355-76

INSTITUTE FOR RESEARCH ON POVERTY DISCUSSION PAPERS

AN EQUILIBRIUM MODEL OF URBAN POPULATION AND THE DISTRIBUTION OF INCOME

John Yinger and
Sheldon Danziger

UNIVERSITY OF WISCONSIN-MADISON

AN EQUILIBRIUM MODEL OF URBAN POPULATION AND THE DISTRIBUTION OF INCOME

John Yinger and Sheldon Danziger

July 1976

This research was supported by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Department of Health, Education, and Welfare pursuant to the provisions of the Economic Opportunity Act of 1964. The opinions expressed are solely those of the author.

ABSTRACT

The relationship between the level of income and the population of an urban area is a familiar concern in urban economics. In an equilbrium model of urban location, households will migrate among urban areas until real income levels are equalized across all areas. This equality of real incomes implies differentials in money incomes across urban areas. Money incomes will be higher in urban areas with larger populations to compensate households for the higher housing prices and higher transportation costs in those areas.

Existing models of the relationship between income levels and urban population assume that there is a homogeneous labor force and, hence, a world in which there is no inequality in the size distribution of income within an urban area. In this paper we model a world in which there are two classes of workers and examine the relationship between urban population and the distribution of income between these two classes. In particular, we determine what happens to the degree of inequality in money incomes as urban population increases, if each class of worker is compensated for the higher costs associated with larger urban size.

The analysis in this paper is based on a mathematical model of an urban area which allows us to calculate how the income distribution--as measured by a Gini coefficient--changes as urban population changes. This model is solved numerically for a variety of assumptions about the initial size of the urban population and its division between skill classes, initial income distribution, commuting costs, and several other parameters. The findings of the simulations using this model are then tested using data for a sample of metropolitan areas.

AN EQUILIBRIUM MODEL OF URBAN POPULATION AND THE DISTRIBUTION OF INCOME

The relationship between the level of income and the population of an urban area is a familiar concern in urban economics (Mansfield 1949; Evans 1972; Hoch 1972; Richardson 1973). In an equilibrium model of urban location in which the labor force is both homogeneous and perfectly mobile, households will migrate among urban areas until real income levels are the same for all households in all areas. This equality of real incomes implies differentials in money incomes across urban areas. Money incomes will be higher in urban areas with larger populations to compensate households for the higher housing prices and higher transportation costs in those areas

In the first section of this paper, a mathematical model of an urban area with a homogeneous labor force is presented. Relationships are derived between urban population and the level of income within an urban area, and between areas of various sizes and the distribution of income levels across areas. This one-skill-class model is simulated and tested using data for a sample of metropolitan areas in an attempt to gauge the magnitude of these relationships. In the second section, the model is respecified under the assumption that there are two types of workers. As population increases, each class of worker must be compensated for the higher costs associated with the larger urban size. These costs are shown to differ between the skill classes. Thus, the degree of inequality in money incomes within an urban area differs across urban areas when the real income level of all workers of a given skill class is constant across areas. The relationship between this intraurban degree of inequality and urban size is shown to be a function of the initial population size and its distribution between the skill classes, the initial income distribution, commuting costs, and several other parameters. The two-skill-class model is also simulated and tested with data for a sample of urban areas.

1. One-Skill-Class, Open Urban Model

The Level of Income Within an Urban Area

By assuming that migration among urban areas is costless, an open model of the urban area implies that, in equilibrium, members of a given skill class who have the same tastes receive the same utility in all cities. Our analysis is simplified by two additional assumptions: (a) that members of a given skill class all have the same tastes, and (b) that some members of every skill class live in every city.[1]

In an open model with one skill class, money incomes increase with the population of an urban area. As more people move into an urban area, the prices of land and housing at all locations in the area are bid up, and the boundaries of the city are extended. In short, both the price of housing and the average commuting distance in the city increases. Thus, workers will receive the same real income--and hence the same utility--after the population increase as before, only if their money incomes rise to reflect the higher cost of living. This model, like those of Alonso (1964), Mills (1967, 1972a, 1972b), Muth (1969), and others, simultaneously determines the prices and quantities in each of four urban markets: housing, labor, land, and capital.

The housing market is the central focus of the model. On the demand side, households are assumed to have a single earner who works in the central business district (CBD) and to face the following maximization problem:

$$\begin{aligned} &\text{Maximize} && U(Z,H) \\ &\text{Subject to} && Y = Z + P(u)H + tu \end{aligned} \tag{1}$$

where Z is a composite consumption good with a price of unity;[2] H is units of housing services; Y is money income; u is miles from the CBD that the household lives; P(u) is the price per unit of housing services at location u; and t is the per mile cost of a round trip to the CBD.

Two further assumptions complete the demand side of the housing market: first, that the utility function is Cobb-Douglas, or

$$U(Z,H) = Z^{\alpha}H^{\beta}; \tag{2}$$

and second, that t consists of operating costs (t^o) and time costs (t^yY), or

$$t = t^o + t^yY = 500c^o + (.25w/MPH)Y \tag{3}$$

where c^o is operating costs per mile, w is the value of travel time as a fraction of Y, and MPH is the speed of commuting.[3]

The central relationship in the model--the relationship between location and the unit price of housing--can be derived from this household maximization problem. The demand function for housing is derived from the first-order conditions:

$$H = (1/k)(Y - tu)/P(u) \tag{4}$$

where $k = (\alpha + \beta)/\beta$. The substitution of this demand function and the analogous demand function for Z into (2) leads to the following indirect utility function (see Solow 1972):

$$V = [(\alpha/(\alpha + \beta))(Y - tu)]^{\alpha}[(1/k)(Y - tu)/P(u)]^{\beta}. \tag{5}$$

Household mobility insures that utility (=V) will be the same at all locations within an urban area. Furthermore, if utility is constant at all locations, (5) indicates that

$$P(u) = (\gamma/V)^{1/\beta}(Y-tu)^{k} \tag{6}$$

where

$$\gamma = \alpha^{\alpha}\beta^{\beta}/[(\alpha + \beta)^{\alpha+\beta}].$$

The price-distance function (6) is a market locational equilibrium condition that defines the pattern of housing prices that makes households indifferent among locations within a city.

Following Mills (1972), the supply of housing services is assumed to be given by a Cobb-Douglas production function

$$H_s(u) = AK(u)^{1-a}L(u)^a \tag{7}$$

where K stands for capital and L stands for land. The equality of the supply and demand for housing is guaranteed by the equation

$$H_s(u) = N(u)H(u) \tag{8}$$

where N(u) is the number of households at location u--each of which consumes H(u) units of housing services.

The production function (7) and the assumptions of perfect competition lead to demand functions for land and capital which are found in the usual way from (7):

$$aP(u)H_s(u)/L(u) = R(u). \tag{9}$$

where R(u) is the land rental rate, and

$$(1 - a)P(u)H_s(u)/K(u) = r \tag{10}$$

where r is the capital rental rate.

It is also assumed that a fixed number of radians of land (ϕ) are available for residential development. Thus the supply of land is

$$L(u) = \phi u. \tag{11}$$

A final assumption about the land market is that residential uses must outbid agricultural uses for land so that the urban area extends to the point at which

$$R(\bar{u}) = \bar{R} \tag{12}$$

where $\bar{R}$ is the agricultural rental rate and $\bar{u}$ is the outer edge of the urban area.

On the assumption that capital is supplied by a national market, the supply function for capital is

$$r = \text{constant}. \tag{13}$$

In the labor market, an exogenous demand for N workers in the CBD is assumed. Since the supply of workers from location u is equal to N(u), the equality of supply and demand in the labor market is insured by

$$\int_0^{\bar{u}} N(u)du = N. \tag{14}$$

Equation (14) completes the model. To prove that income must rise as population rises so that household utility remains constant, we determine how V and N are related to the other variables in the model. From equations (7), (9), and (10), the price-distance function can be expressed as

$$P(u) = CR(u)^a \tag{15}$$

where

$$C = [Aa^a(1 - a)^{1-a}]^{-1}r^{1-a}.$$

The relationship between the price-distance and rent-distance functions defined by (15) is an important feature of the model, indicating that statements about land rents can be easily translated into statements about the unit price of housing and vice versa. For example, the unit price of housing at the outer edge of a city, which equals $P(\bar{u})$, is defined as the opportunity cost of housing (or $\bar{P}$). According to (15), this opportunity cost is related to the agricultural rental rate:

$$P(\bar{u}) = \bar{P} = CR(\bar{u})^a = C\bar{R}^a. \tag{16}$$

Substituting (16) into (6), the opportunity cost of housing is

$$P(\bar{u}) = \bar{P} = (\gamma/V)^{1/\beta}(Y-t\bar{u})^k \tag{17}$$

or

$$V = \gamma(1/\bar{P})^{\beta}(Y-t\bar{u})^{\alpha+\beta}. \tag{18}$$

Equation (18) indicates the relationship of utility, income, and city size.

A price-distance function that does not depend on utility is derived by substituting (18) into (6),

$$P(u) = \bar{P}[(Y - tu)/(Y - t\bar{u})]^k. \tag{19}$$

Now using (15) and (16), a rent-distance function that also does not depend on utility is,

$$R(u) = \bar{R}[(Y - tu)/(Y - t\bar{u})]^b \tag{20}$$

where $b = k/a$.

The relationship between population size and the other variables in our model is now apparent. Substituting (9) and (10) into (7),

$$H_s(u) = DR(u)^{1-a}L(u) \tag{21}$$

where

$$D = A((1-a)/ar)^{1-a}.$$

Substituting (4), (8), (19), (20), and (21) into (14) yields

$$\begin{aligned} N &= \int_0^{\bar{u}} [\bar{R}a\phi u(Y - tu)^{b-1}/(Y - t\bar{u})^b]du \\ &= (\bar{R}\phi/t)[Y^{b+1}/t(b + 1)(Y - t\bar{u})^b - (Y - t\bar{u})/t(b + 1) - \bar{u}]. \end{aligned} \tag{22}$$

Equations (18) and (22) describe the relationship between urban population and the level of income. Although they cannot be explicitly solved for the income level, Y, the effect of changes in population on the income level can be determined by totally differentiating the two equations with respect to V, Y, $\bar{u}$, and N (remembering from (3) that t depends on Y), and solving for dY/dN. Since this derivation requires exogenous values for dV and dN, two alternative interpretations of the results are available. First, the differentials can be interpreted as changes in a given urban area over some time period. In this case, dN is simply the change in the population of the area and dV is the change in utility in the system of urban areas (as measured, say, by the changes in median real income). Second, the differentials can be interpreted as differences between areas at a point in time. In this case, dN is the difference between population in an urban area and population in the next largest area, and, since an open model requires equal utilities across areas, dV is equal to zero. We will simplify our presentation by assuming that dV = 0, but an expression for dY/dN can be obtained for any other value of dV.

Differentiating (18) yields

$$dY = [t/(1 - t^{y}\bar{u})]d\bar{u}. \tag{23}$$

This equation has strong intuitive appeal. For example, a household at the outer edge of an urban area must receive the same utility as a household at the outer edge of an area with a larger population. Since the unit price of housing equals $\bar{P}$ at each boundary, regardless of population, the difference in spending between two such households consists entirely of commuting costs. Thus, as indicated in equation (23), the income compensation that accompanies a population increase is based on the commuting costs between the original city's edge and the city's edge after a population increase (or $td\bar{u}$). However, since a dollar increase in Y increases the time costs of commuting to $\bar{u}$, such a dollar of compensation is only worth $(1 - t^{y}\bar{u})$
$td\bar{u}$ must be divided by $(1 - t^{y}\bar{u})$ to obtain the desired income compensation.

Differentiating (22) yields

$$dN = (\bar{R}\phi/t)\{(1/t)[Y/(Y - t\bar{u})]^{b}dY - [b/t(b + 1)][Y/(Y - t\bar{u})]^{b+1}$$

$$\cdot[dY(1 - t^{y}\bar{u}) - td\bar{u}] - [dY(1 - t^{y}\bar{u}) - td\bar{u}]/t(b + 1) - d\bar{u}\}. \tag{24}$$

Setting $dV = 0$ and substituting (23) into (24) yields

$$dY/dN = 1/\{(\bar{R}\phi/t^2)[(Y/(Y - t\bar{u})^b - (1 - t^{y}\bar{u})]\}. \qquad (25)$$

Since $Y > Y - t\bar{u}$ and $b > 1$ and $\bar{R}$ and $\phi > 0$, it follows that $dY/dN > 0$; when population increases, money incomes also increase.

The Distribution of Income Among Urban Areas

The distribution of income among urban areas can be determined using equation (25). The data required to calculate the distribution of income levels among M areas (as measured, for example, by a Gini coefficient) are the pairs of numbers $(N_1, Y_1), \ldots, (N_m, Y_m)$, where the i subscript refers to area i. If the distribution of urban populations (that is, the N_i's) and one value of Y (say Y_1) are known, then $dN_i = N_{i+1} - N_i$ for $i = 2$ to M, and $Y_i = Y_{i-1} + dY_{i-1}$, where dY_{i-1} is given by (25).

By making use of the rank-size rule--that the population of area i is equal to the population of the largest area in the system divided by the population rank of area i--the distribution of income among urban areas can be calculated when the income of only one area and the population of only the largest area are known. In this case, the distribution of urban populations is determined using the rank-size rule and the distribution of income levels among areas is calculated as described above.

Simulations and Empirical Results

Given values for Y, N, and the parameters in (25), the precise relationship between income and population can be calculated. One must first solve equation (22) for the boundary of the urban area, $\bar{u}$ (see the Appendix), and then determine dY/dN using equation (25). By way of example, the values of (dY/dN) and ((dY/Y)/(dN/N)) calculated in this way for four hypothetical urban areas are given in Table 1. According to Table 1, an increase of 1000 people in area 1 leads to an increase of $.0078 in daily income. Similarly, a 1 percent increase in population leads to a .03 percent increase in income. As income increases for a given population size (comparing areas 1 and 3, and areas 2 and 4), this elasticity increases. Similarly, as population increases for a given income level (comparing areas 1 and 2, and areas 3 and 4), the elasticity also increases.

These simulated values are similar to the values estimated for a sample of 89 large Standard Metropolitan Statistical Areas (SMSAs) for 1960 and 1970. These areas have an average of about 250,000 male workers, and a range from about 60,000 to 3 million workers in 1960. Table 2 presents some simple regressions in which mean male wages in 1960 and 1970 are the dependent variables and total male employment and a set of regional dummies are the independent variables.[4] In the simulations of Table 1, the elasticity of income with respect to population was between .031 and .053, while in both 1960 and 1970, the elasticity of male wages was about .065.

Table 1

Simulation of Relationship Between Income Level and Urban Population

Description of Urban Area	$(\partial Y/\partial N) \times 1000$	$(\partial Y/Y)/(\partial N/N)$
1. N = 250,000 workers Y = \$64 per day	.0078	.0306
2. N = 1 million workers Y = \$64 per day	.0025	.0392
3. N = 250,000 workers Y = \$81 per day	.0101	.0312
4. N = 1 million workers Y = \$81 per day	.0043	.0531

Note: All areas are assumed to be circular with $\phi = 2\pi$ radians of land available. The value of $\bar{u}$ is calculated using the first step of the procedure described in the Appendix. In computing transportation costs, t, it was assumed that workers travel at 20 miles per hour, value their travel time at one-half the wage rate, and spend 10 cents per mile on pecuniary travel costs. Households are assumed to spend one-quarter of their budget on housing (k = 4), and land receives one-fifth of housing expenditures (a = .2). The agricultural rental rate is set at \$1500 per square mile per day. These assumptions are discussed in more detail in the next section.

Table 2

Estimation of Relationship Between Wage Level and Urban Population

	Mean Male Wage, 1960	Mean Male Wage, 1970
Constant	7.898	8.242
Male Employment	0.0646 (5.77)	0.0663 (6.72)
Northeast	0.0206 (0.83)	0.0506 (2.36)
Northcentral	0.0754 (3.09)	0.0798 (3.76)
West	0.0947 (3.58)	0.0711 (3.08)
R^2	.402	.461

Note: The sample contains 89 large SMSAs. The regressions are estimated using the logarithms of the wage and employment levels; t-statistics appear in parentheses below the regression coefficients. The Southern Region is the omitted regional dummy.

2. Two-Skill-Class, Open Urban Model

The Level and Distribution of Income within an Urban Area

Statement of the Problem. Having demonstrated that in equilibrium income levels increase with urban size, the model developed above can be extended to consider the relationship between urban population and the distribution of income within an urban area. For analytic manageability, two simplifying assumptions are made: first, that there are only two skill classes in an urban area; and second, that the distribution of income can be accurately measured by a Gini coefficient.[5]

In an urban area with only two skill classes, the Gini coefficient is given by the formula

$$G = N_l/(N_l + N_h) - N_l Y_l/(N_l Y_l + N_h Y_h) \qquad (26)$$

where N_i is the number of households in skill class i, Y_i is the income of households in skill class i, and the subscripts "l" and "h" refer to the low and high skill classes, respectively.[6]

The relationship between urban population and the distribution of income can be determined by examining how G changes as N_l and N_h change. However, as shown in section 1, changes in population in an open urban model lead to changes in income, so that the change in G depends on changes in all four of the variables on the right-hand side of (26). Differentiating (26) with respect to N_l, N_h, Y_l, and Y_h, and rearranging terms yields the precise statement of this relationship:

$$dG = (N_h dN_l - N_l dN_h)[(Y_h - Y_l)(N_h^2 Y_h - N_l^2 Y_l)/(N_l + N_h)^2$$

$$\cdot (N_l Y_l + N_h Y_h)^2)] - (Y_h dY_l - Y_l dY_h)[N_l N_h/(N_l Y_l + N_h Y_h)^2]. \qquad (27)$$

Equation (27) indicates that two elasticities are important in determining how G changes with urban population.[7] These elasticities are

$$e_N = (dN_l/N_l)/(dN_h/N_h) \qquad (28)$$

and

$$e_Y = (dY_l/Y_l)/(dY_h/Y_h). \qquad (29)$$

The first elasticity measures the rate of change of the low-skill population relative to the rate of change in the high-skill population; the second elasticity measures the rate of change in low-skill income relative to the rate of change in high-skill income. If both of these elasticities are equal to unity (so that $(N_h dN_l - N_l dN_h)$ and $(Y_h dY_l - Y_l dY_h)$ are both equal to zero), equation (27) indicates that the distribution of income does not change as urban population changes. But as long as $N_h > N_l$, inequality will increase with urban population if $e_N > 1$. Furthermore, inequality will decrease with urban population if $e_Y > 1$. If both e_N and e_Y are greater than unity (or both less than unity), the net effect of a change in urban population on G is ambiguous.[8]

In the model there is a fundamental difference between e_N and e_Y; the former reflects forces that are assumed to be exogenous to the model whereas the latter is determined within the model. Indeed, the changes in N_l and N_h that are reflected in e_N cause incomes to change.

The precise relationship between urban population and the distribution of income can be derived by examining the determinants of e_Y in an open urban model.

The only previous discussion of e_Y of which we are aware is found in Evans (1972). Evans argues that

> There is evidence to justify the assumptions that, firstly, the value of time spent travelling is a constant fraction of the wage rate and this fraction does not vary with income, and secondly, the average income elasticity of demand for housing is equal to one so that the amount spent on housing is a constant fraction of total income. Therefore the amount necessary to compensate households in each income group for increased rents and increased time spent travelling [due to increased population] will increase proportionately with income. On the other hand the amount necessary to compensate households for the increased direct financial costs of travel will not increase with income. Hence we would expect that the increase in wages necessary to compensate households for living in a larger city would be proportionately smaller but absolutely larger, the higher the household's income (p. 55).

In other words, as population increases the income of the poor will increase proportionately more than the income of the rich.

Evan's argument that e_Y is greater than unity can be formally related to our model. Let u* be the border between the areas inhabited by a low- and a high-skill class and note that competition insures that the price-distance functions of the two skill classes intersect at u*.[9] Individuals in a given skill class are assumed to be indifferent to their location within an urban area or across all areas. Thus, the value of e_Y can be derived by comparing the compensation received by a low-skill household that moves from this skill-class boundary, u*, in one area, to the u* in a more populous area, with the compensation received by a high-skill household making the same move. Both these households will face an identical increase in both commuting distance (the difference between the u*'s in the two areas) and in the unit price of housing (the difference between the two P(u*)'s). Since Evans assumes that total commuting expenditure over a given distance increases less than proportionately with income, it follows that the low-skill household will receive a proportionately larger compensation for the move than the high-skill household (that is, e_Y will exceed one).

Even if Evans' empirical assertions are true,[10] however, there are three errors in his argument. First, the high- and low-skill households may commute at different speeds, so that the per-mile time cost of commuting may be different for the two classes even if the valuation of travel time (as a proportion of income) is the same (see equation (3)). If high-skill households commute faster than low-skill households, then Evans' argument understates the value of e_Y; if high-skill households commute more slowly, then Evans' argument overstates e_Y. Indeed, in the latter case it is possible that e_Y will be less than one.

Second, because population density can change, the area inhabited by low-skill households may decrease in size even if the low-skill population increases. In this case, households living at u* face <u>lower</u> commuting costs in cities with larger populations, so that the compensation for commuting costs, which is less than proportional to income, is negative, and low-skill households end up with a smaller compensation (as a proportion of their income) than the high-skill households. This case occurs when the high-skill population increases much more rapidly than the low-skill population, so that the low-skill population is outbid for housing inside the original u*.

Third, Evans' argument that there is a unitary elasticity of demand for housing with respect to total income has the implausible implication that people with a given total income spend the same amount on housing regardless of their commuting costs. A much more plausible assumption is that there is a unitary elasticity of demand for housing with respect to income net of commuting costs. This alternative assumption is equivalent to the assumption of a Cobb-Douglas utility function, as seen in equation (5) above. The difference between total and net income is important becasse Evans argues that t increases less than proportionately with total income so that Y - tu increases more than proportionately with total income. Thus, the compensation associated with housing expenditures (which depend on Y - tu) will increase more than proportionately with total income.

In summary, when du* is positive, total compensation has one component (commuting costs) that increases less than proportionately with total income and another component (housing costs) that increases more than proportionately with income, so that it cannot be determined a priori whether e_Y is greater or less than one. If du* is negative, this indeterminacy remains since the relationship between the two components of compensation is reversed; commuting costs decrease less than proportionately with income and housing costs increase less than proportionately with income. Thus, the relationship between population and the distribution of income, which depends on both e_Y and e_N, is also indeterminate a priori. We will therefore extend to a two-skill-class economy the open urban model developed above and determine more precisely how the distribution of income is related to the many parameters of the model.

The Model Extended. The mathematical model discussed in section 1 can now be extended to consider two skill classes. Clearly, each class lives in that part of the city where it outbids the other class for housing. In this model, the high-skill class will always live in the outer part of the city and the low-skill class will live in the city center.[11]

Three changes are required in order to extend the mathematics to the two-skill-class case. First, the equations[12] must be doubled and given the subscripts "l" and "h" to refer to the low- and high-skill groups, respectively. Second, the equation for the opportunity cost of housing (16)--and the equivalent equation for land--must be revised so that the high-skill class extends to the outer edge of the city, or

$$P_h(\bar{u}) = \bar{P} \tag{30}$$

and the price-distance function of the low-skill class meets the price-distance function of the high-skill class at u*, or

$$P_l(u^*) = P_h(u^*). \tag{31}$$

Substituting (30) and (31) into the price-distance function derived earlier (equation (18) with subscripts added), we have

$$V_h = (\gamma_h/\bar{P})^{\beta_h}(Y_h - t_h\bar{u})^{k_h\beta_h} \tag{32}$$

and

$$V_l = (\gamma_l/\bar{P})^{\beta_l}[(Y_h - t_h\bar{u})/(Y_h - t_h u^*)]^{k_l\beta_l}(Y_l - t_l u^*)^{k_l\beta_l}. \quad (33)$$

The third change in the one-skill-class model is to add skill-class segregation to the labor market by replacing equation (14) with

$$\int_o^{u^*} N_l(u)du = N_l \quad (34)$$

and

$$\int_{u^*}^{\bar{u}} N_h(u)du = N_h. \quad (35)$$

Following the derivation of (22), these two equations lead to

$$\begin{aligned} N_h = (\bar{R}\phi/b_h)\{ &- (Y_h - t_h\bar{u})/t_h^2 b_h(b_h + 1) - \bar{u}/t_h b_h \\ &+ (Y_h - t_h u^*)^{b_h+1}/t_h^2 b_h(b_h + 1)(Y_h - t_h\bar{u})^{b_h} \\ &+ (u^*/t_h b_h)[(Y_h - t_H u^*)/(Y_h - t_h\bar{u})]^{b_h}\} \end{aligned} \quad (36)$$

and

$$\begin{aligned} N_l = (\bar{R}\phi/t_l)[(Y_h - t_h u^*)/(Y_h - t_h\bar{u})]^{b_h}\{&Y_l^{b_l+1}/t_l(b_l + 1)(Y_l - t_l u^*)^{b_l} \\ &- (Y_l - t_l u^*)/t_l(b + 1) - u^*\}. \end{aligned} \quad (37)$$

The four equations, (32), (33), (36), and (37), can now be used to determine the relationship between population and the distribution of income. By differentiating the four equations with respect to V_l, V_h, Y_l, Y_h, N_l, N_h, u^*, and $\bar{u}$ (remembering from (3) that t_l depends on Y_l and t_h depends on Y_h); by setting $dV_l = dV_h = 0$; and by treating dN_l and dN_h as exogenous, a system of four linear equations in the four unknowns, dY_h, dY_l, du^*, and $d\bar{u}$, is obtained. The solutions to this set of equations indicate the changes in income for each class and in the areas inhabited by each class that accompany any given changes in the populations of the two classes. The solutions are too complicated to yield qualitative results, so the four equations are relegated to the Appendix and the system is analyzed numerically.

Simulation Results

Numerical analysis of the model was carried out in four steps. First, values of the parameters were chosen. As described below, several different sets of parameter values were used. Second, the values for u^* and $\bar{u}$ that result from the parameters chosen in the first step were calculated. The iterative method used for these calculations is described in the Appendix. Third, the set of equations in the Appendix was solved for dY_l, dY_h, du^*, and $d\bar{u}$, using a packaged computer program. Fourth, the effect of changes in urban population on the relative incomes of the two skill classes was determined by substituting dY_l and dY_h into the formula for e_y, equation (29). Similarly, the value of dG, which summarizes the effect of a population change on the distribution of income, was calculated using (27).

For convenience, the parameters of the model are divided into two types. The first type defines a particular city and the second type describes the market conditions that hold in all cities. The first type of parameter consists of income and population levels for the two skill classes, the rates of population growth for the two skill classes, and the number of radians of land in the city. For the initial simulations, two different divisions of a population into skill classes were examined. The first division considers the bottom 20 percent of the income distribution to be in the low-skill class and assigns an income of $15 per day to the low-skill class and an income of $64 per day to the high-skill class.[13] The second division considers the bottom half of the income distribution to belong to the low-skill class and assigns an income of $28 per day to the low-skill class and $81 per day to the high-skill class. Each of these divisions of a population into skill classes is then simulated for an area with 250,000 workers and one with one million workers. Note that the division of the population into skill classes is implied by the assumption about the income distribution; for example, 50,000 workers make up the bottom 20 percent of the income distribution in an urban area of 250,000 workers.

All four combinations of an income distribution and a total population are then simulated for a circular area of 2π and a semi-circular area of π radians of land. Finally, each of the eight areas is simulated for three different assumptions about the growth rates of the two skill classes. The first assumption is that both groups grow at a 10 percent rate, the second is that the low-skill class grows at a 5 percent rate and the high skill class at a 15 percent rate, and the third is that the low-skill class grows at a 15 percent rate and the high-skill class at a 5 percent rate.

Parameters of the second type describe commuting costs and the demand and supply conditions in the four markets. The values chosen for these parameters are similar to the values chosen by Mills (1972à) and Solow (1973) for similar simulation models. The basic set of parameters consists of "best guesses" about the actual values of these parameters in a typical American metropolitan area. The basic set of parameters is as follows: agricultural rental rate ($\bar{R}$) = 1500 per square mile per day; share of land in the production of housing (a) = .20; inverse of the proportion of net income spent on housing for both skill classes (k_1 and k_h) = 4.0; per mile operating costs (c_1^o and c_h^o) = .10; commuting speed (MPH_1 and MPH_h) = 20 miles per hour; valuation of travel time (w_1 and w_h) = .5 times the wage rate.

Given this basic set of parameters, four sets of simulations were performed.

Basic Simulations. This set of simulations examines the values of e_Y and dG generated by a basic set of parameters. Thus, this set consists of simulations of eight basic urban areas (each with three different patterns of population growth rates) using the basic set of market parameters.

Sensitivity Simulations. This set calculates the effects on e_Y and dG of changes in the values of the commuting and market parameters.

Asymmetrical Simulations. The asymmetrical simulations determine the effects on e_Y and dG of different assumptions for the two skill classes about commuting costs and the proportion of net income spent on housing.

Actual City Simulations. The final set of simulations calculates the values of e_Y and dG using values of the parameters Y_1, Y_h, N_1, N_h,

dN_l, dN_h and $\bar{\phi}$ for actual urban areas. These simulated values are then compared with the actual values of e_Y and dG in the urban areas. Each set of simulations will be described in turn.

Basic Simulations. The relationship between urban population and the distribution of income cannot be determined a priori. The value of e_Y can be greater or less than one depending on the sign of du* and the relative sizes of the transportation cost and housing cost components of compensation. The results in Table 3 show that, within the range of parameters used for the basic simulation, the transportation cost component dominates the housing cost component, and that this dominance increases with the absolute value of du*. If $dN_l/N_l = dN_h/N_h = .10$, du* is positive and e_Y is slightly greater than one. If $dN_l/N_l = .05$ and $dN_h/N_h = .15$, the low-skill class is outbid for housing inside u* by the rapidly growing high-skill class, so that du* is negative and e_Y is less than one. Finally, if $dN_l/N_l = .15$ and $dN_h/N_h = .05$, then du* is large and positive, the transportation component of compensation is much larger than the housing component, and e_Y is considerably greater than one.

Several other characteristics of the model are suggested by Table 3. First, by comparing areas of 250,000 workers with the corresponding areas of 1,000,000 workers (such as cases I.A.1 and I.B.1) it can be seen that e_Y decreases slightly as total population increases. Furthermore, Table 1 indicates that e_Y decreases somewhat as the number of radians in an area decreases. However, both these effects are small and are not always true in other simulations (not shown).

The simulations in panels I and II of Table 3 cannot separate the effects on e_Y of Y_l/Y_h and N_l/N_h. Therefore, two additional simulations

Table 3

Basic Simulations

Description of Urban Area	Population Growth Rates								
	$dN_l/N_l = dN_l/N_h = .10$			$dN_l/N_l = .05, dN_h/N_h = .15$			$dN_l/N_l = .15, dN_h/N_h = .05$		
	e_Y	du*	dG/G	e_Y	du*	dG/G	e_Y	du*	dG/G
I. Low-Skill Class is Bottom 20 Percent of Income Distribution ($Y_l=15$, $Y_h=64$)									
A. 250,000 Total Workers ($N_l=50{,}000$; $N_h=200{,}000$)									
1. Radians: $\phi=2\pi$	1.075	.025	.000	.927	−.025	−.066	1.468	.075	.069
2. Radians: $\phi=\pi$	1.050	.021	.000	.906	−.038	−.066	1.433	.081	.069
B. 1,000,000 Total Workers ($N_l=200{,}000$; $N_h=800{,}000$)									
1. Radians: $\phi=2\pi$	1.029	.017	.000	.889	−.051	−.066	1.403	.084	.069
2. Radians: $\phi=\pi$	1.014	.012	.000	.877	−.062	−.066	1.381	.087	.069
II. Low-Skill Class is Bottom Half of Income Distribution ($Y_l=28$, $Y_h=81$)									
A. 250,000 Total Workers ($N_l=N_h=125{,}000$)									
1. Radians: $\phi=2\pi$	1.109	.071	.000	.927	−.043	−.023	1.432	.186	.020
2. Radians: $\phi=\pi$	1.077	.064	.000	.895	−.076	−.023	1.406	.203	.020
B. 1,000,000 Total Workers ($N_l=N_h=500{,}000$)									
1. Radians: $\phi=2\pi$	1.049	.052	.000	.867	−.111	−.023	1.381	.215	.020
2. Radians: $\phi=\pi$	1.026	.039	.000	.845	−.144	−.023	1.360	.221	.020
III. Change Income Distribution Hold Population Distribution Constant ($Y_l=28$, $Y_h=81$)									
A. 250,000 Total Workers ($N_l=50{,}000$; $N_h=200{,}000$)									
1. Radians: $\phi=2\pi$	1.041	.031	.000	.948	−.040	−.064	1.276	.103	.066
IV. Change Population Distribution, Hold Income Distribution Constant ($Y_l=15$, $Y_h=64$)									
A. 250,000 Total Workers ($N_l=125{,}000$; $N_h=125{,}000$)									
1. Radians: $\phi=2\pi$	1.183	.053	.000	.900	−.027	−.029	1.774	.132	.026

are carried out using the basic set of commuting and market parameters. First, the simulation in line I.A.1 is repeated, varying only the income distribution--that is, the ratio Y_l/Y_h. The results of this simulation, shown in Panel III, indicate that an increase in income equality, holding the population distribution constant, moves e_Y closer to unity. This result has strong intuitive appeal, since a perfectly equal income distribution ($Y_l = Y_h$) represents a one-skill-class model where, by definition, e_Y is equal to unity. Second, the simulation in line I.A.1 is repeated, varying only the population distribution--that is, the ratio N_l/N_h. As shown in panel lV, this simulation indicates that an increase in the proportion of the population that is low-skill, holding income distribution constant, increases e_Y. This result is also not surprising; the increase in N_l/N_h is equivalent to an increase in the absolute change in low-skill workers and therefore should have the same effect as an increase in dN_l/N_l, which (as seen earlier) increases e_Y.

The values of dG associated with this first set of simulations are given in the last column of Table 3. Before examining these results, it should be remembered that population changes have a direct effect on dG, as shown in the first term in equation (27), and an indirect effect through changes in incomes of the two skill classes. Furthermore, the direct and the indirect effect typically work in opposite directions. For example, a value of e_N that is greater than unity leads, in general, to a value of e_Y that is also greater than one so that, according to equation (27) the sign of dG is indeterminate. However, the results in the last column of Table 3 are unambiguous; the sign of dG is always determined by the direct effect of the population elasticity. If e_N is greater than one, dG is positive; if e_N is less than one, dG is negative; and if e_N is equal to one, dG is essentially zero. The results are true for all simulations; the direct effect of e_N in equation (27) always swamps the indirect effect of e_N that operates through incomes. Changes in relative incomes in an open urban model are quite small and have virtually no impact on the Gini coefficient. This result does not depend on the magnitudes of population changes. If N_1 and N_h change by several hundred percent, rather than by the 5, 10 or 15 percent shown in Table 3, the direct effect of e_N still dominates the indirect effect that operates through changes in incomes. Since e_Y is very close to one when e_N equals one, the indirect effect is very small even when the direct effect is zero.[14]

Sensitivity Simulations. The second set of simulations is summarized in Table 4. These simulations take the first case in Table 3 (line I.A.1) and vary the market and transportation parameters to determine the sensitivity of the results to changes in the second type of parameter. Table 4 indicates that in the range of parameters examined, an increase in the agricultural rental rate ($\bar{R}$), an increase in the share of land in the production of housing ($\bar{a}$), and an increase in the proportion of income spent on housing (1/k), all increase the value of e_Y. Simulations of the other cases from Table 3 lead to the same conclusions.

Table 4

Sensitivity Simulations

	Population Growth Rates								
	$dN_1/N_1=dN_h/N_h=.10$			$dN_1/N_1=.05$, $dN_h/N_h=.15$			$dN_1/N_1=.15$, $dN_h/N_h=.05$		
Description of Urban Area	e_Y	du*	dG/G	e_Y	du*	dG/G	e_Y	du*	dG/G
V. First Case (I.A.1 from Table 1)									
A. Basic Market Parameters ($\bar{R}=1500$, $a=.2$, $k_h=k_1=4$, $t_h^0=t_1^0=.10$, $W_h=W_1=.5$, $MPH_h=MPH_1=20$)	1.075	.025	.000	.927	-.025	-.066	1.468	.075	.069
B. Change Agricultural Rental Rate									
1. $\bar{R}=1000$	1.060	.023	.000	.914	-.032	-.066	1.447	.078	.069
2. $\bar{R}=2000$	1.086	.026	.000	.936	-.020	-.066	1.484	.072	.069
C. Change Share of Land in Housing									
1. $a=.10$	1.057	.011	.000	.914	-.019	-.066	1.438	.041	.069
2. $a=.30$	1.085	.039	.000	.933	-.027	-.066	1.488	.104	.069
D. Change Proportion of Net Income Spent on Housing									
1. $k_h=k_1=3.0$	1.082	.034	.000	.932	-.026	-.060	1.482	.095	.069
2. $k_h=k_1=6.0$	1.064	.016	.000	.919	-.022	-.066	1.450	.053	.069
E. Change Travel Costs									
1. Low Costs ($t_1^0=t_h^0=.05$, $W_1=W_h=.25$, $MPH_1=MPH_h=30$)	1.193	.064	.000	.992	.001	-.066	1.724	.127	.069
2. High Costs ($t_1^0=t_h^0=.20$, $W_1=W_h=.75$, $MPH_1=MPH_h=10$)	1.009	.004	.000	.902	-.024	-.066	1.293	.033	.069
VI. Other Cases									
A. Case II.A.1 from Table 1 with Low Travel Costs	1.247	.162	.000	1.019	.024	-.023	1.642	.300	.020
B. Case I.B.1 from Table 1 with High Travel Costs	.994	.002	.000	.890	-.030	-.066	1.272	.034	.069
C. Case II.B.1 from Table 1 with High Travel Costs	.996	.006	.000	.856	-.074	-.023	1.251	.086	.020

Transportation costs effect e_Y in two ways. First, higher transportation costs lead to a more compact city and to smaller values for du*. Thus, the transportation component of compensation becomes relatively less important and e_Y falls. Second, changes in the elements of transportation costs change the ratio of total transportation costs from u* to income and thereby affect e_Y. In the simulation, in rows V.E.1 and V.E.2 of Table 4, it can be seen that the value of e_Y decreases as transportation costs increase. Since the ratio of total transportation costs from u* to income increases faster in the low-cost case than in the high-cost case, it follows that the first of the two effects mentioned above is stronger in these simulations.

The transportation cost simulations bring out several other characteristics of the model. In case V.E.1, du* is positive even when dN_1/N_1 is less than dN_h/N_h. This result illustrates how low transportation costs lead to areas that grow rapidly in area as population rises. The same case also illustrates that the housing component of compensation can be larger than the transportation cost component. Case VI.B., with $dN_1/N_1 = dN_h/N_h$, illustrates the same result: $e_Y > 1$ even though du* is positive.

The simulations in Table 4 tell the same story about dG as our earlier simulations: dG is positive when e_N is greater than one, negative when e_N is less than one, and essentially zero when e_N is equal to one.

<u>Asymmetrical Simulations</u>. The third set of simulations, which is presented in Table 5, examines the effects of differences between the two skill groups in the transportaion and housing parameters.

Table 5

Asymmetrical Simulations

	Population Growth Rates								
	$dN_l/N_l=dN_h/N_h=.10$			$dN_l/N_l=.05, dN_h/N_h=.15$			$dN_l/N_l=.15, dN_h/N_h=.05$		
Description of Urban Area	e_Y	du*	dG/G	e_Y	du*	dG/G	e_Y	du*	dG/G
VII. First Case (I.A.1 from Table 1)									
A. Same Parameters for Both Skill Classes ($t_l^0=t_h^0=.10$, $W_h=W_l=.5$, $MPH_l=MPH_h=20$, $k_h=k_l=4$)	1.075	.025	.000	.927	−.066	−.010	1.468	.075	.069
B. High-Skill Commute Faster ($MPH_h=30$, $MPH_l=10$)	1.205	.030	.000	.891	−.066	−.010	2.054	.078	.069
C. High-Skill Pay Higher Operating Costs and Value Travel Time More ($t_h^0=.20$, $t_l^0=.10$, $W_h=.75$, $W_l=.25$)	.991	.004	.000	1.024	−.066	−.010	.905	.038	.069
D. Low-Skill Spend Higher Proportion of Income on Housing ($k_l=3$, $k_h=6$)	1.983	.023	−.001	1.958	−.067	−.010	2.041	.084	.069
E. High-Skill Spend Higher Proportion of Income on Housing ($k_l=6$, $k_h=3$)	.590	.023	.001	.451	−.065	−.010	.983	.061	.069

If the high-skill class commutes faster than the low-skill class, then, as shown in row VII.B, the effect of the transportation cost component of compensation is magnified and the values of e_Y are farther from one than in the symmetrical case. If, on the other hand, the high-skill class has higher operating costs (because they drive luxury cars while the low-skill class drives compacts) and if the high-skill class values travel time more, then transportation costs may, as illustrated in case VII.C, increase more than proportionately with income. In this case, both components of compensation imply that e_Y will be less than one whenever du* is positive, and greater than one whenever du* is negative.

Finally, the effects on e_Y of differences between the two skill classes in the proportion of income spent on housing is determined. These differences are similar in their effect to the assumption that the elasticity of demand for housing with respect to net income is not equal to unity. If the low-skill class spends a higher proportion of its income on housing than the high-skill class (that is, $k_h > k_l$), then, as shown in row VII.D, the value of e_Y is always much greater than one--even when du* is negative. This result obtains because housing expenditures now increase less than proportionately with income. Conversely, if the high-skill class spends a higher proportion of its income on housing than the low-skill class, then the values of e_Y are significantly less than one. Even when dN_h/N_l is greater than dN_h/N_h, the value of e_Y may be less than one, as shown in Table 5.

Despite the rather dramatic effects on e_Y of the simulations in Table 5, conclusions about dG are not altered: the sign of dG is still determined by e_N. Even when e_N is equal to one, so that there is no direct effect of population changes on G, the large values of e_Y have little effect on dG. For example, for both cases VII.D and VII.E in Table 5, with $dN_l/N_l = dN_h/N_h = .10$, the value of dG is less than 0.1 percent of G. However, if the changes in the area's population were several hundred percent, the area's value of e_Y in these cases would lead to a significant increase in G.

Simulations of Actual Cities. The fourth set of simulations takes Y_l, Y_h, N_l, N_h, dN_l, and dN_h and ϕ from 17 actual SMSAs using 1960 and 1970 Census data. The values of e_Y for each urban area were then calculated using the basic set of market parameters.[15] The results are presented in Table 6. The simulated values of e_Y in Table 6 can be compared with the actual values of e_Y--that is, with the values obtained using dY_l and dY_h from the Census data. In nine of the urban areas, the difference between the simulated and the actual is less than .25. However, the difference in some areas is enormous. In Des Moines, for example, the model estimates that $e_Y = 2.116$, whereas the actual value is .963.

The divergence between the simulated and the actual values of e_Y suggests some ex post hypotheses. In the sun belt cities (Phoenix, Atlanta, Oklahoma City, Houston, San Jose, Nashville, Richmond and Fresno), the actual elasticity is greater, and often much greater, than the simulated elasticity. In midwestern cities (Milwaukee, Minneapolis, Indianapolis, Omaha, Des Moines, and Flint), the actual elasticity is much less than the simulated elasticity. In fact, in every Midwestern city except Flint, the actual elasticity is less than one and the simulated elasticity is greater than one. This result can be explained by two assumptions: (1) the high-skill class is

Table 6

Simulations for Actual Urban Areas

Urban Area	Y_1	Y_h	N_1	N_h	dN_1/N_1	dN_h/N_h	ϕ	e_Y	e_Y^*	du*	dG/G	dG*/G
Phoenix	10.8	27.1	33,000	129,688	.259	.424	4.97	1.001	1.736	.013	-.085	-.224
Atlanta	10.6	27.4	35,615	216,168	.304	.423	5.59	.994	1.447	.006	-.068	-.165
Denver	13.2	27.9	32,833	200,603	.396	.293	6.28	1.081	1.000	.093	.061	.062
Houston	12.3	27.8	48,841	275,323	.510	.601	6.28	1.023	1.262	.067	-.045	-.108
Milwaukee	16.2	27.3	37,507	284,239	.358	.075	3.49	1.260	.900	.171	.201	.260
Minn.-St. Paul	14.7	28.4	48,802	329,237	.364	.201	5.59	1.110	.955	.120	.104	.123
Baltimore	13.6	25.5	67,286	370,901	.195	.143	5.32	1.069	1.160	.055	.034	-.034
Oklahoma City	11.1	25.2	16,013	113,649	.368	.211	5.50	1.169	1.221	.093	.103	.040
Hartford	14.2	28.6	17,033	126,501	.314	.197	5.59	1.116	1.114	.091	.078	.034
San Jose	14.7	33.5	22,092	139,371	.678	.666	6.28	1.057	1.296	.144	.005	-.073
Nashville	10.1	24.1	14,915	82,612	.328	.368	6.28	1.061	1.560	.047	-.023	-.148
Indianapolis	13.7	28.0	23,861	158,349	.576	.518	6.28	1.056	.986	.116	.030	-.053
Richmond	11.1	26.2	14,244	89,613	.233	.243	4.97	1.071	1.740	.044	-.007	-.166
Omaha	13.8	26.7	16,722	97,621	.236	.078	6.28	1.295	.944	.090	.109	.139
Des Moines	13.7	27.6	8,986	60,386	.285	.013	6.28	2.116	.963	.093	.202	.218
Flint	14.9	24.2	10,048	86,850	.583	.200	6.28	1.169	1.008	.168	.248	.243
Fresno	11.6	26.2	24,081	67,912	-.006	.036	4.89	.815	1.460	-.013	-.025	-.138

much more mobile than the low-skill class, and (2) sunbelt cities are strongly preferred to other cities, particularly to Midwestern cities. These assumptions imply that high-skill households will move to sunbelt cities without receiving the monetary compensation predicted by our model, so that our model will underestimate e_Y. Similarly, high-skill households will only move to Midwestern cities if they receive more compensation than predicted by the model, so that the model overestimates e_Y. This hypothesis receives some tentative confirmation in our data. In every sunbelt city except San Jose and Oklahoma City, the high-skill population grew at a faster rate than the low-skill population; in every Midwestern city, the low-skill population grew at a faster rate than the high-skill population.

As before, the wide variation in e_Y in these simulations does not affect conclusions about dG. In every one of these simulations of an actual city, the sign of dG is determined by the direct effect of e_N.

A clear conclusion emerges from these simulations: the relationship between urban population and the distribution of income is a complex one, even in this simple two-skill-class model. Evans was incorrect in arguing that the incomes of the low-skill class would increase proportionately more than those of the high-skill class as urban size increased. Nevertheless, this section has shown that several statements can be made about the relationship between urban population and the distribution of income within an urban area in the range of parameters examined here. In the next section, an attempt is made to test these statements, using data for a sample of metropolitan areas.

Estimation Results

In the simulations presented above, the value of e_Y depends primarily on the value of e_N, but is also influenced by the values of the many parameters in the model. Two main difficulties prevent precise tests of these simulation results: first, data are not available on most of the parameters for individual urban areas, and second, the functional forms relating the various parameters and e_N to e_Y are not known.

It is possible, however, to perform a simple, but not very powerful, test of the main simulation result by performing the following regression:

$$\log(Y_1/Y_h) = a_o + a_1\log(N_1) + a_2\log(N_h). \tag{38}$$

In this regression--which is a straightforward extension of the regressions reported for the one-skill class model--the coefficients can easily be related to the values of e_Y and e_N. Differentiating (38) yields

$$d\log(Y_1/Y_h) = a_1 d\log(N_1) + a_2 d\log(N_h)$$

or

$$(dY_1/Y_1) - (dY_h/Y_h) = a_1(dN_1/N_1) + a_2(dN_h/N_h) \tag{39}$$

The left-hand side of this regression is clearly positive if e_Y is greater than unity and negative if e_Y is less than unity, so that it conveys the same qualitative information as e_Y. According to our simulations, e_Y increases with (dN_1/N_1) and decreases with (dN_h/N_h) so a_1 should be positive and a_2 should be negative. Furthermore, e_Y is almost always slightly greater than unity if (dN_1/N_1) is equal to (dN_h/N_h) so that the coefficient of (dN_1/N_1) should be slightly greater than the coefficient of (dN_h/N_h).

Equation (30) was estimated for a sample of 89 SMSAs using the skill class divisions described in footnote 15.[16] Regional dummy variables were included in the regressions to capture the amenity affects discussed above. The results are presented in Table 7. In both the 1960 and the 1970 regressions, a_1 and a_2 have the expected signs, a_1 is slightly greater than a_2, and the regional dummy variables are highly significant. Furthermore, a_1 and a_2 are highly significant in the 1970 regression, although they are not quite significant at the 10 percent level in the 1960 regression. Thus these regression results are consistent with the simulation model.

A further test of the simulation model can be made by comparing the simulated values of $((dY_1/Y_1) - (dY_h/Y_h))$ with the values of that quantity predicted by the regressions. A simulated value is obtained using the values of Y_1, Y_h, dY_1, and dY_h from a particular simulation and the comparable predicted value is obtained by substituting the values of (dN_1/N_1) and (dN_h/N_h) from that simulation and the values of a_1 and a_2 from the 1970 regression into equation (39). Several such comparisons are presented in Table 8. The striking result in Table 8 is that the regressions predict a much larger difference in $[(dY_1/Y_1) - (dY_h/Y_h)]$ due to changes in $\log(N_1)$ and $\log(N_h)$ than do the simulations. In the first row of the table, the predicted difference (.0533 percent) is about 2.5 times as large as the simulated difference (.0222 percent); in the second row, the predicted difference is about 40 times the simulated difference. Even in the most extreme simulations--when k_h and k_1 differ--the simulated difference is, as shown in the last two lines of Table 8, a fraction of the predicted response.

Table 7

Regression Estimates for Two-Skill-Class Model

	$\log(Y_1/Y_h)$ 1960	$\log(Y_1/Y_h)$ 1970
Constant	-.791 (4.87)	-.531 (3.99)
$\log(N_1)$	.058 (1.11)	.137 (2.74)
$\log(N_h)$	-.051 (0.99)	-.132 (2.67)
Northeast	.223 (7.65)	.136 (6.33)
Northcentral	.200 (6.83)	.108 (4.97)
West	.082 (2.74)	.014 (0.59)
R^2	.492	.391

Note: Number of observations is 89; t-statistics appear in parentheses below regression coefficients.

Table 8

Comparison of Simulation and Regression Results

Case	dN_l/N_l	dN_h/N_h	$(dY_l/Y_l - dY_h/Y_h)$x100	
			Simulated	Predicted
I.A.1	.10	.10	.0222	.0533
	.05	.15	-.0315	-1.2944
	.15	.05	.0758	1.4010
I.B.1	.10	.10	.0113	.0533
	.05	.15	-.0623	-1.2944
	.15	.05	.0849	1.4010
VII. D	.15	.05	.1383	1.4010
VII. E	.05	.15	-.2922	-1.2944

Note: The cases refer to the Simulations of Table 3 (I.A.1 and I.B.1) and Table 5 (VII.D and VII.E).

Thus the simulations severely underestimate the variation in $((dY_l/Y_l) - (dY_h/Y_h))$ and, therefore, underestimate the variation in inequality across urban areas. It is hoped that future research on this topic will increase understanding of this variation.

APPENDIX

Deriving the Equations

By totally differentiating equations (32), (33), (36), and (37) with respect to V_l, V_h, Y_l, Y_h, N , N_h, u^*, and $\bar{u}$ (remembering from (3) that t_l depends on Y_l and t_h depends on Y_h); by setting $dV_l = dV_h = 0$; by treating dN_l and dN_h as exogenous; and by rearranging terms the following system of equations is obtained:

$$dY_h - [t_h/(1 - t_h^y\bar{u})]d\bar{u} = 0, \tag{A1}$$

$$\begin{aligned} & dY_l[k_l(1 - t_l^y u^*)/(Y_l - t_l u^*)] + dY_h \\ & \cdot \{k_h[(1 - t_h^y\bar{u})/(Y_h - t_h\bar{u}) - (1 - t_h^y u^*) \\ & /(Y_h - t_h u^*)]\} + du^*[k_h t_h/(Y_h - t_h u^*) \\ & - k_l t_l/(Y_l - t_l u^*)] + d\bar{u}[-k_h t_h/(Y_h - t_h\bar{u})] = 0 \end{aligned} \tag{A2}$$

$$\begin{aligned} dN_h = {} & \bar{R}\phi b_h\{dY_h(1/t_h)\} - (1 - t_h^y\bar{u})/t_h b_h(b_h + 1) \\ & - [(1 - t_h^y\bar{u})/t_h(b_h + 1)][(Y_h - t_h u^*) \\ & /(Y_h - t_h\bar{u})]^{b_h+1} + [(1 - t_h^y u^*)/t_h b_h][(Y_h - t_h u^*) \\ & /(Y_h - t_h\bar{u})^{b_h}] + u^*(1 - t_h^y u^*)(Y_h - t_h u^*)^{b_h-1} \\ & /(Y_h - t_h\bar{u})^{b_h} - u^*(1 - t_h^y\bar{u})(Y_h - t_h u^*)^{b_h} \\ & /(Y_h - t_h\bar{u})^{b_h+1}\} + d\bar{u}\{-1/t_h(b_h + 1) + [(Y_h - t_h u^*) \\ & /(Y_h - t_h\bar{u})]^{b_h+1}/t_h(b_h+1) + u^*(Y_h - t_h u^*)^{b_h} \\ & /(Y_h - t_h\bar{u})^{b_h+1}\} + du^*[-u^*(Y_h - t_h u^*)^{b_h-1}/(Y_h - t_h\bar{u})^{b_h}] \end{aligned} \tag{A3}$$

$$dN_l = dY_h(\bar{R}\phi Zb_h/t_l)\{(1 - t_h^y u^*)(Y_h - t_h u)^{b_h - 1}/(Y_h - t_h\bar{u})^{b_h} - (1 - t_h^y\bar{u})(Y_h - t_h u^*)^{b_h}/(Y_h - t_h\bar{u})^{b_h+1}\} + dY_l(\bar{R}\phi X/t_l^2)\{(Y_l/(Y_l - t_l u^*)^{b_l} - [b_l(- t_l^y u^*)/(b_l + 1)][Y_l/(Y_l - t_l u^*)]^{b_l+1} - (1 - t_l^y u^*)/(b_l + 1)\} + du^*(\bar{R}\phi X/t_l)\{[b_l/(b_l + 1)][Y_l - t_l u^*)]^{b_l+1} - b_l/(b_l + 1)\} - (\bar{R}\phi Z/t_l)\{b_h t_h(Y_h - t_h u^*)^{b_h - 1}/(Y_h - t_h\bar{u})^{b_h}\} + d\bar{u}(\bar{R}\phi Z/t_l)\{b_h t_h(Y_h - t_h u^*)^{b_h}/(Y_h - t_h\bar{u})^{b_h+1}\} \quad (A4)$$

where $Z = Y_l^{b_l+1}/(Y_l - t_l u^*)^{b_l} t_l^2 b_l(b_l + 1) - (Y_l - t_l u^*)/t_l^2 b_l(b_l + 1) - u^*/t_l b_l$, and

$X = [(Y_h - t_h u^*)/(Y_h - t_h\bar{u})]^{b_h}$.

Solving for u^* and $\bar{u}$

After the parameters are chosen, the next step in solving the model is to determine u^* and $\bar{u}$. The values of u^* and $\bar{u}$ are found using equations (36) and (37) and the following iterative procedure:

1. Define $N = N_l + N_h$, set $u^* = 0$, and solve (36) numerically for $\bar{u}$ by increasing the value of $\bar{u}$ by small increments from zero until

$$\int_{u^*}^{\bar{u}} N_h(u)du > N.$$

Note that the value of $\bar{u}$ calculated in this step will always exceed the final value of $\bar{u}$ since the high-skill class lives in less densely settled areas than the low-skill class.

2. Solve (37) numerically for u^* using the $\bar{u}$ calculated in step 1.

3. Solve (36) numerically for $\bar{u}$ using the u^* calculated in step 2.

4. Solve (37) numerically for u^* using the $\bar{u}$ calculated in step 3.

5. Repeat steps 3 and 4 until the change in both u^* and $\bar{u}$ from one calculation to the next is less than some specified value (.001).

Although it has not been proven that this iterative procedure has any desirable mathematical properties, it converged rapidly in all simulations.

NOTES

[1]The assumption that some members of every skill class live in every city is an assumption about the demand for labor. Although the demand for labor is not considered in much detail in this paper, it should be pointed out that our model is based on the implicit assumption that economies of scale and external economies allow employers to pay higher wages in larger cities and still compete with firms in smaller cities that pay lower wages. In addition, we assume that all income is received as wages.

[2]The assumption that P_Z is the same in all cities is important. If, for example, large cities have amenities (such as theatres or parks) that are not present in smaller cities, then P_Z will be lower in large cities and our model will overestimate the compensation required to induce people to live in large cities. This problem can, in principle, be included in a model like ours, but for simplicity is not considered here.

[3]Two comments should be made on equation (3). First, it may vary across cities due to different public transportation systems or different fuel costs. This complication, like that of amenities, is not considered here. Second, the expression for operating costs is equal to c_o times working days per year (250) times two (because we are considering round-trip commuting costs). The expression for time costs is equal to working days per year times minutes per mile of commuting times the dollar value of a minute spent commuting times two. The dollar value of a minute spent commuting is w times Y divided by minutes worked per year (for 250 eight-hour days).

[4] Since the model assumes that all income is derived from labor supplied at the CBD by a single worker per household, male wages are the appropriate variable for the estimation.

[5] The Gini coefficient was chosen as the measure of income inequality because of its widespread use in the literature. Any other summary statistic could have been chosen, but as is shown below, the Gini can be conveniently decomposed when there are only two income classes.

[6] Bhattacharya and Mahalanobis (1967) present a decomposition of the Gini coefficient into a between-group variance, a within-group variance and an interaction term. In the one-skill-class case, in which all individuals in a given skill class have identical incomes, these last two components are zero, so the Gini coefficient reduces to:

$$G = \sum_{i \neq j} P_i P_j |u_i - u_j| / 2u$$

where P_i is the proportion of the population in each group, u_i the group mean income, and u the mean income for the entire population. In the two-skill-class case,

$$G = \frac{2(Y_h - Y_l)[N_l/(N_l + N_h)][N_h/(N_l + N_h)]}{2[(N_l Y_l + N_h Y_h)/(N_l + N_h)]}$$

$$= [N_l/(N_l + N_h)][N_h Y_h/(N_l Y_l + N_h Y_h)] - [N_h/(N_l + N_h)][N_l Y_l/(N_l Y_l + N_h Y_h)]$$

$$= [N_l/(N_l + N_h)]]N_h Y_h/(N_l Y_l + N_h Y_h)] - \{1 - [N_l/(N_l + N_h)]\}[N_l Y_l/(N_l Y_h + N_h Y_h)]$$

$$= [N_1/(N_1 + N_h)]\{[N_hY_h/(N_1Y_1 + N_hY_h)] + [N_1Y_1/(N_1Y_1 + N_hY_h)]\}$$

$$= [N_1/(N_1 + N_h)] - [N_1Y_1/(N_1Y_1 + N_hY_h)].$$

[7] These elasticities do not actually appear in equation (27). However, e_N conveys the same qualitative information as $(N_h dN_1 - N_1 dN_h)$ and e_Y conveys the same qualitative information as $(Y_h dY_1 - Y_1 dY_h)$. Thus e_N is equal to one when $(N_h dN_1 - N_1 dN_h)$ is equal to zero, greater than one when $(N_h dN_1 - N_1 dN_h)$ is positive, and less than one when $(N_h dN_1 - N_1 dN_h)$ is negative. For ease of exposition, the discussion will focus on the elasticities instead of the more cumbersome formulas in equation (27).

[8] Assuming that $N_h > N_1$ (Y_h is greater than Y_1 by definition), then the change in Gini is related to the elasticities as follows:

	$e_N < 1$	$e_N = 1$	$e_N > 1$
$e_Y < 1$	?	+	+
$e_Y = 1$	-	0	+
$e_Y > 1$	-	-	?

[9] At this point it is appropriate to assume that the two skill classes are completely segregated so that there is a border between them. However, this result can be derived from the model instead of assumed. (See also Solow 1972).

[10] It should be pointed out that Evans' empirical assertions are open to some question. For example, Beesley (1965) found that high-income commuters value their travel time more highly than low-income commuters, and many studies (some of which are reviewed

in DeLeeuw (1971) have found an income elasticity of demand for housing that is not equal to one. However, these assumptions are convenient and not unreasonable simplifications, and will be used throughout this paper.

[11]This result obtains because the price-distance function for the high-skill class is flatter than the price-distance function for the low-skill class so that high-skill households outbid low-skill households for housing far from the CBD. For proofs of this income-sorting result in similar models, see Mills (1972b) or Muth (1969).

[12]The equations are (4) and (6) through (14).

[13]The mean daily income for families in the bottom 20 percent of the income distribution in American cities in 1971 was about \$15 and the mean daily income for families in the top 80 percent of the income distribution was about \$64. Similarly, the mean daily income for families in the bottom half of the income distribution was about \$28 and for families in the top half of the income distribution was about \$81 (U.S. Bureau of the Census 1973).

[14]It should be pointed out that the focus on the elasticity e_Y is somewhat misleading here. As pointed out in note 7, the precise measure of the impact of changes in relative incomes on G depends (unlike e_Y) on the absolute changes in Y_l and Y_h. Since, as shown in section 1, these absolute changes are fairly small (on the order of 0.03 - 0.06 percent), changes in relative income do not affect the Gini even when e_Y is considerably greater than one.

[15] In order to compute e_Y for the sample of SMSAs it was necessary to divide the employed population into two skill classes. Laborers, farm laborers, and service workers were assumed to be in the low-skill class, while all other occupations were assumed to be in the high-skill class. Since the model assumes only one worker per household, employed males were the units chosen for analysis. First, the mean incomes of the two groups and the number employed in each group were calculated for each SMSA for 1960 and 1970 (U.S. Bureau of the Census 1960, Table 124; U.S. Bureau of the Census 1970, Table 175). The low-skill group is about 20 percent of the size of the high-skill group, and has average wages about one-half the level of the high-skill group's average. Once these levels were computed, e_Y, e_N, and dG were known from equations (27), (20) and (29). An estimate of radians of available land was, then, derived by using the urbanized area maps provided by the Census. Finally, these actual values of Y_1, Y_h, N_1, N_h, and ϕ, and the basic market parameters (assumed constant for all areas), $\bar{R}$, a, K_1, K_h, t_1^o, t_h^o, w_h, w_1, MPH_h, and MPH_1, were used in the computer simulation. The simulation produced the estimated elasticity, e_Y, change in skill-class boundary, u*, and change in Gini coefficient, dG/G, shown in Table 6.

[16] The division of workers into skill classes is arbitrary and these results could be replicated with other divisions.

References

Alonso, W. 1971. "The Economics of Urban Size." Papers and Proceedings of the Regional Science Association 26: 67-83.

Beesley, M.E. 1965. "The Value of Time Spent Travelling; Some New Evidence." Economica 37: 174-185.

Bhattacharya N. and Mahalanobis, B. 1967. "Regional Disparities in Household Consumption in India." Journal of the American Statistical Association 62: 143-161.

De Leeuw, F. 1971. "The Demand for Housing. A Review of Cross-Section Evidence." Review of Economics and Statistics, 53: 1-8.

Evans, A. 1972. "The Pure Theory of City Size in an Industrial Economy." Urban Studies 9: 49-77.

Hoch, I. 1972. "Income and City Size." Urban Studies 9: 299-328.

Mansfield, F. 1949. "City Size and Income." in Studies in Income and Wealth, Vol. 21. New York: National Bureau of Economic Research.

Mills, E. 1967. "An Aggregative Model of Resource Allocations in a Metropolitan Area." American Economic Review Papers and Proceedings 57:197-210.

Mills, E. 1972a. Studies in the Structure of the Urban Economy. Baltimore: Johns Hopkins University Press.

Mills, E. 1972b. Urban Economics. Glenview, Ill.: Scott Foresman.

Muth, R. 1969. Cities and Housing. Chicago: University of Chicago Press.

Richardson, H. 1973. The Economics of Urban Size. Lexington, Mass.: D.C. Heath.

Solow, R. 1973. "Congestion Cost and the use of Land for Streets." Bell Journal of Economics and Management Science 4: 602-618.

Solow, R. 1972. "On Equilibrium Models of Urban Location." In J. M. Parkin, ed., Essays in Modern Economics. London: Longmans,

U.S. Bureau of the Census. 1973. Money Income in 1971 of Families and Persons in the United States. Current Population Reports, Series P-60, No. 85. Washington, D.C.: Government Printing Office.