Journal of Urban Economics **47,** 165–179 (2000)
doi:10.1006/juec.1999.2137, available online at http://www.idealibrary.com on IDEAL®

The Panexponential Monocentric Model[1]

Alex Anas

Department of Economics State University of New York at Buffalo, Amherst, New York 14260

Richard J. Arnott

Department of Economics, Boston College, Chestnut Hill, Massachusetts 02467

and

Kenneth A. Small

Department of Economics, University of California at Irvine, Irvine, California 92717

Received October 23, 1998; revised April 14, 1999

We derive declining exponential rent and density functions for a monocentric city from a new set of assumptions, which place restrictions on commuting costs rather than on the demand for land. The utility function is Cobb–Douglas with unrestricted income expenditure shares for land and for the numeraire good. The marginal commuting cost is assumed to be proportional to income-earning potential and exponentially declining in distance from the center at a particular constant rate. These assumptions capture realistic properties of congested cities. Under these assumptions, equilibrium land rent, residential density, and numeraire consumption all decline exponentially with distance, although at different rates. If it is also assumed that traffic speed at the edge of a city is equal to free-flow speed, then the rates of decline in rent, residential density, and numeraire consumption all increase with the city's physical size. We also suggest a new statistical procedure for estimating negative exponential density functions from a cross section of cities of various sizes.

1. INTRODUCTION AND BACKGROUND

The monocentric model of urban land use has played a central role in the theoretical and empirical understanding of how urban areas function. This role continues despite the understanding that it is only a broad-brush approximation of actual land-use patterns. For example, as we have noted elsewhere [1], it provides an explanation—though not an entirely satisfac-

[1] The authors thank two anonymous referees and the editor for their helpful comments.

0094-1190/00 $35.00

tory one—for the two most robust facts about urban spatial structure: that gross residential density falls with distance from the city center, and that the rate at which it falls has diminished over time.

A particular form of the monocentric model, in which densities decline exponentially, plays an especially useful role by providing a single index summarizing the degree of centralization in land-use patterns. In this model, population density takes the form $D(x) = D(0)e^{-\gamma x}$, where $D(x)$ is the density at distance x from the center and γ is the *density gradient*. This negative exponential functional form for urban population densities was first introduced and tested empirically by Clark [4] in 1951. Researchers have estimated such functions for a large variety of metropolitan areas [6, 7, 10].

However, the theoretical conditions used to derive this negative-exponential form of the monocentric model have proven somewhat unsatisfactory. The first derivation was by Muth [11], who assumed that all residents are identical in income and in preferences, and have a compensated demand for land with a unitary price elasticity. Kim and McDonald [5] showed that Muth's assumption of a unitary compensated land price elasticity of demand is consistent with some utility functions, most notably with, $u = z + a \ln q$, where q is residential lot size, z is "other goods," and a is a positive coefficient. They also showed that a unitary compensated land price elasticity of demand implies a zero income elasticity of the demand for land if income and price elasticities are constants. They concluded, as had Brueckner [3], that empirical support for these assumptions is weak because compensated demand elasticities are known to be smaller than one and income elasticities to be positive. Papageorgiou and Pines [12] provided more general conditions on the utility function that lead to exponential densities while allowing for a positive income elasticity of demand for land.

Anas and Kim [2] showed that a good approximation to the negative exponential density functions is obtained by allowing for nonidentical incomes and that this approach better explains empirical observations. Assuming that all residents have incomes equal to the mean income of the city, as is done in all of the earlier models, causes the central densities—where the poor are located—to be underestimated and the fringe densities—where the rich are located—to be overestimated.

Our approach to deriving negative exponential densities contrasts sharply with the contributions cited above. We assume that all consumers have identical incomes and identical Cobb–Douglas preferences defined over "residential lot size" and "other consumption," with arbitrary expenditure shares. In earlier studies, it was assumed that the cost of commuting increases linearly with distance from the Central Business District (CBD).

This is unrealistic because of traffic congestion and the fact that transport capacity requires land, which is scarcer near the center.[2] We replace the linear-transport-costs assumption with one in which speed near the center is slower than that near the city's fringe. As it turns out, not just densities but all other endogenous distance functions of our model are negative exponential at equilibrium. That is why we refer to this simple model of residential land use as the *panexponential model*.[3]

Two implications of our assumptions about transport costs are that marginal (with respect to distance) commuting cost is a negative exponential function of distance from the CBD and that it is proportional to the commuter's income-earning potential. These characteristics approximate well two facts that are empirically supportable and difficult to ignore: first, that traffic congestion causes the time cost of traversing a unit distance to decline with distance from the CBD, and second, that commuting cost increases with the value of time of the commuter, which in turn increases with the commuter's income.

The price we pay for this more realistic description of transportation cost is that we do not determine transportation cost from primitive assumptions such as the fraction of land devoted to roads, but rather we assume that a specific functional form approximates the outcome of such assumptions. We do, however, anchor transportation cost in a realistic way by assuming that marginal transportation cost at the city's edge is exogenous, which we interpret to mean that it is determined by free-flow highway speeds. Under this assumption, we do comparative static analysis to determine the effect of city population on the physical size of a city and on spatial equilibrium within the city.

Comparative static analysis of our model shows that increasing the city's population, keeping transport costs constant, directly causes the city's land area to increase. However, if it is assumed that per-mile travel time at the city's fringe is equal to the free-flow (uncongested) travel time, then there is an indirect effect as well: increasing city population causes travel times within the city to rise, reflecting the building up of congestion. This indirect congestion effect causes residents to want to relocate closer to the center, and the city's radius is reduced if the indirect effect dominates the direct effect. We propose improved regression equations for estimating the density gradient, such that data from cities of different sizes can be pooled, improving statistical efficiency.

[2] This fact is most obvious for cities where commuters rely mainly on automobile or bus transit. But even for the few cities where commuters rely primarily on rail transit, speeds are slower in the center due to closer spacing between subway stations.

[3] Pan-, from the Greek meaning "all." Hence, *panexponential* means "everything is exponential."

2. THE PANEXPONENTIAL MODEL

Let x be distance from the CBD, and let $t(x)$ be the daily round-trip commuting time for someone at distance x, with $t(0) = 0$ (i.e., there are no transport costs which are not related to distance). We make two further assumptions. The first is about the consumer's time allocation, and the second about how commuting times and costs vary with distance from the CBD.

Assumption 1. The consumer has a fixed daily time budget (or endowment) H which is used for commuting or for working at a fixed wage rate, w. Hours spent at work are then $H - t(x)$. Daily money income is $w(H - t(x))$ or $y - T(x)$, where $y \equiv wH$ is the maximum possible money income and $T(x) \equiv wt(x)$ is transportation cost.

We will sometimes call y "income" for brevity, but it equals the money income only of a commuter residing adjacent to the CBD and, hence, enjoying zero commuting time.

Assumption 2. Within the city, commuting times and commuting costs satisfy the following differential equation, where $b > 0$ is independent of location, x:

$$\frac{t'(x)}{H - t(x)} = b, \qquad \text{or equivalently,} \qquad \frac{T'(x)}{y - T(x)} = b. \quad (1a, b)$$

The first version of this equation (1a) says that, at each x, the ratio of the marginal travel time to work hours equals b. The second (equivalent) version, (1b), says that, at each x, the ratio of the marginal commuting cost to money income equals b. The interpretation of this assumption is that the city's transportation infrastructure as a function of distance from the CBD is such that $\frac{t'(x)}{H - t(x)}$ is independent of x. The assumption captures the empirical regularity that congestion decreases with distance from the CBD; however, the implied pattern of transportation infrastructure is not in general first-best or second-best conditional on congestion pricing.

It will become apparent in the next section that there are several possible ways to close the model. This choice reduces to assumptions concerning which variables are exogenous and which endogenous, just as in the standard monocentric city model. In that model, population is exogenous and utility endogenous if the city is closed, and vice versa if the city is open. In the way we choose to close the model, b will be endogenous.

The solution to the differential equation (1) that meets the requirement of zero fixed transportation time is $t(x) = H(1 - e^{-bx})$, or equivalently,

$T(x) = y(1 - e^{-bx})$. From this solution several properties follow directly:

Property 1. Marginal commuting time and marginal commuting cost are each negative exponential functions of distance with gradient b. More precisely,

$$t'(x) = bHe^{-bx} \qquad \text{or, equivalently,} \qquad T'(x) = bye^{-bx}. \quad (2a, b)$$

Property 2. Money income $y - T(x)$ and work hours $H - t(x)$ are also exponentially declining with gradient b.

Property 3. Commuting time and cost are each concave with distance. This is because $t''(x) = -b^2He^{-bx} < 0$ and $T''(x) = -b^2ye^{-bx} < 0$.

Property 4. The gradient of the marginal travel time or transport cost is equal to the marginal (per mile) travel time at $x = 0$, divided by the time endowment H: that is, $b = t_0/H$ where $t_0 \equiv t'(0)$. This is seen from Eq. (1a) evaluated at $x = 0$. Note that this is a quantitative restriction on the transport-cost gradient, which is implied by Assumption 2 and the restriction that $t(0) = 0$.

To obtain additional results, we must now assume something about preferences and the distribution of income. Let q be the size of a commuter's residential lot and let z be the (numeraire) good representing all consumption other than that of land.

Assumption 3. Each consumer earns the same potential income as all others, y (or earns the same wage rate, w), and all consumers have the same Cobb–Douglas utility function $u(z, q) = z^{\alpha}q^{1-\alpha}$, where $0 < \alpha < 1$.

The budget constraint is $y - T - z - Rq = 0$, where R is the land rent. As is well known, expenditures on the two goods are then

$$z(x) = \alpha[y - T(x)], \tag{3a}$$

$$R(x)q(x) = (1 - \alpha)[y - T(x)], \tag{3b}$$

yielding achieved utility,

$$v(x) = u[z(x), q(x)] = \lambda[y - T(x)]R(x)^{\alpha-1}, \tag{3c}$$

where $\lambda \equiv \alpha^{\alpha}(1 - \alpha)^{1-\alpha}$. Equation (3b) can be rewritten as

$$R(x) = (1 - \alpha)[y - T(x)]D(x), \tag{3b$'$}$$

where $D(x) \equiv 1/q(x)$ is land-use density at x.

It is well known that the equilibrium requirement that the utility level be invariant with location is expressed by Muth's condition,

$$q(x)R'(x) = -T'(x). \tag{4}$$

Dividing (4) by (3b) and then applying Assumption 2 and Property 4, we get

$$\frac{R'(x)}{R(x)} = -\frac{1}{1-\alpha}\frac{T'(x)}{y-T(x)} = -\frac{b}{1-\alpha} = -\frac{1}{1-\alpha}\frac{t_0}{H}. \tag{5}$$

This result is stated as

Property 5. The equilibrium rent–distance function is negative exponential with gradient $b/(1-\alpha)$. Note that the larger is t_0 (the marginal transport time at the CBD), and the smaller is $1-\alpha$ (the expenditure share of land), the larger will be the gradient $-R'/R$ of the rent–distance function.

Next, differentiate (3b′) with respect to x, divide the resulting equation by (3b′), and apply Properties 4 and 5 to get

$$\frac{D'(x)}{D(x)} = \frac{R'(x)}{R(x)} + \frac{T'(x)}{y-T(x)} = -\frac{b}{1-\alpha} + b = -\frac{\alpha}{1-\alpha}b$$
$$= -\frac{\alpha}{1-\alpha}\frac{t_0}{H}. \tag{6}$$

This result is stated as

Property 6. The residential density function is negative exponential with gradient $\gamma \equiv b\alpha/(1-\alpha)$. Lot size (the reciprocal of density) is positive exponential. Specifically,

$$q(x) = (1-\alpha)wHR(0)^{-1}e^{(\alpha b/(1-\alpha))x}, \qquad x < \bar{x}. \tag{7a}$$

The next property follows from (3a) and Property 2.

Property 7. Numeraire consumption, $z(x)$, is negative exponential with gradient b. Specifically,

$$z(x) = \alpha wHe^{-bx}. \tag{7b}$$

In summary, we have proved that Cobb–Douglas preferences (Assumption 3) together with Assumptions 1 and 2 (which imply negative exponential marginal travel time and transport-cost functions) guarantee that rent, residential density, and other consumption are also negative exponential functions of distance from the CBD. If $\alpha > 0.5$, as is realistic, these four gradients can be ordered in absolute value, as

$$|R'/R| = b/(1-\alpha) > |D'/D| = b\alpha/(1-\alpha) > |T''/T'| = |z'/z| = b. \tag{8}$$

For example, suppose that the expenditure share of residential land is $\frac{1}{10}$ ($\alpha = 0.90$); then the rent gradient is 10 times as large, and the density gradient 9 times as large, as the gradient of marginal transport cost. With $\alpha = 0.75$, the rent gradient is 4 times as large, and the density gradient 3 times as large, as the gradient of the marginal transport cost. And with $\alpha = 0.50$, the rent gradient is twice as large as the gradient of the marginal transport cost, while the density gradient and the gradient of the marginal transport cost are equal.

3. MODEL CLOSURE

As noted earlier, there are several ways in which the model can be closed. We consider only one. In addition to the usual arbitrage on residential rents at the edge of the city, we assume that the city is closed (i.e., population is exogenous) and that the marginal travel time (and, therefore, travel speed) at the boundary of the city, $\bar{t}$, is exogenous. The latter assumption is a realistic approximation, since travel at a city's fringe is normally free of congestion. In the standard closed monocentric city model, the city population, N, is taken as exogenous, while the equilibrium utility level, u, and the location of the boundary of the city, $\bar{x}$, are endogenous. We adopt this convention and add the assumption that $\bar{t}$ is exogenous. Therefore, the gradient b (and, hence, $t_0 = bH$) is endogenous.

Thus, based on the above explanation, there are three closure conditions. First, rent and density must be determined by the usual boundary condition at the edge of the city. Namely, let $\bar{x}$ be the outer edge of residential development, and let R_a be the land rent there for nonresidential purposes. Then, arbitrage between residential and nonresidential land requires that

$$R(\bar{x}) = R_a. \tag{9}$$

Then the rent−distance and density−distance functions are

$$R(x) = R_a e^{(b/(1-\alpha))(\bar{x}-x)}, \tag{10a}$$

and

$$D(x) = \frac{R(0)}{(1-\alpha)wH} e^{-(\alpha b/(1-\alpha))x} = \frac{R_a e^{(b/(1-\alpha))\bar{x}}}{(1-\alpha)wH} e^{-(\alpha b/(1-\alpha))x}. \tag{10b, c}$$

The second closure condition is that an exogenous city population be accommodated (closed city model). That is, for a circular city,

$$N = 2\pi \int_0^{\bar{x}} D(x)\, x\, dx = \frac{2\pi R_a e^{(b/(1-\alpha))\bar{x}}}{(1-\alpha)wH} \int_0^{\bar{x}} x e^{-(b\alpha/(1-\alpha))x}\, dx. \tag{11}$$

The equilibrium level of utility, u, can be calculated from the achieved utility (3c) evaluated at any x, using $u = \nu(x)$. For convenience, we evaluate it at $x = 0$ and $x = \bar{x}$:

$$u = \lambda w H e^{-b\bar{x}} R_a^{\alpha-1} = \lambda w H R(0)^{\alpha-1}. \tag{12a, b}$$

The third closure condition is

Assumption 4. Marginal travel time at the edge of the city is exogenous:

$$t'(\bar{x}) = \bar{t}. \tag{13}$$

The assumption states that an exogenous free-flow travel speed applies at the edge of the city. As stated earlier, this is realistic since travel at a city's fringe is normally free of traffic congestion.

Equations (9), (11), and (13) together determine three endogenous variables, $R(0)$, $\bar{x}$, and b (or $t_0 = bH$), from the exogenous parameters w, H, α, R_a, $\bar{t}$, and N. The utility level, u, is then obtained from (12b); (10b) gives the density–distance function $D(x)$; and (10a) gives the rent–distance function, $R(x)$.

4. THE EFFECT OF POPULATION SIZE ON THE GRADIENTS

We have already pointed out that an advantage of our model is that, in attempting to explain negative exponential rent and density functions, it recognizes the spatial variation of congestion. Although congestion and the allocation of land to roads are not explicitly treated within the model, we will show that, under Assumption 4, an exogenous increase in the population of the city does cause the travel time function to shift and the congested travel times which are approximated by the model to be altered.[4] Are these congestion-related predictions of the model intuitively reasonable? We will show that the model predicts increased congestion, steeper rent–distance and density–distance functions, and lower equilibrium utility and consumption as the city's population rises. But the effect on the city's boundary is ambiguous: the model predicts that very large cities shrink in radius as the population increases, while smaller cities expand in radius.

[4] Given a particular allocation of land to roads (short run), only a change in population, N, can alter t_0. However, because Assumption 4 means that t_0 depends on $\bar{x}$, which is affected by the exogenous variables w, H, α, R_a, and $\bar{t}$, changes in these variables will also affect t_0 in our approximate model. In the long run, changes in these variables would again affect t_0, by altering the allocation of land to roads. But our model does not explicitly treat how land is allocated to roads. For this reason, we do not attempt a comparative static analysis with respect to w, H, α, and R_a. We only present the comparative statics analysis with respect to population, N.

To obtain these results we do a comparative static analysis of the equilibrium with respect to population, N, showing how it affects the endogenous variables and functions, which are, $\bar{x}$, t_0, u, $R(x)$, and $D(x)$. The details of the analysis are in the Appendix; they show that there will be two regimes, as indicated in the previous paragraph. To see this let us recall that, under Assumption 4 and Properties 1 and 4, $t_0 = \bar{t}e^{b\bar{x}} = \bar{t}e^{(t_0/H)\bar{x}}$. Taking logarithms and solving for $\bar{x}$, we can write this as

$$\ln t_0 = \ln \bar{t} + (t_0/H)\bar{x}. \tag{14}$$

This states a fundamental relationship between $\bar{x}$ and t_0. It is easy to show, by total differentiation of (14), that

$$\frac{d\bar{x}}{dt_0} = \frac{H - t_0\bar{x}}{t_0^2} \geq (<)0 \qquad \text{as } t_0\bar{x} \leq (>)H. \tag{15}$$

Regime 1 ($t_0\bar{x} < H$). This case corresponds to cities which are not extremely large in radius. The inequality says that the city radius should not be so large that commuting from the edge of the city to the CBD and back would consume the entire daily time budget of the commuter, were he/she to travel at a constant speed equal to the speed prevailing at the CBD. Suppose, for example, that travel speed at the CBD is 5 miles per hour, reflecting highly congested conditions. Then, t_0 is 24 minutes per mile (recalling that it should be measured for a round trip). This means that if $H = 600$ minutes (i.e., 10 hours are available for work and commuting each day), then such a city should be less than 25 miles in radius for this regime to hold.

It is shown in the Appendix that an increase in N causes t_0, the marginal travel time at the CBD, to increase. This happens because the higher N causes more congestion, which is approximated by the model. Then (15) immediately implies that for a city in this regime, the radius expands. Since $t(x) = H(1 - e^{-bx})$ and since $b = t_0/H$ is higher, it follows that travel time $t(x)$ rises everywhere within the city. The Appendix also shows that the equilibrium level of utility, u, falls and that land rent and residential density rise at every location in the city. Since t_0 is higher, it follows from (5) and (6) that the rent and density gradients increase. These results are intuitively satisfactory; they indicate that population growth causes changes which resemble those that would occur with the buildup of congestion in a growing city.

Regime 2 ($t_0\bar{x} > H$). This regime holds for cities which are large according to the criterion described above: a commuter traveling from the city fringe to the CBD and back at the speed which prevails at the CBD would require more than the entire time budget to complete the round trip. Using the numbers in the example given above, the city would have to be more than 25 miles in radius. In this case the comparative static results are somewhat different. The Appendix shows that, as in Regime 1, the increase in N causes the equilibrium level of utility to fall, and the rent and residential density near the CBD to rise. Also, as in Regime 1, travel time at the CBD increases as well and, as before, this causes the travel time function $t(x)$ to rise everywhere. In turn, we can see from (5) and (6) that the rent and density gradients increase, as was also true under Regime 1. However, while rent and density rise near the CBD, there is an intermediate distance, $\hat{x}$, beyond which rent falls. To express this differently: the rent–distance function rotates around the point $\hat{x}$, rising for lower x and falling for higher x. It follows from the arbitrage condition given by (9) that, since the nonresidential rent has not changed, the city radius shrinks. Meanwhile, the density–distance function can either rise throughout the city or fall beyond some distance $\tilde{x} > \hat{x}$ while rising for shorter distances. That $\tilde{x} > \hat{x}$ means that the point around which the density–distance function rotates (if such a point exists) is always farther from the CBD than is the point around which the rent–distance function rotates. That $\tilde{x} > \hat{x}$ follows because the density at $\hat{x}$ cannot have decreased. To see this, consider that rent at $\hat{x}$ is unchanged, while travel time at $\hat{x}$ has increased. It can be seen from (3b) that to maintain a constant expenditure share for land at $\hat{x}$, lot size must decrease (i.e., density must increase).

In summary, the panexponential model predicts that for cities which are initially not very large in radius (Regime 1), an increase in population expands the city, drives up rents and densities, and causes travel times to rise. The response is similar in cities which are initially very large in radius (Regime 2): travel times rise and rents and densities climb near the center. However, in this case the population increase causes travel congestion to increase so much that the tendency of households to locate more centrally, to economize on their travel times, causes the city to shrink in radius, and rents and possibly also densities to fall near the fringe. This result of Regime 2 is not necessarily inconsistent with the comparative static analysis of a monocentric city model in which congestion and the allocation of land to roads are endogenous. Although there is no formal analysis of this type of model in the literature, the result can be conjectured in piecemeal fashion, by drawing on the well-known results from the comparative static analysis of a standard monocentric city without congestion (e.g., Wheaton [14]).

Drawing on the standard model, we know that an increase in population has two effects on a city's expansion: a direct effect and an indirect effect. The direct effect would expand the city, keeping all else, including travel speeds, constant. The indirect effect would operate as follows: the population increase would increase congestion and travel times. The comparative static analysis of the standard model shows that the higher travel times, in turn, would cause residents to want to relocate more centrally, which would tend to reduce the city's radius. The combination of the direct and indirect effects is ambiguous. Regime 1 of the panexponential model can be interpreted as saying that the direct effect dominates and the city expands in radius. Regime 2 can be interpreted as saying that the indirect effect dominates and the city shrinks in radius.

We have cautioned that these results are approximations of reality, since congestion and the allocation of land to roads are not explicit in the panexponential model. The purpose of the model is to suggest a family of interrelated exponential distance functions describing transport costs, rents, and densities which can be fit to data from cities of various sizes, as discussed in the next section. The comparative static exercise summarized above (and detailed in the Appendix) demonstrates that the variation in these exponential functions across cities of different sizes should yield conceptually consistent results.

5. CONCLUSION

The panexponential model offers an alternative, within the monocentric framework, to previous theoretical derivations of negative exponential residential density functions. Instead of making restrictive assumptions about the compensated demand for land, we make a restrictive assumption about marginal transportation cost, namely that it rises toward the center according to Eq. (1b). This may be regarded as an implicit assumption about congested conditions and the allocation of land to roads within the city. We also assume that transportation cost arises from a value of time equal to the wage rate and that free-flow speed prevails at the city's outer edge. These assumptions are at least qualitatively realistic in their depiction of congestion and of the covariance between marginal transportation cost and income, both of which are entirely absent in other models leading to negative exponential densities.

Like other monocentric models, ours generates a relationship between the rent and density gradients: namely, the latter is a fraction α of the former, where $1 - \alpha$ is the proportion of income spent on residential land. This is a special case of a more general power−law relationship between density and rent, demonstrated by Mills and Hamilton [8, p. 430].

How might our model be tested empirically? Taken literally, the model implies a density gradient equal to $\frac{\alpha t_0}{(1-\alpha)H}$, which could be tested directly. A test more focused on the novel features of our model is whether the *ratio* of the density gradient to the gradient of marginal transportation time is equal to $\frac{\alpha}{1-\alpha}$, as implied by Eq. (8). We do not advocate such tests because cities are far from truly monocentric and their residents do not have identical incomes, time budgets, or preferences. A less direct test would involve the comparative static implications of the model, which are not identical to those of other versions of the monocentric model because of our more realistic assumptions about transport costs. For example, as shown in Section 4, our model potentially generates a nonmonotonic relationship between city population and area.[5] This could be tested by using comparable criteria in defining each city's boundary, as was attempted by the US Census Bureau in 1985 [15, p. 856].

A rather general way of testing comparative statics, which appears not to have been noted in the literature, would be to estimate density gradients from observations pooled across cities. One would start with the traditionally employed regression equation, which has the form $\ln D(x) = \ln D(0) = -\gamma x + \varepsilon$, where ε is a random error term. One would then rely on the comparative static results to specify functional relationships $\ln D(0) = f(Y)$ and $\gamma = g(Y)$, where Y is a set of city characteristics taken to be exogenous, such as population, wage rate, land value in nonurban uses, and travel speed near the edge. Then, the regression to be estimated is of the form

$$\ln D_{ij} = f(Y_i) - g(Y_i)x_{ij} + \varepsilon_{ij},$$

where $[D_{ij}, x_{ij}]$ is the jth observation on density and distance in city i. Supposing, for example, that the f and g functions are specified as being linear in coefficients, the above regression is linear in the coefficients of the Y_i variables which measure city characteristics and linear in the coefficients of the interaction terms $Y_i x_{ij}$. Such a regression would enable the researcher to compare how well alternative theoretical derivations of the negative exponential model hold up, or to predict density gradients for cities not in the sample, without the loss of statistical efficiency inherent in the two-step process used previously [7, 9, 10].

APPENDIX: EFFECT OF AN INCREASE IN N

The two market equilibrium conditions are Eqs. (11) and (13). Equation (11) says that the integral of population density over the residential area of the city equals the exogenous population. Equation (13) says that the marginal travel time at the edge of the city equals the free-flow travel time, which is taken as exogenous. These two equations must be solved

[5] Models in which transport costs are exogenous (uncongested) imply that a city's area increases with population. In Pines and Sadka [13] a nonmonotonic relationship between population and city area arises from having closure with respect to rent distribution. In our case, it arises from congestion that varies with distance. We are not aware of any formal comparative static analysis of a monocentric model with an explicit congestion technology.

simultaneously to find the city's radius, $\bar{x}$, and the CBD travel speed, t_0, given the exogenous population, N, the free-flow travel time, $\bar{t}$, the wage rate, w, the agricultural rent, R_a, and other parameters.

Defining $\theta \equiv \bar{x}b$ and using integration by parts to evaluate the integral in (11) gives

$$b = k_1\left(e^{\alpha\theta/(1-\alpha)} - 1 - \frac{\alpha\theta}{1-\alpha}\right), \qquad \text{where } k_1 = \frac{2\pi R_a(1-\alpha)}{w\alpha^2 N\bar{t}}. \quad \text{(A.1)}$$

From Property 1 (Eq. (2a)) the exogenous free-flow travel time condition is $bHe^{-b\bar{x}} = \bar{t}$ and, again using $\theta \equiv \bar{x}b$, it can be rewritten as

$$b = k_2 e^{\theta}, \qquad \text{where } k_2 = \bar{t}/H. \quad \text{(A.2)}$$

Now imagine plotting (A.1) and (A.2) with b on the vertical axis and θ on the horizontal axis. It is easy to show that, in the positive quadrant, (i) (A.1) is a strictly convex function passing through the origin; (ii) (A.2) is a strictly convex (exponential) function with a positive b-intercept equal to k_2; and (iii) when $\alpha \geq 0.5$, which has been assumed and covers realistic cases, (A.1) and (A.2) intersect once and only once, with (A.1) cutting (A.2) from below and, hence, (A.1) having the steeper slope at the point of intersection. That (A.1) and (A.2) must intersect is proved by showing that, for large enough θ and $\alpha \geq 0.5$, the value of b calculated from (A.1) exceeds that calculated from (A.2).

A rise in N has no effect on (A.2) but causes $\frac{db}{d\theta}$ computed from (A.1) to decrease at any θ, which means that (A.1) pivots clockwise around the origin, causing the intersection point to move up on the (A.2) curve. This establishes that

$$\frac{d\theta}{dN} > 0 \qquad \text{and} \qquad \frac{db}{dN} > 0. \quad \text{(A.3a, b)}$$

Since $t_0 = bH$, it also follows immediately from (A.3b) that $\frac{dt_0}{dN} > 0$.

Differentiating (A.2) yields

$$\frac{d\bar{x}}{db} = \frac{(1-\theta)}{b^2}. \quad \text{(A.4)}$$

Now note, from (15), that $\theta \equiv \bar{x}b = \bar{x}t_0/H < (\geq) 1$ as $t_0\bar{x} < (\geq) H$, and combine (A.3b) and (A.4). Then

$$\frac{d\bar{x}}{dN} > (\leq) 0 \qquad \text{as } t_0\bar{x} < (\geq) H. \quad \text{(A.5)}$$

This proves the nonmonotonic relationship between population and city radius which was stated in the text and separates the two regimes. We now turn to the effect of population on rent, density, and utility. From (10a), $R(0) = R_a e^{\theta/(1-\alpha)}$, and it can be seen that $\frac{dR(0)}{dN}$ and $\frac{d\theta}{dN}$ have the same signs and, from (10b), $\frac{dD(0)}{dN}$ and $\frac{dR(0)}{dN}$ have the same signs. From (12b), the sign of $\frac{du}{dN}$ is the opposite of the sign of $\frac{dR(0)}{dN}$. Then, we have established that

$$\frac{dR(0)}{dN} > 0, \qquad \frac{dD(0)}{dN} > 0, \qquad \text{and} \qquad \frac{du}{dN} < 0. \tag{A.6}$$

Now consider what happens under the two regimes discussed in Section 4. Under regime 1 ($t_0\bar{x} < H$ or $\theta < 1$), we have shown that $\frac{db}{dN} > 0$, and it follows from (A.5) that $\frac{d(\bar{x}-x)}{dN} > 0$ and $\frac{d(\bar{x}-\alpha x)}{dN} > 0$ for any x. It then follows from (10a) and (10b) that $\frac{dR(x)}{dN} > 0$ and $\frac{dD(x)}{dN} > 0$ for any $x \in [0, \bar{x})$. Under Regime 2 ($t_0\bar{x} > H$ or $\theta > 1$), however, since $-\frac{R'(x)}{R(x)} = \frac{b}{1-\alpha}$ rises, the rent at the CBD rises (see (A.6)), and the city shrinks in radius; there is an $\hat{x} \in [0, \bar{x})$ such that for $x < \hat{x}$ rent rises while for $x > \hat{x}$ rent falls. In other words, the rent–distance function rotates clockwise around $\hat{x}$. All that remains is to characterize how the rise in N affects $D(x)$ under this regime. Note that $-\frac{D'(x)}{D(x)} = \frac{\alpha b}{1-\alpha}$ rises. Also, from (10c),

$$\frac{dD(x)}{dN} = D(x)\left(\frac{1}{1-\alpha}\frac{d\theta}{dN} - \frac{\alpha x}{1-\alpha}\frac{db}{dN}\right). \tag{A.7}$$

Substituting $\frac{d\theta}{dN} = \bar{x}\frac{db}{dN} + b\frac{d\bar{x}}{db}\frac{db}{dN}$ and using (A.4) for $\frac{d\bar{x}}{db}$, we get

$$\operatorname{sign}\left(\frac{dD(\check{x})}{dN}\right) = \operatorname{sign}\left[(1 - \alpha\check{x}b) = \left(1 - \alpha\frac{\check{x}}{\bar{x}}\theta\right)\right]. \tag{A.8}$$

Recall from the discussion in the text that the density–distance function could rotate around some point $\tilde{x} < \bar{x}$. In (A.8), $\check{x} \in (\tilde{x}, \bar{x})$ is a location near the outer edge of the city, so that $\frac{\check{x}}{\bar{x}} \cong 1$. Thus, for $\alpha\theta < 1$ or $\alpha t_0\bar{x} < H$, the rise in N causes density to rise everywhere in the city, while for $\alpha\theta > 1$ or $\alpha t_0\bar{x} > H$, density rises for $x < \tilde{x}$, where $\tilde{x} = \bar{x}/\alpha\theta$, while density falls for $x > \tilde{x}$. Finally, recall that an argument in the text established that $\tilde{x} > \hat{x}$, where $\hat{x}$ is the point around which the rent–distance function rotates clockwise.

REFERENCES

1. A. Anas, R. Arnott, and K. A. Small, Urban spatial structure, *Journal of Economic Literature*, **36**, 1426–1464 (1998).
2. A. Anas and I. Kim, Income distribution and the residential density gradient, *Journal of Urban Economics*, **31**, 164–180 (1992).

3. J. K. Brueckner, A note on sufficient conditions for negative exponential population densities, *Journal of Regional Science*, **22**, 353–359 (1982).
4. C. Clark, Urban population densities, *Journal of the Royal Statistical Society A*, **114**, 490–496 (1951).
5. K. H. Kim and J. F. McDonald, Sufficient conditions for negative exponential density: Further analysis, *Journal of Regional Science*, **27**, 295–298 (1987).
6. J. F. McDonald, Econometric studies of urban population density: A survey, *Journal of Urban Economics*, **26**, 361–385 (1989).
7. E. S. Mills, "Studies in the Structure of the Urban Economy," Johns Hopkins Press, Baltimore/London (1972).
8. E. S. Mills and B. W. Hamilton, "Urban Economics," Scott, Foresman and Co., Glenview, IL (1989).
9. E. S. Mills and R. Price, Metropolitan suburbanization and central city problems, *Journal of Urban Economics*, **15**, 1–17 (1984).
10. E. S. Mills and J. P. Tan, A comparison of urban population density functions in developed and developing countries, *Urban Studies*, **17**, 313–321 (1980).
11. R. F. Muth, The spatial structure of the housing market, *Papers and Proceedings of the Regional Science Association*, **7**, 207–220 (1961).
12. Y. Y. Papageorgiou and D. Pines, The exponential density function: First principles, comparative statics and empirical evidence, *Journal of Urban Economics*, **26**, 264–268 (1989).
13. D. Pines and E. Sadka, Comparative static analysis of a fully closed city, *Journal of Urban Economics*, **20**, 1–20 (1986).
14. W. Wheaton, A comparative static analysis of urban spatial structure, *Journal of Economic Theory*, **9**, 223–237 (1974).
15. U.S. Bureau of the Census, "Statistical Abstract of the United States: 1994," U.S. Government Printing Office, Washington, DC (1994).