city model in an economy experiencing rural-urban migration. When rural and urban incomes are set equal to guarantee migration equilibrium, the model generates an equilibrium city size that depends on the rural-urban income ratio, the ratio of agricultural land rent to urban income, and the ratio of commuting cost per mile to urban income. The empirical work, which attempts to relate urbanization measures and urban growth rates to these variables, shows mixed results. In one sample that uses reliable income data, the signs of the regression coefficients are consistent with the predictions of the model (the coefficients, however, do not pass the usual significance tests). However, a second sample in which urban incomes are represented by less reliable wage proxies gives less encouraging results. The upshot is that, while the empirical results are suggestive, they offer at best a weak confirmation of the fundamental hypothesis that the urbanization process tends to equalize real incomes between city and countryside. In spite of this, the present paper makes a distinct contribution to the literature on third-world urbanization. By formulating and testing with cross-section data a simple urbanization model, the paper fills in a gap in the literature left by rapid progress in computable general equilibrium modelling. Small-scale models like the present one can yield useful insights and are worth elaborating in future research.

Appendix

This appendix proves that when the utility function takes the Cobb-Douglas form $c^{\alpha}q^{\beta}$, the m equation in (7) will be homogeneous in its last three arguments. First, solving for the demand functions using (1) and substituting in (2) yields

$$[\alpha(y-tx)/(\alpha+\beta)]^{\alpha}[\beta(y-tx)/(\alpha+\beta)r]^{\beta} = u, \tag{A1}$$

which can be solved for r to yield

$$r = au^{-1/\beta}(y - tx)^{(\alpha+\beta)/\beta}, \tag{A2}$$

where a is a constant. Substituting for r in the demand function for q using (A2) gives

$$q = bu^{1/\beta}(y - tx)^{-\alpha/\beta}, \tag{A3}$$

where b again is a constant. Using (A2) and (A3), equations (5) and (6) become

$$au^{-1/\beta}(y - tm)^{(\alpha+\beta)/\beta} = r_A \tag{A4}$$

$$\int_0^m (2\pi/b)xu^{-1/\beta}(y - tx)^{\alpha/\beta}dx = P. \tag{A5}$$

Eliminating u in (A5) using (A4), the equation reduces to

$$\int_0^m wr_A x(y - tm)^{-(\alpha+\beta)/\beta}(y - tx)^{\alpha/\beta}dx = P, \tag{A6}$$

where w is a constant. Finally, after rearrangement, (A6) can be written

$$\int_0^m w(r_A/y)x[1 - (t/y)m]^{-(\alpha+\beta)/\beta}[1 - (t/y)x]^{\alpha/\beta}dx = P. \qquad (A7)$$

Equation (A7) determines the solution for m in terms of the parameters P, y, t, and r_A. Since the last three parameters enter the equation only in the terms r_A/y and t/y, it follows that proportional increases in these variables leave m unchanged, establishing zero degree homogeneity of (7).

Table 1
Variable Definitions

Y - the ratio of agricultural income to urban income

R - the ratio of agricultural land rent urban income

T - the ratio of commuting cost to urban income

LGCITP75 - the 1975 population of the country's largest city

UP75, UP60 - the percentages of the country's population living in urban areas in 1975 and 1960

UG7080, UG6080, UG6070 - the average annual growth rates of the urbanized population over the periods 1970-1980, 1960-1980, 1960-1970

P75 - the 1975 population of the country

Table 2
The Income-Based Sample

Bangladesh
Brazil
Colombia
Costa Rica
Ecuador
Honduras
India
Korea
Malaysia
Pakistan
Phillipines
Sri Lanka
Thailand

variable	mean	minimum	maximum
Y	.524	.185 (Honduras)	.739 (India)
R	.281	.108 (Brazil)	.644 (Korea)
T	.491E-03	.171E-03 (Colombia)	.110E-02 (Sri Lanka)
UP75	.345	.090 (Bangladesh)	.660 (Colombia)
UG6080	.044	.033 (India and Malaysia)	.068 (Bangladesh)

Table 3
Wage-Based Samples

Argentina
Bangladesh*
Burma
Burundi*
Cameroon*
Chile
Colombia
Costa Rica*
Ghana*
Honduras*
India
Kenya*
Korea*
Malawi*
Mexico*
Morocco
Nicaragua*
Nigeria*
Pakistan
Sri Lanka*
Syria
Tanzania*
Turkey*
Upper Volta*
Zambia*

Manufacturing Wage = y (all 25 countries)

variable	mean	minimum	maximum
Y	.589	.221 (Cameroon)	1.242 (Burma)
R	5.517	.299 (Upper Volta)	26.312 (Korea)
T	.120E-02	.928E-04 (Colombia)	.509E-02 (Burundi)
UP75	.342	.020 (Burundi)	.810 (Argentina)
UG6080	.049	.020 (Argentina)	.098 (Malawi)

Construction Wage = y (17 countries with asterisks)

variable	mean	minimum	maximum
Y	.700	.280 (Cameroon)	1.518 (Burundi)
R	6.076	.376 (Upper Volta)	16.396 (Korea)
T	.184E-02	.158E-03 (Mexico)	.118E-01 (Burundi)
UP75	.278	.020 (Burundi)	.630 (Mexico)
UG6080	.054	.024 ((Burundi)	.098 (Malawi)

Table 4
Regression Results for the Income-Based Sample*

Dependent variable	const	Y	R	T	P75	UP75	UP60	R^2
LGCITP75								
ols	3.752E+03 (.87)	-6.875E+02 (-.10)	1.512E+02 (.25)	-4.960E+05 (-.12)				.0085
2sls	1.215E+04 (1.10)	-1.876E+04 (-.86)	6.732E+03 (.69)	-1.311E+06 (-.24)				
ols	5.050E+03 (1.34)	-6.662E+03 (-.98)	3.052E+03 (.58)	2.270E+05 (.06)	1.294E-02 (2.04)			.3485
UP75								
ols	5.678E-01 (2.67)	-3.942E-01 (-1.12)	4.896E-02 (.16)	-6.059E+01 (-.30)				.1307
2sls	1.415E+00 (1.80)	-2.039E+00 (-1.32)	2.276E-01 (.33)	-1.335E+02 (-.12)				
UG7080								
ols	4.360E-02 (2.71)	-1.589E-02 (-.75)	3.317E-02 (1.97)	4.122E+00 (.13)		-1.544E-02 (-.82)		.3689
2sls	-3.304E-02 (-.28)	1.083E-01 (.56)	1.312E-02 (.24)	1.136E+01 (.40)		2.416E-02 (.33)		
UG6080								
ols	4.308E-02 (3.13)	-1.997E-02 (-1.15)	3.931E-02 (2.72)	3.847E+00 (.40)			-7.501E-03 (-.34)	.5160
2sls	8.462E-03 (.19)	3.763E-02 (.52)	3.068E-02 (1.14)	7.191E+00 (.47)			1.220E-02 (.11)	
UG6070								
ols	4.115E-02 (3.46)	-2.545E-02 (-1.70)	4.923E-02 (3.94)	3.769E+00 (.46)			8.702E-03 (.46)	.6713
2sls	2.171E-02 (.70)	6.295E-03 (.12)	4.539E-02 (2.40)	5.624E+00 (.52)			1.994E-02 (.72)	

*observations = 13; t-ratios in parentheses

Table 5
Regression Results for the Manufacturing-Wage Sample*

Dependent variable	const	Y	R	T	P75	UP75	UP60	R^2
<u>LGCITP75</u>								
ols	2.791E+03	1.054E+03	2.286E+01	-7.395E+05				.0879
	(1.53)	(.39)	(.20)	(-1.25)				
2sls	-4.673E+03	1.541E+04	-7.674E+01	-1.108E+06				
	(-.75)	(1.36)	(-.38)	(-1.16)				
ols	1.735E+03	1.929E+03	3.571E-01	-6.061E+05	1.081E-02			.2392
	(.97)	(.75)	(.003)	(-1.09)	(1.99)			
<u>UP75</u>								
ols	3.781E-01	1.284E-01	-7.140E-03	-6.042E+01				.1202
	(3.15)	(.72)	(-.97)	(-1.55)				
2sls	-1.624E-01	1.244E+00	-2.040E-02	-9.675E+01				
	(-.36)	(1.51)	(-1.37)	(-1.39)				
<u>UG7080</u>								
ols	7.583E-02	-1.424E-02	-5.386E-05	-4.892E+00		-4.108E-02		.4002
	(8.17)	(-1.24)	(-.11)	(-1.86)		(-2.95)		
2sls	8.498E-02	-3.311E-02	1.368E-05	-4.361E+00		-3.829E-02		
	(4.85)	(-1.00)	(.02)	(-1.46)		(-2.45)		
<u>UG6080</u>								
ols	7.488E-02	-1.602E-02	6.709E-05	-3.605E+00			-4.976E-02	.4387
	(8.57)	(-1.39)	(.14)	(-1.39)			(-3.20)	
2sls	7.706E-02	-1.570E-02	-1.811E-04	-4.025E+00			-5.171E-02	
	(4.59)	(-.46)	(-.30)	(-1.40)			(-2.96)	

*observations = 25; t-ratios in parentheses

Table 6
Regression Results for the Construction-Wage Sample*

Dependent variable	const	Y	R	T	P75	UP75	UP60	R^2
<u>LGCITP75</u>								
ols	1.943E+03	6.645E+02	4.163E+01	-3.775E+05				.0928
	(.83)	(.20)	(.27)	(-.98)				
2sls	-2.004E+03	6.962E+03	9.856E+01	-8.150E+05				
	(-.54)	(1.23)	(.50)	(-1.53)				
ols	5.224E+02	9.107E+02	-4.016E+01	-2.698E+05	6.981E-02			.3881
	(.25)	(.33)	(-.29)	(-.81)	(2.41)			
<u>UP75</u>								
ols	2.391E-01	2.008E-01	-1.996E-03	-4.845E+01				.3757
	(2.25)	(1.35)	(-.28)	(-2.77)				
2sls	8.647E-03	6.199E-01	-3.252E-03	-7.836E+01				
	(.04)	(2.15)	(-.33)	(-2.90)				
<u>UG7080</u>								
ols	9.060E-02	-2.281E-02	-7.680E-04	-2.364E+00		-4.826E-02		.6059
	(9.28)	(-1.84)	(-1.39)	(-1.37)		(-2.23)		
2sls	8.717E-02	-2.042E-02	-4.182E-04	-2.494E+00		-4.874E-02		
	(7.47)	(-1.06)	(-.68)	(-1.13)		(-2.06)		
<u>UG6080</u>								
ols	9.178E-02	-2.588E-02	-6.795E-04	-1.954E+00			-6.199E-02	.6938
	(11.42)	(-2.18)	(-1.36)	(-1.22)			(-2.59)	
2sls	7.896E-02	6.664E-03	-5.818E-04	-5.008E+00			-9.207E-02	
	(5.90)	(.23)	(-.84)	(-1.61)			(-2.40)	

*observations = 17; t-ratios in parentheses

References

William Alonso, 1964, Location and Land Use (Harvard University Press, Cambridge, MA).

Charles M. Becker and Edwin S. Mills, 1986, Studies in Indian Urban Economic Development (Oxford University Press, New York).

Charles M. Becker, Edwin S. Mills, and Jeffrey G. Williamson, 1986, "Modeling Indian Migration and City Growth, 1960-2000," Economic Development and Cultural Change 35, 1-33.

Jan K. Brueckner, 1987, "The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model," in E.S. Mills, ed., Handbook of Urban Economics (North-Holland, Amsterdam), forthcoming.

Jan K. Brueckner and David A. Fansler, 1983, "The Economics of Urban Sprawl: Theory and Evidence on the Spatial Sizes of Cities," Review of Economics and Statistics 55, 479-482.

Gary Fields, 1982, "Place-to-Place Migration in Colombia," Economic Development and Cultural Change 30, 539-558.

John R. Harris and Michael P. Todaro, 1970, "Migration, Unemployment and Development: A Two-Sector Analysis," American Economic Review 60 126-142.

J. Vernon Henderson, 1982, "The Impact of Government Policies on Urban Concentration," Journal of Urban Economics 12, 280-303.

J. Vernon Henderson, 1986, "Efficiency of Resource Usage and City Size," Journal of Urban Economics 19, 47-70.

International Union of Public Transport, 1979, International Statistical Handbook of Urban Public Transport (Brussels).

Jain, Shail, 1975, Size Distribution of Income (World Bank, Washington, D.C.).

Allen C. Kelley and Jeffrey G. Williamson, 1984, What Drives Third World City Growth? (Princeton University Press, Princeton, NJ).

Edwin S. Mills, 1967, "An Aggregative Model of Resource Allocation in a Metropolitan Area," American Economic Review 57, 197-210.

Richard F. Muth, 1969, Cities and Housing (University of Chicago Press, Chicago).

Andre Rogers, 1982, "Sources of Urban Population Growth and Urbanization, 1950-2000: A Demographic Accounting," Economic Development and Cultural Change 30, 483-506.

T. Paul Schultz, 1982, "Lifetime Migration within Educational Strata in Venezuela: Estimates of a Logistic Model," Economic Development and Cultural Change 30, 559-593.

Michael P. Todaro, 1969, "A Model of Labor Migration and Urban Unemployment in Less Developed Countries," American Economic Review 59, 138-148.

George S. Tolley, 1987, "Market Failures as Bases of Urban Problems," in G.S. Tolley and Vinod Thomas, eds., The Economics of Urbanization and Urban Policies in Developing Countries (World Bank, Washington, D.C.).

United Nations Food and Agriculture Organization, 1970, 1976, FAO Production Yearbook (Rome).

United Nations International Labor Organization, 1970-1984, Yearbook of Labor Statistics (Geneva).

William Wheaton, 1974, "A Comparative Static Analysis of Urban Spatial Structure," Journal of Economic Theory 9, 223-237.

World Bank, 1979, 1982, World Development Report (Washington, D.C.).

World Bank, 1983, World Tables, vol. I (Johns Hopkins University Press, Baltimore).

Footnotes

*I wish to thank Kangoh Lee for excellent research assistance and James Follain for useful comments. Any errors are mine.

[1] Kelley and Williamson's work is also described in a number of papers that have appeared in various books and journals (these are cited in the 1984 monograph).

[2] Henderson (1982) has analysed the effects of government policies on the equilibrium of a system of cities, deriving results of interest in the third-world context. Rural-urban migration does not play an important role in his model, however. Also, Henderson (1986) offers an empirical analysis of agglomeration economies in Brazilian manufacturing. While the agglomeration issue is certainly relevant to the urbanization process in third-world countries, rural-urban migration is again largely a separate concern.

[3] Kelley and Williamson (1984) identified technical progress in the modern urban sector as a major source of third-world city growth.

[4] Note that this formulation is similar to the standard "open-city" model, where the urban utility level in the economy is given and population P adjusts to ensure that the residents of the city achieve this utility. In the present context, the utility of rural residents is parametric, being determined by the values of y_A and r_A. As in the open-city model, urban population adjusts to equate urban utility to this parametric level.

[5] Formally, this follows from the result that disposable income at the urban boundary is increasing in y.

[6] The decline in the urban standard of living can be inferred from the fall in the boundary resident's disposable income.

[7] This occurs because the city shrinks in area in response to the higher r_A.

[8] In order for the consumer's budget constraint to make sense when t incorporates time cost, income y must include the monetary value of commuting time and leisure time (see Muth (1969)).

[9] To be precise, α equals the fraction of the total hours available in each period that would be expended in commute trips with a (one-way) distance of one mile.

[10] This result is actually inconsistent with theory, which predicts that spatial area should rise less than proportionally with population (the reason is that a higher population leads to higher density).

[11] Cost per mile is $AVGFR/(m/2) = (2/k)(AVGFR/P^{1/2})$.

[12] The years used ranged from 1967 to 1972.

[13] Since an unknown constant has already been suppressed in deriving the t proxy, multiplication by 288 may seem pointless. However, since incomes in the second sample are on a monthly rather than yearly basis, sample-specific scaling of the t variable is warranted. The number 288 comes from the assumption made below that workers work 24 days per month (multiplication by 12 gives days per year). It should also be noted that the fare data used to compute AVGFR are from a variety of years (1973-1979). No attempt was made to deflate these values to 1970 given that fare changes are likely to be infrequent and that $AVGFR/LGCITP75^{1/2}$ is in any case a fairly crude proxy for the t variable.

[14] Where the original data applied to a year other than 1975, wages were converted to 1975 values using the consumer price index (the discrepancies were never greater than a few years).

[15] Since the number of months worked per year is likely to be less in agriculture than in manufacturing or construction, the ratio of the agricultural wage to the wage in either of these sectors is likely to overstate Y. However, if the length of the agricultural work year is similar across countries, then the wage ratio will be proportional to Y, eliminating any problem.

[16] Box-Cox transformations of the equations were explored with little effect on the results.

[17] It should be noted that a possible explanation for the negative signs of the T coefficients in the LGCITP75 and UP75 regressions is that the square root of LGCITP75 is in the denominator of T. While this could produce a negative association that has nothing to do with the model predictions, it is still appropriate on theoretical grounds to normalize fares by city population. Evidence for this comes from regressing the fare variables themselves (the minimum fare and AVGFR) on LGCITP75. The coefficients in these regressions are positive, indicating that fare levels are higher in relation to income in large cities given the longer trips involved.

[18] The R elasticities in the UP75 and UG6070 regressions are .04 and .30 respectively.

[19] In the equations without the UP variables, the only coefficient with a t-ratio larger than unity is that of AGDEN in the R equation. Its positive sign and high t-ratio (4.84) strongly indicate that countries with extreme population pressure in rural areas have high ratios of r_A to urban income. When either UP variable is added, notable changes in the reduced form are large increases in the t-ratios of GDPAG in the R equation, indicating a significantly positive impact of the percent of GDP in agriculture on R. The coefficients of the UP variables are also positive and significant in the respective R equations.

[20] The critical value is 2.16.

[21] Since the explanatory variables are for 1975, no 1960-1970 growth equation was estimated for the wage-based sample.

[22] In addition to showing a strong positive effect of AGDEN on R, as in the income-based sample, the reduced-form equations for the wage-based sample show some other interesting effects. First, the per capita GNP variable has a positive effect on Y (PCGNP's t-ratios are at least as large as unity in the various forms of the Y equation). Also, R is positively affected by GDPAG (the percent of GDP in agriculture). GDPAG's t-ratios again are in the respectable range, especially in the construction-wage sample.