NBER WORKING PAPER SERIES

URBAN GROWTH

Duncan Black
Vernon Henderson

Working Paper 6008

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1997

The authors acknowledge gratefully support of the National Science Foundation for this research. This paper is part of NBER's research programs in Economic Fluctuations and Growth and Public Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Urban Growth
Duncan Black and Vernon Henderson
NBER Working Paper No. 6008
April 1997
JEL Nos. O4, R00, E2, J24
Economic Fluctuations and Growth
and Public Economics

ABSTRACT

This paper models and examines empirically the evolution of cities in an economy. Twentieth century evolution in the USA is characterized by parallel growth of cities of different types and on-going entry of new cities, together maintaining a stable relative size distribution of cities. Each type of city has a particular industrial composition and good(s) it specializes in and corresponding equilibrium size. This evolution is modeled in an economy with exogenous population growth and endogenous human capital accumulation. Within cities, there are knowledge spillovers as well as scale externalities. Individual city sizes grow with human capital accumulation; and cities grow in number if national population growth is high enough. Different types of cities grow in parallel in size and human capital accumulation. However, per capita income and human capital levels differ across city types by production process and benefits of human investments and spillovers, so there is observed inequality across cities among otherwise identical individuals.

Duncan Black
Department of Economics
Box B
Brown University
Providence, RI 02912

Vernon Henderson
Department of Economics
Box B
Brown University
Providence, RI 02912
and NBER
j_henderson@brown.edu

Urban Growth[†]

Duncan Black and Vernon Henderson

Brown University

This paper models and examines empirically the evolution of cities in an economy. Key features of economic expansion are population growth and human capital accumulation. To illustrate, in the USA from 1900-50, a period of rapid urban evolution, national population grew at an average annual rate of 1.4% and urban population at a 2.2% annual rate. The percent of the 17-year old population completing high school rose from 6.3 to 57.4%.[1] Population growth and human capital accumulation interact with the urbanization process to raise two sets of issues. The first involves the nature and efficiency of the evolution of the spatial organization of an economy. The second concerns the relationship between urbanization and evolving income inequality.

For the first set of issues, consider that each economy has a size distribution of cities. National population growth can be accommodated by increases in the sizes of existing cities, formation ("entry") of new cities, or both. As either or both occur, the relative size distribution of cities can remain stable or exhibit a tendency to collapse ("converge" to a common city size) or spread. Apart from intellectual curiosity about our geography, these issues have public policy relevance. For example, whether population growth is primarily absorbed through expansion of existing cities or creation of new ones has implications for the efficient intercity allocation of public infrastructure investments. Governments in developing countries such as Thailand, Mexico, Indonesia, and Bangladesh focus infrastructure investments on just one or two mega cities, fueling top heavy urban development. If, in fact, markets would tend towards more parallel development of existing cities along with increases in numbers of cities instead of tending towards mega city development, these governments should

[†]The authors acknowledge gratefully support of the National Science Foundation for this research.

[1]The average annual number of days students attend school annually also rose from 99 to 158. Adult illiteracy fell from 10.7 to 2.7%. All numbers are from *Historical Statistics of the US: Colonial Times to 1970*, U.S. Bureau of the Census.

be spreading investment out across many cities. A mega-city development policy may lead to inefficient crowding of tens of millions of people into tiny areas, with resulting deplorable living conditions.

In thinking about the tendency of markets towards urban population distribution, traditional urban analysis suggests two important considerations. First, underlying the city size distribution, at any point in time, there are different types of cities where urban models predict that different types and sizes of cities have different industrial bases. Such models may be based on Henderson's (1974) model of city specialization derived from the nature and extent of urban scale economies or on spatial hierarchy models of cities (Fujita, Krugman, and Mora (1995)). Later in the paper we will present evidence on urban specialization. Second, the efficiency of the urbanization process and the possibility of excessive urban concentration depend on the role which land developers and autonomous local governments play in the national city formation and growth process (Helsley and Strange (1993), Krugman (1993), and Becker and Henderson (1996)). Both considerations will be incorporated into our model.

Human capital accumulation and endogenous growth introduce additional key considerations in the analysis of spatial evolution. Urbanization involves economies of scale in production, as Alfred Marshall wrote about so cogently in 1890. Individual human capital accumulation and localized knowledge spillovers will accentuate the scale benefits of urbanization, with spillovers creating a stock of what could be labelled localized "trade secrets." By emphasizing the local nature of knowledge spillovers in an endogenous growth context, as suggested in Lucas (1988), we will show that human capital accumulation fuels urban growth, bearing directly on the question of whether a developing economy grows with bigger or more cities.

Human capital accumulation and localized knowledge spillovers in an urbanization context also have other implications. Given that different technologies are used in the industrial sectors of different types of cities, private returns to human capital investment and the magnitude of knowledge spillovers will vary naturally across city types. Consequently, the equilibrium levels of per person human capital and contemporaneous real incomes will vary across different types of cities.

That fact suggests the second set of issues concerning the interaction between urbanization and national growth: evolving income inequality in an economy. We will see that an economy grows at least with measured income and educational inequality across different types of cities for otherwise identical people, although in our formulation per person consumption levels will be equal. In extensions we will ask, under what formulations parental human capital and locational choices of initial generations can lead to inequality in per person consumption levels in later generations. In future work underway, by examining the interplay of ability levels, given either by nature or influenced by the choices of prior generations, with endogenous location and education choices which determine how heterogeneous agents sort themselves out spatially in an economy, it will be possible to further examine how the evolution of inequality can be affected by urban choices.

The empirical side of this paper is related to work by Dobkins and Ioannides (1995) examining USA urban evolution from 1900 to the present. On the theoretical side, apart from late 1970's urban exogenous growth models, summarized in Henderson (1988), the paper is most closely related to Eaton and Eckstein (1994), who consider human capital spillovers within and also transmitted across cities, an exciting extension. However, in Eaton and Eckstein, the number of cities is fixed and the endogenous growth model is not fully specified. In this paper, we step back and solve a fully specified growth model, focusing on issues of city formation and the effect of endogenous growth on changes in city sizes, numbers, and human capital levels over time. In terms of evolving income inequality, the work in this paper and extensions on the interplay between inequality, ability differences, human capital accumulation and urbanization are most closely related to that of Benabou (1993, 1996), with his rich specification of peer group effects and parental choices. However, Benabou is not concerned with national spatial evolution per se, and does not allow for city formation and growth, multiple numbers and types of cities, transformation of local industrial bases and other key considerations in this paper.

In section 1 of the paper, we present and analyze a simple growth model of an urbanized economy, experiencing exogenous population growth, in which endogenous growth occurs through human capital accumulation. In section 2, we present empirical evidence on the evolution of cities in the economy, urban specialization, and the transformation of local industrial bases. Section 2 establishes basic information about cities. Over time, in the USA, with national population growth and human capital accumulation, the numbers and especially the sizes of cities have grown each decade. Despite this growth the relative size distribution of cities has remained remarkably stable, showing no tendency to collapse or spread. We will show that, as noted above, underlying this size distribution are different types of cities, which appear to differ significantly in typical size and educational attainment. In addition, the relative educational distribution of cities appears to also be stable over time. Section 1 develops a model of national urban growth which is consistent with these facts and gives us a framework to start to discuss issues of efficiency, institutions, and inequality.

I. A MODEL OF EVOLUTION OF AN URBANIZED ECONOMY

The growth model of an urbanized economy consists of two components. First is the urban part which describes the spatial organization of production and population. In this paper, the economy consists of two types of cities, each performing different functions and having different equilibrium sizes, per worker human capital levels and incomes. While having two types of cities does not make for a very sophisticated size distribution, it is sufficient to establish basic principles. Type 1 cities in the economy produce the numeraire good, an intermediate input (e.g., materials, disposable machines), that is purchased by firms in type 2 cities. Firms in type 2 cities specialize in production of the economy's consumption good, priced at P relative to the numeraire. This characterization of cities as being absolutely specialized in traded good production with no costs of inter-city trade begs questions about the role of fairly diversified mega-cities in an economy. Again we are simplifying to establish principles about growth, avoiding some of these broader issues analyzed in Henderson (1988). Finally as characterized below, there is a city formation process in national land markets, involving either land developers or autonomous local governments.

The second component to an urban growth model involves family migration and human capital investment decisions. Workers are members of dynastic families. At the end of this section, we will show that our basic results — allocations of people and human capital across cities — hold in a simple overlapping generations model. For dynasties, each family starts with the same per person human capital and each family's size grows at the same rate, g. Each family discounts the future at a rate ρ, where $\rho > g$ to help ensure well-behaved solutions. At each instant, dynasties choose how much total family income to allocate to per member consumption, c, and how much to allocate to increasing the family's human capital stock. Families must allocate also their members across city types and decide on the per person human capital investments for members by the type of city in which they live. For existing family members, current own human capital endowments are nontransferable, except to newborns. Family decisions govern human capital accumulation, in the absence of formal markets for human capital which we rule out under the usual "no slavery" constraint.

For any dynasty, using a common form to utility of per person consumption, c, the optimization problem is (without subscripting for t)

$$\max_{c, h_1, h_2, z} \int_0^\infty \left(\frac{c^{1-\sigma} - 1}{1 - \sigma} \right) e^{-(\rho - g)t} dt \qquad \sigma > 0,\ \rho > g \tag{1}$$

$$\text{s.t.} \quad P\dot{H} = ze^{gt}I_1 + (1-z)e^{gt}I_2 - Pce^{gt} \qquad \text{(a)}$$

$$H = ze^{gt}h_1 + (1-z)e^{gt}h_2 \qquad \text{(b)}$$

$$\dot{H} \geq 0;\ \dot{h}_1/h_1 + g \geq 0,\ \dot{h}_2/h_2 + g \geq 0 \quad \text{(c)}$$

In equation (1), given an initial normalized family size of 1, family size at time t is e^{gt}. H is the family's human capital stock. z proportion of family members are assigned to type 1 cities and $(1-z)$ to type 2 cities. I_1 and I_2 represent net per worker incomes earned by workers living in type 1 and type 2 cities respectively, with h_1 and h_2 representing their human capital levels.

Constraint (a), the equation of motion, states that the value of family human capital growth $(P\dot{H})$ is the difference between total family income $(ze^{gt}I_1+(1-z)e^{gt}I_2)$ and value of consumption Pce^{gt}. In (a) it is assumed noncritically that human capital is formed by conversion of the consumption good produced in type 2 cities and sold at price P. Constraint (b) states that total family human capital is the sum of individual human capitals (h_1 and h_2) of members in type 1 (ze^{gt}) and type 2 $((1-z)e^{gt})$ cities.

In our formulation of the dynasty's problem in equation (1), we make two important assumptions that merit some discussion. First, we assume that human capital is transferable to a limited degree. Specifically, the first constraint in (c) $(\dot{H} \geq 0)$ states that families can neither borrow nor consume their human capital; conversion of the consumption good to human capital is irreversible. The second constraint in (c) states that once installed, human capital is only transferable as an endowment to newborns in the same city type. $\dot{h}_i/h_i + g \geq 0$ states that the maximal percentage drop in per member human capital in a city type is the growth rate of their offspring. Neither of these constraints is binding in equilibrium and this formulation is also consistent with an additional constraint that human capital is specific to either a city or an industry, partially or fully non-transferable through migration across cities (see later).

Second, given that family members will generally earn different incomes I_1 and I_2 by city type, there must generally be intra-family transfers across cities to maintain equality of per member consumption, c. Although in many countries transfers from residents in large cities to relatives in smaller towns can amount to 10% of family income, in modern economies people may not think in these terms. The rigidity of this formulation can be relaxed in several ways. First, the need for income transfers can be avoided if all members of a generation are born with identical initial endowments of wealth which can be invested in personal human capital or in the human capital of siblings; individuals born in low human capital cities can "invest" informally in the human capital of family members in high human capital cities and receive the returns. That is, rather than a family matriarch stating "George, give Harry $2,000 every year," as implied by (1), George can promise "Harry, if you lend me money for education, I will reimburse you at an (appropriate) rental rate." Secondly, dynasties can splinter as long as each splinter starts with the same per person stock of human capital (H/e^{gt}). Third, following from the first, if there were a formal market for human capital, each family or family splinter could reside entirely in one type of city or the other, borrowing/investing in the human capital of workers in the other type of city.

In order to proceed with optimization problem in (1), it is necessary to determine the expressions for net real incomes, I_1 and I_2, which family members can earn in city types 1 and 2. To solve for these, as well as to detail the nature of local human capital spillovers, we need to turn to an analysis of production in cities, determination of city sizes (which affect the returns to human capital investment), and the like. Given that analysis, we can then return to the problem in (1) to study investment and migration decisions of families. Those will determine the evolution of overall human capital levels, the formation of cities and their growth, and the size distribution of cities.

The Structure of Cities

Contemporaneous city formation and size determination involves a tradeoff within cities between the benefits and costs of changing city sizes. We assume production in a city occurs under "localization" economies of scale — own industry local external economies of scale. Contemporaneous efficiency of each firm is enhanced by having more firms in the same industry in a city, with whom the firm communicates about what inputs to buy from whom, what product lines to emphasize, or how to organize production. This variant of a communications model may involve exogenous spillovers (Fujita and Ogawa, 1982) or endogenous information exchange (Kim, 1988). Other externalities may also be at work (Helsley and Strange, (1990)), (Abdel-Rahman and Fujita, 1990), or (Becker and Henderson, (1996)). Over time, firm efficiency and the benefits of larger cities will be enhanced by local human capital accumulation. In achieving contemporaneous equilibrium city size, scale benefits are traded off against the higher per person internal commuting (and potentially congestion and pollution) costs of supporting larger cities. We start by examining the structure of type 1 cities.

Production in Type 1 Cities

Consider a representative city of type 1. Each firm in the city is composed of one worker. Each period that worker decides how much to produce and how much to invest in private human capital accumulation. Having single worker firms is a convenience, so that human capital spillovers only exist across firms, not within firms.[2] Output of firm i of the numeraire intermediate input X_1 (to be sold to type 2 cities) is given by

$$X_{1i} = D_1[n_1^{\delta_1} h_1^{\psi_1}] h_{1i}^{\theta_1} \tag{2a}$$

where

$$W_{1i} = X_{1i}. \tag{2b}$$

In (2a), n_1 is employment in industry 1 in this city, h_1 is the average level of human capital of workers in the city, and h_{1i} is the human capital of the worker in firm i. δ_1 represents scale economies arising from the total volume of local communications which are proportional to n_1. δ_1 is the elasticity of firm output with respect to total local employment (holding own firm inputs fixed). ψ_1 is elasticity of firm i's output with respect to the average level of human capital in the city, which represents the spillover benefits of local levels

[2]If firms are multi-worker, of course, in theory the firm could solve the internal coordination problem of each employee's investment decisions by imposing employment requirements. We didn't add that complication here, because it is not relevant to the problem.

of human capital, or knowledge accumulation.[3] We use the average local level of human capital, rather than total local human capital, since scale economies are already captured in the $n_1^{\delta_1}$ term. $h_1^{\psi_1}$ could be thought of as representing the "richness" of the information spillovers $n_1^{\delta_1}$. Equation (2b) tells us that a worker i's wage in a type 1 city, W_{1i}, is simply the output of that worker.

Given all workers are inherently identical in a symmetrical equilibrium as developed later, within city type 1, $h_{1i} = h_1$. Total city output then is $n_1 * X_{1i}$ or

$$X_1 = D_1 n_1^{1+\delta_1} h_1^{\theta_1+\psi_1} \tag{3}$$

The specification of technology assumes human capital spillovers and scale externalities are purely localized. In addition, scale externalities are own industry, meaning that the presence of a different industry in the locality would not benefit the X_1 industry. In the analysis to follow, since agglomerating people into cities is costly on the commuting side, developers will form specialized cities, as in Henderson (1974). For the same population and commuting costs, a specialized city with greater per industry scale will have greater output per worker than a diversified city where each separate industry operates at a lower scale.[4]

Commuting. All production in a city occurs at a point, the Central Business District [CBD]. Surrounding the CBD is a circle of residences, where each resident lives on a lot of unit size, and commutes to the CBD (and back) at a constant cost per unit distance of τ (paid in units of type 1 city output). Adding in considerations of infrastructure investments, congestion, pollution and the like is critical for analyzing some features of urban growth but not the

[3] In specifying X_1 technology in equation (1), we use the simple specification in Romer (1986), where the effect of human capital on productivity is treated as a black box. One could also adapt the elaborate structure in Romer (1990) in specifying the technology in city type 1. One method of doing this would be to retain X_1 as an inter-city traded, competitive good, but have it produced with nontraded machinery inputs, as well as labor and private human capital investment. Nontraded machinery inputs enter in X_1 production in Dixit-Stiglitz (1977) fashion, generating local economies of scale (Abdel-Rahman and Fujita, 1990), and sold locally under monopolistic competition. However, the local span or degree of diversity of these nontraded machinery inputs would increase with local human capital accumulation.

[4] Of course, with inter-industry spillovers of communications that argument is weakened, although not actually eliminated provided δ's (applied to, say, all industries' total local employment) vary by industry.

ones in this paper. For tractability, without loss of generality, we use the simplest standard version of the internal spatial structure of cities. In this version, equilibrium in the land market is characterized by a rent gradient, declining linearly from the CBD to the city edge where rents (in the best alternative use) are zero. Standard analysis gives us expressions for total city commuting costs and rents in terms of city population where[5]

$$\text{total commuting costs} = bn_1^{3/2} \tag{4}$$

$$\text{total land rents} = \frac{1}{2}bn_1^{3/2} \tag{5}$$

$$b \equiv \frac{2}{3}\pi^{-1/2}\tau$$

Equation (4) is a critical resource cost to the city, where average commuting costs $(bn_1^{1/2})$ rise with city size, with an elasticity of 1/2. That is the force limiting city sizes. Equation (4) constitutes the gross rental income of the city developer.

City Developers. Type 1 cities form in the competitive context of a large economy with many type 1 cities in the national land market. Each city is operated by its developer who collects urban land rents, offers inducements to firms to locate in the city, and specifies city population (although people are free to move in equilibrium). Nationally, there are an unexhausted number of potential identical sites upon which cities can form, and each developer controls only one site. The resulting solutions can be obtained in other ways. In

[5] An equilibrium in residential markets requires all residents (living on equal size lots) to spend the same amount on rent, $R(u)$, plus commuting costs, τu, for any distance u from the CBD. Any consumer then has the same amount left over to invest or spend on all other goods. At the city edge at a radius of u, rent plus commuting costs are τu_1 since $R(u_1) = 0$; and elsewhere they are $R(u) + \tau u$. Equating these at the city edge with those amounts elsewhere yields the rent gradient $R(u) = \tau(u_1 - u)$. From this, we calculate total rents in the city to be $\int_0^{u_1} 2\pi u R(u) du$ (given lot sizes of 1 so each "ring" $2\pi u du$, contains that many residents), or $\frac{1}{3}\pi\tau u_1^3$. Total commuting costs are $\int_0^{u_1} 2\pi u(\tau u) du = \frac{2}{3}\pi\tau u_1^3$. Given city population of n, and lot sizes of 1, $n_1 = \pi u_1^2$ or $u_1 = \pi^{-1/2}n^{1/2}$. Substitution gives us equations (4) and (5).

a static context, Henderson and Becker (1996) show this solution (1) is the only coalition-proof equilibrium and (2) will occur also in a model with only "self-organization" where each existing city is governed by an autonomous local government. In a growth context, they show an equivalent formulation is that developers start, or set up new cities to maximize profits, while existing cities are taken over and governed by local autonomous governments seeking to maximize incomes of existing residents.

Within a representative city, the developer's profits are residential land rents (eq. (5)) less any transfer payments, T_1, to each worker/firm. The developer faces a free migration constraint that each worker's net income (after paying rents and commuting costs) equals the prevailing net income available in national labor markets to workers in other type 1 cities, I_1. The developer chooses city population, n_1, and transfer payments, T_1, to maximize current profits. Since there is only private human capital in the model, developers cannot, for example, invest in capital accumulation of residents. Later in the section on efficient growth we will relax this assumption (also, cf. Deo and Duranton, 1995). For now, each developer

$$\max_{n_1, T_1} \quad \Pi_1 = \frac{1}{2} b n_1^{3/2} - T_1 n_1 \tag{6}$$

$$\text{s.t.} \quad W_1 + T_1 - \frac{3}{2} b n_1^{1/2} = I_1$$

where from (1a) $W_1 = D_1 n_1^{\delta_1} h_1^{\theta_1 + \psi_1}$, given symmetry within the city. In the constraint the first term is per worker firm income and the third term is per resident rent plus commuting costs anywhere in the city from (4) and (5).

Solving (6), substituting for T_1 back into Π_1 (6), setting $\Pi_1 = 0$ (through folk theorem free entry of developers/cities in national land markets), together yields two key traditional results:

$$T_1 = \frac{1}{2} b n_1^{1/2} \tag{7}$$

$$n_1 = [\delta_1 2 b^{-1} D_1]^{\frac{2}{1-2\delta_1}} \; h_1^{2\epsilon_1} \tag{8}$$

for

$$\epsilon_1 \equiv \phi_1 + \frac{\psi_1}{1-2\delta_1}, \; \phi_1 \equiv \frac{\theta_1}{1-2\delta_1} < 1.$$

Equation (7) states the "Henry George Theorem" (Flatters, Henderson, and Mieszkowski (1974) and Stiglitz (1977)), that total transfers to firms $(T_1 n_1)$ equal total urban land rents $(\frac{1}{2} b n_1^{3/2})$. The per worker transfer closes the gap between private marginal product (eq. (2)) and social marginal product due to enhanced scale benefits when a worker/firm enters a city, where the gap is $\delta_1 W_1$.

Equation (8) tells us equilibrium city size is a function of scale and other parameters and per worker human capital. Second-order conditions and equation (8) reveal the parameter restriction $\delta_1 < 1/2$, necessary to have multiple type 1 cities in the economy. Equation (8) shows city sizes increase as the scale elasticity, δ_1, rises towards the commuting cost elasticity, 1/2, from equation (4). For $\delta_1 > \frac{1}{2}$, all X_1 production would occur in just one city, because marginal scale benefits of increasing city size would always outweigh marginal costs.

In terms of the relationship between city size and human capital, $2\epsilon_1$ defines the elasticity of city size with respect to per worker human capital, which is increasing in the private (θ_1) and external (ψ_1) elasticities of productivity with respect to human capital. ϵ_1 is decomposed into a private return portion ϕ_1, and an externality return portion $\frac{\psi_1}{1-2\delta_1}$, according to the private, θ_1, and spillover, ψ_1, returns to human capital. Regularity (see later) requires $\phi_1 < 1$. ϵ_1 rises as the degree of scale economies, δ_1, rises towards the commuting resource cost elasticity, 1/2. Scale benefits augment human capital returns. As we will see next in equation (9), ϵ_1 is the elasticity of net income in a city with respect to average human capital levels. With human capital accumulation not only do incomes rise, but so do city sizes.

For later use, we solve for wages, W_1, and net income, I_1 by substitution of (7) and (8) into (5) and (2), as well as city output X_1. For Q_1 a parameter cluster,[6]

$$I_1 = (1 - 2\delta_1)W_1 = Q_1 h_1^{\epsilon_1} \tag{9}$$

Type 2 Cities

Type 2 cities specialize in production of the economy's consumption good, sold in competitive national markets at a price, P. A single worker firm's output is

$$X_{2j} = D_2 \left[n_2^{\delta_2} h_2^{\psi_2}\right] h_{2j}^{\theta_2} x_{1j}^{1-\alpha} \tag{10}$$

Corresponding to equation (2) for X_{1i}, in equation (10) for X_{2j} are external scale $(n_2^{\delta_2})$ and human capital $(h_2^{\psi_2})$ terms. h_{2j} is worker j's human capital. x_{1j} is the jth firm's use of imported intermediate inputs of type 1 cities. Profits for a firm are $PX_{2j} - x_{1j}$. Maximizing and substituting in $PX_{2j} - x_{1j}$ for the choice of x_{1j} gives the residual return to the worker-firm:

$$W_2 = \alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}} D_2^{1/\alpha} P^{1/\alpha} (n_2^{\delta_2} h_2^{\psi_2})^{1/\alpha} h_{2j}^{\theta_2/\alpha} \tag{11}$$

As for type 1 cities, developers of type 2 cities choose T_2 and n_2 to max $\Pi_2 = \frac{1}{2}bn_2^{3/2} - T_2 n_2$ s.t. $W_2 + T_2 - \frac{3}{2}bn_2^{1/2} = I_2$. W_2 is given in (11) for $h_{2j} = h_2$ under symmetry, and commuting and rents for a representative type 2 city are derived as for type 1 cities. Note commuting costs and rents are paid and enumerated in units of X_1, the numeraire good. We solve this problem as before, maximizing, setting Π_2 equal to zero, and solving for

[6] $Q_1 \equiv (\delta_1 2b^{-1} D_1)^{\frac{1}{1-2\delta_1}} b(2\delta_1)^{-1}(1-2\delta_1)$ and $X_1 = [(Q_1/(1-2\delta_1))^{\frac{1+\delta_1}{\delta_1}} D_1^{-\delta_1}] h_1^{3\epsilon_1}$

the Henry George result, $T_2 = \frac{1}{2} b n_2^{1/2}$. With substitutions, we have equations corresponding to (8) and (9) defining equilibrium city size and income:[7]

$$n_2 = C_2 \, P^{\frac{1}{(\alpha/2) - \delta_2}} h_2^{2\epsilon_2}, \qquad \delta_2 < \frac{\alpha}{2} \tag{12}$$

where

$$\epsilon_2 \equiv \phi_2 + \frac{\psi_2}{\alpha - 2\delta_2}, \quad \phi_2 \equiv \frac{\theta_2}{\alpha - 2\delta_2} < 1.$$

$$I_2 = (\alpha - 2\delta_2)\alpha^{-1} W_2 = Q_2 P^{\frac{1}{\alpha - 2\delta_2}} h_2^{\epsilon_2} \tag{13}$$

While these expressions have similar properties to those for type 1 cities, they contain the relative price P. We need first to determine migration and human capital investment decisions based on the family's problem in (1) which we are now ready to solve, given we know I_1 and I_2. Then we can solve for P in national output markets and proceed to growth properties.

<u>Investment and Migration Decisions</u>

Given the family's dynamic optimization in (1), we form the Hamiltonian, ignoring for constraints (c), which we show in the Appendix are never binding. A representative family's problem is

$$\max_{c, z, h_1, h_2, H} \mathcal{L} = \frac{c^{1-\sigma} - 1}{1 - \sigma} e^{-(\rho - g)t} \tag{14}$$

$$+ \quad \lambda_1 [z e^{gt} I_1 P^{-1} + (1 - z) e^{gt} I_2 P^{-1} - c e^{gt}]$$

$$+ \quad \lambda_2 [H - z e^{gt} h_1 - (1 - z) e^{gt} h_2]$$

[7] Firm level $x_1 = (1 - \alpha) Q_2 (\alpha - 2\delta_2)^{-1} P^{\frac{1}{\alpha - 2\delta_2}} h_2^{\epsilon_2}$ $C_2 \equiv [(1 - \alpha)^{\frac{1-\alpha}{\alpha}} \delta_2 2 b^{-1} D_2^{1/\alpha}]^{\frac{\alpha}{\alpha/2 - \delta_2}}$ and $Q_2 \equiv [(1 - \alpha)^{\frac{1-\alpha}{\alpha}} \delta_2 2 b^{-1} D_2^{1/\alpha}]^{\frac{\alpha}{\alpha - 2\delta_2}} b (2\delta_2)^{-1} (\alpha - 2\delta_2)$.

In (14), as perceived by family i, for workers in city type 1, based on equations (6) and (2a), $I_{1i} = W_{1i} + T_1 - 3/2bn_1^{1/2}$ for $W_{1i} = D_1(n_1^{\delta_1} h_1^{\psi_1}) h_{1i}^{\theta_1}$ and T_1, h_1 and n_1 perceived as fixed by the family. Thus $\partial I_{1i}/\partial h_{1i} = \theta_1 W_{1i}/h_{1i}$. Then imposing symmetry $(h_{1i} = h_1)$ and using equations (8) and (9), in equilibrium, the value of $\partial I_{1i}/\partial h_{1i} = \frac{\theta_1}{1-2\delta_1} I_1 h_1^{-1} = \phi_1 I_1 h_1^{-1}$. Similarly in city type 2 $\partial I_{2i}/\partial h_{2i} = \frac{\theta_2}{\alpha - 2\delta_2} I_2 h_2^{-1} = \phi_2 I_2 h_2^{-1}$. The first-order conditions for (14) imposing symmetry after differentiation are

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial c} &= c^{-\sigma} e^{-(\rho - g)t} - \lambda_1 e^{gt} = 0 && (a) \\
\frac{\partial \mathcal{L}}{\partial z} &= e^{gt}[\lambda_1 (I_1 P^{-1} - I_2 P^{-1}) + \lambda_2(-h_1 + h_2)] = 0 && (b) \\
\frac{\partial \mathcal{L}}{\partial h_1} &= z e^{gt}[\lambda_1 I_1 h_1^{-1} P^{-1}\ \phi_1 - \lambda_2] = 0 && (c) \qquad (15) \\
\frac{\partial \mathcal{L}}{\partial h_2} &= z e^{gt}[-\lambda_1 \phi_2 I_1 h_2^{-1} P^{-1} + \lambda_2] = 0 && (d) \\
\frac{\partial \mathcal{L}}{\partial H} &= -\dot{\lambda}_1 = \lambda_2 && (e)
\end{aligned}
$$

The transversality condition requires

$$\lim_{t \to \infty} [\lambda_1(t) H(t)] = 0 \qquad (f) \qquad (15)$$

In (15) (c) and (d), families allocate human capital across city types to equalize private returns on investment. Combining the two yields $I_1/I_2 = (\phi_2/\phi_1) h_1/h_2$, which then, when combined with the result from solving λ_2/λ_1 in (15b) and (15c) yields

$$
\begin{aligned}
h_1 &= \left[\frac{\phi_1(1-\phi_2)}{\phi_2(1-\phi_1)}\right] h_2 && (a) \\
I_1 &= \left(\frac{1-\phi_2}{1-\phi_1}\right) I_2 && (b).
\end{aligned}
\qquad (16)
$$

Note the time invariant ratios of h_1/h_2 and I_1/I_2. To proceed further to solve for z, the relative allocation of family members by city type, we need to examine equilibrium in national markets.

National Market Equilibrium

Equilibrium in national output markets requires a balance of trade among cities, or national demand and supply of X_1 to be equalized. National supply of X_1 is $m_1 X_1$ where

m_1 is the number of type 1 cities. The m_2 type 2 cities import X_1 as an intermediate input x_1 and X_1 is used to produce commuting (equation (4)) in both types of cities. Trade balance requires $m_1X_1 = m_2n_2x_1 + m_1(bn_1^{3/2}) + m_2(bn_2^{3/2})$. Then imposing symmetry across dynasties nationally, so each dynasty sends the same proportion of workers to each type of city, we then know at any instant, $z = m_1n_1/N$ and $1 - z = m_2n_2/N$, where N is national population. Combining all relationships yields[8]

$$z = \frac{(1-\phi_1)(1-\alpha+2\delta_2)}{(1-\phi_1)(1-\alpha+\delta_2)+(1-\phi_2)(\alpha-2\delta_2)}. \tag{17}$$

Any family's proportion of workers z, going to type 1 city in equilibrium is invariant to $h(h_1$ or $h_2)$ and is constant over time. All workers once assigned a city type never need to change that type. Migration typically only involves assignment of newborns, especially to new cities (see below). Equation (17) together with $z = m_1n_1/N$ tell us the number of cities of each type at any instant. Specifically,

$$m_1 = zNn_1^{-1}; \quad m_2 = (1-z)Nn_2^{-1}. \tag{18}$$

Having solved for z which reflects migration decisions, we can solve for capital usages, h_1 and h_2, as functions of per person family stock, h. Given $h = He^{-gt}$ from (1b), $h = zh_1 + (1-z)h_2$. Substituting in (16) and (18) yields[9]

$$h_2 = \frac{\phi_2}{1-\phi_2}Kh, \; h_1 = \frac{\phi_1}{1-\phi_1}\,Kh \tag{19}$$

---

[8] Rearranging demand equals supply, $m_1(X_1 - bn_1^{3/2}) = m_2n_2(x_1/n_2 + bn_2^{1/2})$. From (6), (1) and (7), $n_1I_1 = X_1 - bn_1^{3/2}$. From (11), (12) and footnote (9), we know $x_1/n_2 + bn_2^{1/2} = I_2(1-\alpha+2\delta_2)/(\alpha-2\delta_2)$. Combining these relationships gives $m_1n_1I_1 = m_2n_2I_2(1-\alpha+2\delta_2)/(\alpha-2\delta_2)$. Substituting this and equation (16) with the expressions for z and $(1-z)$ gives (17).

[9]
$$K \equiv \frac{(1-\phi_1)(1-\alpha+2\delta_2)+(1-\phi_2)(\alpha-2\delta_2)}{\phi_1(1-\alpha+2\delta_2)+\phi_2(\alpha-2\delta_2)}.$$

Equation (17) directly gives us an unchanging relative allocation of family members by city type. Below we will show that the constraint $\dot{H} \geq 0$ is generally satisfied along equilibrium growth paths. From (16) and (22) by time differentiating, $\dot{h}_1/h_1 + g = \dot{h}_2/h_2 + g = \dot{h}/h + g = \dot{H}/H \geq 0$. Human capital grows in parallel at the same rate in the two types of cities. The only capital transfers need be from each worker type to their own children. In our equilibrium human capital can be nontransferable across existing people and specific to a technology (either X_1 or X_2). In (1c) constraints are never binding.

Finally, by combining various relationships we get[10]

$$P = Qh^{(\epsilon_1 - \epsilon_2)(\alpha - 2\delta_2)}. \tag{20}$$

In (20) as h grows, the relative price of the consumption good rises if $\epsilon_1 > \epsilon_2$. Not surprisingly, P rises so the consumption good becomes more expensive, if the elasticity of income in the numeraire good city with respect to human capital exceeds that in the consumption type city.

The results in this section are summarized in

Proposition 1. Over time, the equilibrium allocation of resources across cities involves the following characteristics:

(a) The ratios of per person human capital and income, h_1/h_2 and I_1/I_2, are time invariant. This implies persistent measured income inequality where $I_1 > [<] I_2$ and $h_1 > [<] h_2$ iff $\phi_1 > [<] \phi_2$, where ϕ_i is the private return on income to human capital investment in city type i.

---

[10] We combine (16) with $z = m_1 n_1/N$ and $1 - z = m_2 n_2/N$, and do substitutions.

$$Q \equiv \left[\frac{\phi_1 Q_1}{\phi_2 Q_2} \; \frac{(\frac{\phi_1}{1-\phi_1})^{\epsilon_1 - 1}}{(\frac{\phi_2}{1-\phi_2})^{\epsilon_2 - 1}} \; K^{\epsilon_1 - \epsilon_2} \right]^{\alpha - 2\delta_2}.$$

(b) The relative allocation of population across cities is time invariant.

(c) The price, P, of X_2, the consumption good rises [falls] with human capital accumulation iff $\epsilon_1 > [<] \epsilon_2$, where ϵ_i is the social return to human capital investment in city type i.

We can now solve for urban and economic growth features of the economy.

Urban Growth

Although the per person human capital levels employed in each type of city differ at any instant, as we just saw, human capital in each type of city grows at the same rate, or $\frac{\dot{h}_1}{h_1} = \frac{\dot{h}_2}{h_2} = \frac{\dot{h}}{h}$. Then in the city size equations (8) and (12) combined with (20) for P, we know

$$\frac{\dot{n}_2}{n_2} = \frac{\dot{n}_1}{n_1} = 2\epsilon_1 \frac{\dot{h}}{h}. \tag{21}$$

Individual city sizes grow with human capital accumulation at a rate $2\epsilon_1$ times the rate of human capital accumulation. Recall ϵ_1 is the elasticity of income with respect to human capital levels in city type 1. Below we will distinguish two cases. One is where the economy experiences steady-state growth, in which case ϵ_1 is generally close to one. Then city sizes grow at approximately twice the rate of human capital accumulation. Through externalities raising the marginal benefits of adding population to cities relative to the marginal costs, human capital accumulation enhances per worker productivity directly and indirectly sufficiently to cause cities to grow at about twice the rate of capital accumulation.

What about growth in the number of cities, m_1 and m_2? From (18) defining m_1 and m_2, for example, $\dot{m}_1/m_1 = \dot{N}/N - \dot{n}_1/n_1$, where national population growth $((\dot{N}/N)$ is g and $\dot{n}_1/n_1$ is given by (21). Thus

$$\frac{\dot{m}_1}{m_1} = \frac{\dot{m}_2}{m_2} = g - 2\epsilon_1 \frac{\dot{h}}{h}. \tag{22}$$

City numbers increase with human capital accumulation, if the rate of individual city size growth fueled by human capital accumulation is not high enough to accommodate the expanding national population growth rate. Regardless, equations (21) and (22) imply by inspection

> Proposition 2. Urban growth across city types is parallel, maintaining a constant relative size distribution of cities. That is, relative sizes and numbers of the two types of cities are time invariant, with all cities growing in size at the same rate.

While this urban growth process seems simple, the underlying institutional and economy reality can be quite complex. Formation of an appropriate number of new cities at any instant conceptually seems to require "large" agents such as developers who set up new cities in a conducive institutional framework (Henderson and Becker, 1996). Absent such agents, cities in general will tend to be too large and too few in number. In fact, part of the problem of top-heavy urban development in some developing countries may be central government hindrance of effective functioning of land markets and local governments. Fortunately, the existence and widespread operation of developers who set up new cities seem, at least for the USA, to be a fact (e.g., Garreau, 1991). Formation of new cities in and of itself efficiently limits the contemporaneous sizes of existing cities; and, in theory, autonomous local governments in existing cities have the incentives to offer appropriate local subsidies, T_1 (equation (7)) and the corresponding T_2, to local businesses (Henderson and Becker, 1996). In summary, the process works if new cities are started by developers and existing cities have traditional USA-style local governments.

From equations (21) and (22) it appears once type 1 cities are set up they stay and grow as type 1 cities and the same for type 2. At each instant, new type 1 and type 2 cities form with newborns. However when we turn to the data in section 2, we will see that this does not appear to be the way the process actually works. If, say, type 1 cities are smaller than type 2 cities, empirically, new cities coming into existence then appear to all be smaller type 1 cities. Additional type 2 cities arise by type 1 cities transforming into type 2 cities. Given type 1 and type 2 cities operate with different per person human capital levels, that means converting type 1 cities must upgrade or downgrade human capital levels. With specific human capital, transforming the human capital base would require migration – exit of type 1 workers from transforming type 1 cities to new type 1 cities and entry of type 2 workers (who could be newborns).

The big question is why do type 1 cities transform to type 2 cities to accommodate growth in numbers of type 2 cities, rather than entirely new cities of both types forming? There seem to be two potential explanations. The first is conceptual and goes beyond the formal scope of this paper. We assumed all potential city sites in the economy are identical — offer identical (unspecified) natural public amenities such as climate, coastal location, harbor facilities, etc. In reality, there is a spectrum of site qualities. Models that start to deal with this problem (Upton, 1981 and also Henderson, 1988, pp. 71-73) appear to have two features to equilibrium. The best sites are occupied first and the best sites go to bigger types of cities, which can bid more for the amenities. Here that means as the number of cities grows, additional bigger type 2 cities outcompete existing type 1 cities (who initially got reasonable quality sites) for the sites they are on, and new type 1 cities form on the lowest quality sites occupied to date. The second explanation is institutional, although it can be specified to have market foundations (Helsley and Strange, 1993). Developers who start new cities have either or both limited financial resources and ability to assemble large pieces of land. It is thus "easier" for them to start new smaller types of cities. Later with growth these initial smaller types may transform into bigger types of cities. In both cases of the site quality and the limited size developer models, an issue concerns the transformation process. To enact mass conversion of production in a city to another type involves large scale movement/conversion of firms, which is not readily attained through atomistic behavior. With scale economies, local developers and/or local governments are needed to facilitate timely transformation (see Rauch, 1993).

Economic Growth

The final part to urban evolution is to solve for growth paths in the economy. Once in place, the analysis is fairly standard and our treatment is brief.

We need to study the representative family's evolution of consumption and human capital. For consumption, we first time differentiate (15a) and combine with (15c) and (15e) to get $\gamma^c \equiv \frac{\dot{c}}{c} = \frac{1}{\sigma}\,[\phi_1 I_1 h_1^{-1} P^{-1} - \rho]$. For the human capital growth path, we focus on the average level of human capital per member, h, where $\frac{\dot{h}}{h} = \frac{\dot{H}}{H} - g$, so from (14a) $\gamma^h \equiv \frac{\dot{h}}{h} = zI_1P^{-1}h^{-1} + (1-z)I_2P^{-1}h^{-1} - ch^{-1} - g$. Into these equations we substitute for I_1 and I_2, from (9) and (13), for h_1 and h_2 from (19), and for P from (20). The results are

$$\gamma^c \equiv \frac{\dot{c}}{c} = \frac{1}{\sigma}\,[Ah^{\epsilon-1} - \rho] \tag{23}$$

$$\gamma^h \equiv \frac{\dot{h}}{h} = Bh^{\epsilon-1} - ch^{-1} - g \tag{24}$$

where

$$\epsilon \equiv \epsilon_1 - (\epsilon_1 - \epsilon_2)(\alpha - 2\delta_2) > 0$$

$$A \equiv \frac{\phi_1 Q_1}{Q}\left(\frac{\phi_1}{1-\phi_1}\right)^{\epsilon_1 - 1} K^{\epsilon_1 - 1}, \quad B \equiv A[\phi_1(1-\alpha+2\delta_2) + \phi_2(\alpha - 2\delta_2)]^{-1}$$

$$A > B.$$ [11]

In the analysis to follow we distinguish between two cases: steady-state growth where $\epsilon = 1$ and steady-state levels where $\epsilon < 1$. Details of proofs and analysis of global stability and uniqueness are given in the Appendix A. The basic proposition is

> Proposition 3. If $\epsilon = 1$, the economy achieves steady-state growth, where consumption and human capital grow at the rate $(A-\rho)\sigma^{-1}$ and city sizes grow at $2\epsilon_1$ times this rate. If $\epsilon < 1$, the economy converges to steady-state levels of consumption and human capital and cities achieve a stationary size.

[11] $B/A = [\phi_1(1-\alpha+2\delta_2) + \phi_2(\alpha-2\delta_2)]^{-1} > 1$, given $z > 0$. $z > 0$ requires $1 > \phi_1(1-\alpha+2\delta_2) + \phi_2(\alpha - 2\delta_2)$, because the denominator of z can be written as $1 - [\phi_1(1-\alpha+2\delta_2) + \phi_2(\alpha-2\delta_2)]$.

If $\epsilon = 1$, by inspection of (23) the steady-state growth rate of c is $(A-\rho)/\sigma$. By differentiating (24), for γ^h to be constant, the steady-state growth rate of h must equal that of c. Denoting the steady-state growth rates as $\overline{\gamma}^c$ and $\overline{\gamma}^h$, then

$$\overline{\gamma}^c = \overline{\gamma}^h = \left(\frac{A-\rho}{\sigma}\right). \tag{25}$$

Positive steady-state growth with bounded utility and satisfaction of transversality conditions require[12]

$$\begin{aligned} &A-\rho > 0 && (a) \\ &\tfrac{1-\sigma}{\sigma}A + g - \tfrac{\rho}{\sigma} < 0 && (b) \end{aligned} \tag{26}$$

With steady-state growth, from equation (21) city sizes grow at a rate $2\epsilon_1\overline{\gamma}^h$ indefinitely. City numbers increase as long as the individual city population growth rate is less than the national population growth rate.

If $\epsilon < 1$, to solve for steady-state levels, we set $\gamma^h = \gamma^c = 0$ and solve

$$\begin{aligned} \overline{h} &= \left(\tfrac{A}{\rho}\right)^{\frac{1}{1-\epsilon}} \\ \overline{c} &= \left(\tfrac{A}{\rho}\right)^{\frac{1}{1-\epsilon}} \left(\tfrac{B\rho}{A} - g\right). \end{aligned} \tag{27}$$

Positive consumption requires $\frac{B\rho}{A} - g > 0$, which is guaranteed given $B > A$ and $\rho > g$. In the Appendix we show $\overline{h}, \overline{c}$ exhibits local saddle path stability and convergence is along a globally stable arm. Along the stable arm $\dot{H} > 0$. Moving upward along the stable arm, h is increasing and hence so are city sizes. However at steady-state levels, since h growth ceases, city sizes stagnate.

[12]For bounded utility, $\lim_{t\to\infty} \frac{c^{1-\sigma}-1}{1-\sigma}e^{-(\rho-g)t}dt = 0$. Given from (25), $c = c_0 e^{\frac{A-\rho}{\sigma}t}$, the limit requires (26b). Similarly for transversality $\lim_{t\to 0} \lambda_1(t)H(t) = 0$ where $H(t) = h(t)e^{gt}$ and $\lambda_1(t) = c^{-\sigma}e^{-\rho t}$ from (15a). Evaluating again requires (26b) to be satisfied.

Efficiency of Urban Resource Allocation

The equilibrium outcome is not optimal. The problem lies with the externalities involved in human capital accumulation decisions and with the corresponding population allocation decisions made by families. The key is the gap between ϵ_i and ϕ_i, where $\epsilon_1 = \phi_1 + \frac{\psi_1}{1-2\delta_1}$ and $\epsilon_2 = \phi_2 + \frac{\psi_2}{\alpha-2\delta_2}$. Families invest based on net private returns, ϕ_i, rather than the net social returns ϵ_i, ignoring spillover returns ψ_i.

In Appendix A, we solve a national social planner's optimization problem in equation (1) for a representative family, where the planner accounts for overall social marginal returns to human capital investment in cities. The solution to the planner's problem has two key aspects. First, the time invariant ratios h_1/h_2, I_1/I_2, $z/(1-z)$, and n_1/n_2 all change; and, secondly, the growth process itself also changes. Of particular interest is

> Proposition 4. With an efficient allocation of resources, compared to the equilibrium, steady-state growth rates (when $\epsilon = 1$) of consumption and capital and steady-state levels (when $\epsilon < 1$) of per person consumption and capital are higher. In terms of allocations to type 1 versus type 2 cities, h_1/h_2 and I_1/I_2 rise [fall] respectively iff $\frac{\epsilon_1}{\epsilon_2}\frac{(1-\epsilon_2)}{(1-\epsilon_1)} > [<] \frac{\phi_1}{\phi_2}\frac{(1-\phi_2)}{(1-\phi_1)}$ and $(\frac{1-\epsilon_2}{1-\epsilon_1}) > (\frac{1-\phi_2}{1-\phi_1})$.

Proofs are in Appendix A. But the intuition is easy. In Proposition 4, since the social returns to capital exceed the private, the efficient solution involves greater capital accumulation. In that solution, the ratios of incomes and capital in the two types of cities depend on social not private returns. But whether, say, I_1/I_2 rises does not simply depend on just whether $\frac{\psi_1}{1-2\delta_1} > \frac{\psi_2}{1-2\delta_2}$, but depends on the initial position of ϕ_1 vs. ϕ_2. So I_1/I_2 rises if $(1-\phi_2)\frac{\psi_1}{1-2\delta_1} > (1-\phi_1)\frac{\psi_2}{\alpha-2\delta_2}$. In comparing relative city sizes (see Appendix) what happens to n_1/n_2 involves a complex expression.

Is there a mechanism in a decentralized market economy to achieve an efficient allocation of resources? Here we present a hypothetical mechanism and then discuss its feasibility. The key is for localities to account for the local human capital spillover benefits, represented by $h_i^{\psi_i}$ in production functions. If they do so, then market allocations will be efficient. In particular, we have

Proposition 5. If city developers can invest in and claim the returns on local human capital accumulation at the margin, they will internalize the local human capital spillovers. Contemporaneous market allocations and the growth process will both be optional.

Suppose any developer can "borrow" human capital in national markets, so her objective function is $\frac{1}{2}bn_i^{3/2} - T_i n_i - rh_i n_i$ where the first term is local land rents, the second tax/subsidies on local employees, and the third capital rental payments for r the human capital rental rate in national markets. Families still accumulate and own human capital stocks, now rented out in national markets. Each developer sees the same labor mobility constraint as before, so for type 1 cities, $I_1 = D_1 n_1^{\delta_1} h_1^{\psi_1+\theta_1} + T_1 - 3/2bn_1^{1/2}$; and each developers chooses h_i and T_i and specifies n_i. Following the solution in Appendix A, we continue to have the Henry George Theorem that local rents equal scale externalities; but now $T_i = 1/2bn_i^{1/2} - rh_i$, where the first term is per person land rents and the second per person human capital rents. Local worker-firms pay for the human capital borrowed by the developer for their use.

In Appendix A, we show that the resulting allocation of human capital and workers across cities is efficient and the accumulation process is optimal. However, due to the nature of human capital it is unlikely that this result can be obtained. Decentralized achievement of optimal growth depends on the ability of city developers (or local governments) to specify and collect returns on local human capital levels. If human capital can only be used in the city in which it is acquired, something which raises the specter of localities having monopsony power in local labor markets, then local governments could provide schooling and/or finance on the job training. Otherwise, if human capital is generally applicable across cities, then the no-slavery constraint makes it difficult to have a "market" for human capital. Individuals in whom a city invests are free to pick up and leave with their individuals h_i's, taking the external benefits with them.

Extensions to an OLG Framework

Our results on the allocation of resources across cities extent directly to a simple overlapping generations framework outlined here. Each generation lives two periods. In the first period, each person lives with their parents and invests in human capital. In the second period, they choose a city in which to live and work, receive their wage, consume, and make bequests to their children. People have a "joy of giving" type bequest motive, with lifetime utility given by $\beta \ln c + (1-\beta) \ln q$, where c is consumption in the second period of life and q is the bequest given to one's children. There are two constraints. First, total lifetime income from working and from bequests received from parents must equal expenditures on consumption, bequests, and human capital investment. Second, lacking a market for human capital, first period investment is constrained by the level of the received bequest. If the latter constraint isn't initially binding for workers intending to locate in either type of city, the economy immediately goes to steady-state levels, with allocations satisfying equations (16) - (18). As in the dynastic family formulation, observed incomes vary by city type, but in equilibrium the proportion of workers living in each city type adjusts to equate incomes net of human capital expenditures, equalizing consumption and bequest levels as well.

In this context, the specification of intergenerational choices and preferences bears critically on inequality issues. For example, if we modify the model in the previous paragraph so that parents choose the educational level of their children investing in human capital specific to their city type and parents care about their children's income level (rather than the joy of giving), the economy evolves with real inequality. The locational choices of the equally endowed initial generation will lead to inequality in human capital levels, income, and consumption of future generations by city type.

In a more general context, an OLG model with a simple bequest motive provides a framework to explore issues of how children's initial characteristics — wealth, ability, and education are influenced by decisions of their parents and characteristics of their environment. If we incorporate simultaneously endogenous location and education choices into a model with parental and peer-group influenced human capital and ability differences, a richer positive analysis of evolving income inequality resulting from location choices will be possible. The basic questions are: how do heterogeneous agents sort themselves out spatially in an economy and how do these decisions affect future generations?

2. Empirical Evidence on Urban Growth

The model in section 1 predicts on-going growth of individual city populations, as long as there is human capital accumulation. The model utilized has also a feature of parallel growth — growth which maintains the same relative size distribution of cities. It leaves open the question of whether, with national population growth, there is positive or negative growth in the number of cities or not. In this section we examine historical patterns of urban growth in the USA in order to analyze the process of individual city growth, the evolution of the size distribution of cities, and the entry and/or exit of cities.

To analyze these issues, we need a data set which utilizes a consistent sensible definition of cities over time. Since we examine cities as economic, not political units, we want to use a metropolitan area concept, where all contiguous urbanized economic activity in an area is lumped together. In the USA metro areas are based on clustering contiguous urbanized counties. Since counties can be very large, a second step is to define what portions of these counties are urban versus rural, particularly for historical contexts when much of these county populations were rural. Finally, a cut-off point defining the minimum size for a metro area must be drawn.

The US Census Bureau didn't start using metro area concepts until 1950. Since 1950 the definitions of metro areas and of urban versus rural have changed significantly several times, so that comparisons of size distributions for years since 1950 are problematical. However, Donald Bogue published a consistent data set in 1953 for the period 1900-1950, based on the 1950 definition of metro areas. He takes the 162 metro areas in 1950 and follows them back in time to 1900. For each decade he calculates their total and urban populations based on a consistent definition of urban — the 1940 urban definition.

For the period 1900-1950 in the USA we analyze the evolution of cities. The population in urbanized metro areas, quadrupled in the 50 years and there was rapid human capital accumulation, as noted in the introduction. It is also a turbulent period, in which there are rapid technological developments, a major economic depression, and two world wars. It will be interesting to observe the evolutionary process of cities during such turbulence.

As an economy evolves, national output composition changes from decade to decade. What cities produce and specialize in, and the numbers of different types of cities will change with the introduction of new products or taste changes. While these items were not part of the model in section 1, we did discuss the notion of transformation of individual city industrial bases. Regardless, while there may be "parallel growth" in the sense that the relative size distribution of cities remains constant over time, individually cities are going to move around in that distribution.

The examination of Bogue's data in the first part of this section will focus on both the transitions and evolution of the distribution. In the second part of this section we establish that (1) cities specialize to some considerable extent, (2) different types of cities tend to have different sizes, and (3) transitions of individual cities are correlated with significant changes in production patterns.

Evolution of the USA Urban System 1900-1950

Following Amemiya (1985), Quah (1993), and Dobkins and Ioannides (1995), we divide the relative size distribution of cities at a point in time into cells, with cut-off points defined for sizes relative to the contemporaneous mean. These cut-off points for six cells are listed in Table 3. f_t denotes the density distribution of the relative sizes $n_{it}/\overline{n}_t$, of cities i at time t, where $\overline{n}_t = \sum_i n_{it}/m_t$ for m_t the number of cities at time t. This distribution evolves over time according to

$$f_{t+1} = (1 - i_t)M_t f_t + i_t Z_t \tag{28}$$

M_t is the transition matrix of existing cities from time t to $t+1$. M_t is hypothesized to be a homogenous stationary first-order Markov process, so M_t equals a common M. i_t is the overall entry rate of new cities between t and $t+1$, and is the increase in number of cities between t and $t+1$, divided by the number of cities at $t+1$. Z_t is the distribution of entrants across cells (the $t+1$ cells) — empirically almost all entry is in the extreme left cell (cell for smallest city sizes).

For a constant i and Z the steady-state distribution, f, is

$$f = [I - (1 - i)M]^{-1} iZ. \tag{29}$$

If we start in 1900 with an arbitrary distribution f_{1900}, the evolution of the system to 1950, for constant M, i, and Z, is

$$f_{1950} = (1 - i)^5 M^5 f_{1900} + [I - (1 - i)M]^{-1}[I - ((1 - i)M)^5] iZ. \tag{30}$$

We will perform both calculations, to compare these hypothetical distributions with the actual 1950 distribution. For i and Z, we use the overall entry rate and its distribution for 1900-50. These calculations, which result in predictions for 1950 of what should be found, are also a casual test of our specification. If the assumption that the system is evolving as a homogeneous first-order stationary process is justified, predictions should be close to outcomes.

In application, there is the problem that the 162 metro areas of 1950 are based on a cut-off point of 50,000 urban residents. In looking at these 162 metro areas in earlier decades, we can either employ the same absolute cut-off point of 50,000 or use a relative cut-off point. We focus on the latter. Since city sizes are increasing over time and since distributions are defined for current size relative to current mean, an absolute cut-off point means effectively truncating the distribution at different (relative) points over time. Instead, we choose to truncate at the same relative point: 50,000 to the 1950 mean of 417,000, or 0.12.[13] For 1900, the cut-off city size is 23,600. A (reasonable) assumption throughout is that in earlier decades, there exist no cities other than the future 1950 metro areas which have urban populations above the cut-off points for our sample.

Results. Table 1 gives some basic facts about the evolutionary process in Column 1 contains the decade percent increase in average sizes of metro areas (for all 162 metro areas, to avoid enlarging the sample each decade). Except for the 1930's, average sizes increase substantially. Column 2 gives the smaller percentage increase in numbers of cities. Overall in thc 50 years while sizes almost triple (for all 162 cities),[14] numbers increase by 41%. Column 3 gives the increase in city numbers using an absolute cut-off point.

Column 4 looks at the subsample of cities which experience significant population losses between decades. Column 4 is looking for "catastrophes" envisioned by Krugman (1993) — high losses from cities splitting in two or dropping precipitously in population in his city formation process. No cities drop from the sample, given cut-off points. Of the 693 city-decade points in the sample, only 40 have absolute urban population losses, most in the Depression involving urban people moving back to rural portions of counties. The maximum decade percent losses over the 50 years for 693 city-decades are 25%, 22%, 16%, and then 10%. Column 4 shows there are only 10 cases of losses over 5%.

In Table 2, we present maximum likelihood estimates, under stationarity of the transi-

[13]In choosing each decade's cut-off point for the base 162 1950 metro areas, we rank cities by size, from highest, 1, to lowest 162 in each decade: We then choose the first s cities as our city set for that decade, such that

$$\min\{s;\ n_{s+1}/\left(\sum_{i=1}^{s+1} n_i/(s+1)\right) > 0.12\}.$$

[14]The 162 cities increase in average size by 290% from 1900 to 1950. The original 115 cities increase in average size by 274%.

tion matrix, for transition probabilities, p_{ij}. p_{ij}, is simply the total number of cities moving from cell i to cell j over the fifty years divided by the number of cities in cell i in each decade summed over all decades (1900-1950). Stationarity is tested against nonstationarity (separate, decade-by-decade transition probabilities); and the restricted model (stationarity) cannot be rejected.[15]

In the transition matrix, as usual, the diagonal numbers are highest, being particularly high for cell 6. There is little mobility from the largest size downwards (up in a column). In terms of off-diagonal elements, there are significant probabilities of moving up a cell (down a column) each decade for at least the four lowest cells. The reason for this comes from the last column, giving the entry rate per decade by cell. 95% of entries occur in the lowest cell. To maintain a stable overall relative size distribution (see below), entrants push existing cities in the lowest cell forward, creating a chain reaction as some cities in the next to lowest cell are also pushed forward into a higher cell. This reaction gives the high upward transition probabilities for the first four cells. Without entry, such high upward transition probabilities would suggest the evolving distribution would end up more concentrated in the higher cells. But with entry, this does not happen.

In Table 3, in column (iii), we show the predicted steady-state (equation (29)) and, in column (iv), the predicted 1950 distribution as evolved from an arbitrary 1900 distribution (equation (30)). Column (v) shows the actual 1950 distribution. Note how close the predicted distributions are to the actual, especially the predicted 1950 distribution. This suggests that analyzing urban evolution as a first-order homogeneous stationary Markov process is quite reasonable.

[15]The χ^2 statistic is

$$-2log[\Pi_t \Pi_i \Pi_j \left[\frac{\hat{p}_{ij}}{\hat{p}_{ij}(t)}\right]^{m_{ij}(t)}]$$

with $(T-1)K(K-1)$ degrees of freedom, where $\hat{p}_{ij}$ is the stationary estimate, $\hat{p}_{ij}(t)$ are the decade-by-decade estimates, $m_{ij}(t)$ is the number of cities moving from i to j in decade t, T is the total number of years, and K the number of cells. The χ^2 statistic is 77.1, where with 120 degrees of freedom the critical value is over 146.6.

For comparison, we show two other predicted distributions. Column (vi) gives the steady state using an absolute cut-off point to define metro area sizes. It tends to over-emphasize concentration at the low end (given a much higher average rate of "entry" as the relative truncation point shifts left over time). Column (vi) shows the steady-state, for transition matrices calculated without entry, so each decade is all 162 metro areas, including seven areas in 1990 with zero urban population (and excluding 1900 areas with positive urban populations, which did not evolve into metro areas). Without entry but still fairly high mobility transition probabilities in lower cells (since by definition, our included 1900 small/zero population metro areas are fast growing), we predict a more uniform steady state with much less concentration of cities in lower cells. This is a poor prediction. If we are drawing a sample by picking 1950 metro areas and looking back at their evolution (only), it is critical to model entry. If one wants to model transitions without entry, the appropriate alternative is to represent the entire US geography every decade, thus including all excluded 1900 cities which did not evolve into 1950 metro areas and model transitions for the entire geography. Unfortunately, there is no existing data set to do this.

Urban Specialization and Transformation

There is no consistent data set on industrial composition of cities prior to 1950. However, from County Business Patterns, we have county data for recent years. We look at current industrial composition of cities to establish the concept of relative urban specialization and look at changes in industrial composition as they relate to changes in cities' positions in the relative size distribution of cities.

Typing of Cities. The common method of typing cities is to use cluster analysis on employment data (see Bergsman, Greenston, and Healy (1972) and Henderson (1988, pp. 12-30)). Prior cluster analysis based on data from around 1970 shows very clear typing of cities by manufacturing activity. Since 1970, manufacturing has declined from over 28% of U.S. non-government employment to under 19%, with many cities losing their manufacturing bases.

So many cities are in transition and classifying the industrial base of transforming cities at a point in time is problematical. Nevertheless, clear patterns emerge.

We classify cities based on their degree of specialization by two-digit SIC activities for private employment in all industries in 1992. We distinguish 317 metro areas using 1990 definitions (by county groupings) of MSA's and PMSA's. The degree of specialization in a particular industry i in city j is measured by the share of industry j in total local private employment, s_{ij}. There are 80 two-digit industries (after dropping the share of the last (99–), non-independent industry). Cities are grouped on the basis of similarity of production patterns, indicated by employment shares of different industries. In clustering cities we use Ward's criterion, grouping cities to minimize the error sum of squares within clusters summed across all clusters. The criterion function is

$$\sum_{c=1}^{n} \delta_c \left(\sum_{j=1}^{m_c} \sum_{i=1}^{80} \left(s_{ijc} - \frac{\sum_{j=1}^{m_c} s_{ijc}}{m_c} \right)^2 \right)$$

where m_c is the number of cities in cluster c and δ_c takes a value 1 if the city is in cluster c and zero otherwise. The clustering algorithm is hierarchical (step-wise) and the number of clusters is chosen in advance (there generally being no globally optimal number of clusters, nor general algorithms that are non-hierarchical).

Table 4 summarizes the cluster results, given in Appendix B in more detail. We originally specified 50 clusters, but broke apart three of the last five clusters formed (in the hierarchy) to distinguish eight sub-clusters among those three, so as to better classify the largest metro areas. In Table 4 there are 8 cluster groups, for the 55 clusters, arranged in some rough product-city categories (not in hierarchical order). We conducted an F-test on whether the clusters in Appendix B are indeed distinct from each other, and strongly reject the null hypothesis that the clusters are similar.[16]

[16] The F-test examines whether the model sum of squares increases significantly (relative to the overall residual sum of squares) from a model where s_{ij}'s differ only by industry compared to a model where they differ by industry clusters. $N = 317 * 80 = 25360$. For 4320 and 20960 degrees of freedom (noting there are $55 * 80 = 4400$ industry-clusters), $F = 11.3$, with a critical value near 1.

In Appendix B, for each cluster, where relevant, we list the dominant industry and report its average s_{ij}, compared to the national $\overline{s_i}$ for that industry. Given 60-70% of local employment is in "nontraded" good activity, individual s_{ij}'s over .05 represent significant specialization. We illustrate with two cluster members for each cluster. Apart from that, Appendix B and Table 4 are self-explanatory; but we make several comments. First, while we focus on industries cities are relatively specialized in, many clusters are also based on what cities don't have in the cluster — e.g., an absence of heavy manufacturing. Second, despite the national move away from manufacturing, some cities are still strongly specialized in manufacturing activity (in particular clusters, 10, 11, 12, 15, 19, 26, 27, 31, 32, 34, 35, 36) or oil and gas (cluster 28). There are a number of emerging or strengthened services centers for health (e.g., clusters 1-4), insurance (46, 47), business (48), transport (49, 50), education (51), hotels and recreation (7) and eating and hotels for military and government (5, 8, 9). While large metro areas are diverse, some have a relative focus on industry (42, 43, 45, 53) while others have a service focus (40, 41, 44, 52).

Average city sizes tend to vary by cluster. Within most clusters, cities have very similar sizes with one or two outliers. Across clusters, if we use cluster 45 with 20 market centers in diverse manufacturing having an average size of 1.3m as the base, all clusters with average sizes under 350,000 (except for one isolette, Reno) have significantly smaller average city sizes, and clusters 28 and 52 with average sizes of 965,000 and 556,000 also have significantly smaller average city sizes.[17] Correspondingly all clusters with average sizes over 2.7m have significantly city sizes. Within broad industrial categories, larger city types appear to be more diversified.

In Appendix B and Table 4, we see that education also varies by city type. Using cluster 45 as the base again, it has a little higher than the national average percent of adults with 4^+ years of college in 1990, or 21.5%. Seventeen other clusters have significantly

[17]This is based on a regression of log (1990 SMSA) population on cluster dummy variables, looking for significant percentage variables.

lower percent college, especially traditional manufacturers — food (38), primary metals (39), furniture (41), and textiles (11), with only 11-12% college educated. Eight other clusters have significantly higher percentages, especially high tech centers (electronics (12), instruments (14), and computers (10)), as well as some of the bigger market centers, with up to 30% college educated. Overall, there is a modest positive correlation between the natural log of city population and educational attainment (simple correlation coefficients for 1980 and 1990 are .15 and .28), driven by the varying needs of different city types for skilled workers.

Education across cities seems to change slowly over time. A five cell transition matrix of percent college educated for 1980 to 1990 has diagonal elements of .91, .73, .84, .67, and .81 going from low to high, with zero off diagonal elements for cells not immediately adjacent to the diagonal. This matrix is for cut-off points (relative to the mean) of .65, .8, 1.1, 1.35, and open with a 1990 distribution of .12, .18, .39, .17 and .13. At least for 80-90 transitions, the ergodic distribution exhibits no tendency to collapse or go bimodal. This is consistent with the model in section 1.

Urban Transformation. In the last 20 years, with the decline of manufacturing and rise of services nationally, many cities have undergone transformation of their industrial bases. While such cities may experience little change in size, there does appear to be some relationship between size and type. We specifically examine for a decade period, 1980-1990, the correlation between changes in city sizes and changes in industrial composition.

We measure changes in industrial composition by the changes in employment shares, so for each city, j, the index of composition change is

$$CC_j = \sum_{i=1}^{80} (s_{ij}^{1990} - s_{ij}^{1980})^2.$$

This index is based on the magnitude of absolute changes in shares; and, with squaring, reflects whether there are significant changes in the major industries in a city. We can also control for national composition changes. In that case we have

$$\overline{CC}_j = \sum_{i=1}^{80} ((s_{ij}^{1990} - \overline{s}_i^{1990}) - (s_{ij}^{1980} - \overline{s}_i^{1980}))^2.$$

To see correlations, we regress CC_j and $\overline{CC}_j$ on 1980-90 city population change, controlling for initial city population and initial manufacturing base. Results are reported in Table 5. First we note that initial population size detracts from composition change, because bigger cities are much more diverse to begin with. Second, initial manufacturing share encourages composition change, since manufacturing cities are subject to the national trend decline in U.S. manufacturing. In terms of the key relationship between composition change and population growth, the relationship is U-shaped with (1) slow growing cities having large changes in mix, (2) the change in mix declining as growth increases achieving a minimum at a growth rate of .29 (.27 for $\overline{CC}$) and (3) very fast growing cities again experiencing big changes in mix. This is almost exactly what we expect: the slowest growing cities which are moving down the size distribution and the fastest growing cities moving up experience the biggest changes in industrial composition. But there are complications.

In Table 5, column 2 identifies cities which transit up or down the size distribution (based on Table 3 categories for 317 metro areas) between 1980 and 1990. The dummy variable outcomes and plots of both CC and conditional residuals (residuals of a CC regression on ln (pop. 80) and manu share 80) against population growth rates suggest the relationship on the downside (big changes in mix with slow growth) is stronger than on the upside. Moreover, cities experiencing the least composition change (at the minimum point of .29 population growth rate) are growing at least a standard deviation above the mean growth rate. While our hypothesis is that cities moving up or down the size distribution are changing type, there is also the notion that cities with historical industrial bases that, by accident, reflect nationally high growth industries (national trends) will enjoy high growth rates per se. Average growth cities are those with a healthy but not spectacular mix.

CONCLUSIONS

Twentieth century urban evolution in the USA is characterized by parallel growth of cities of different sizes and types, maintaining with entry of new cities a stable relative size distribution of cities over time. Cities also appear to evolve with differing per person human capital levels. This paper models these features in a context in which cities experience local information (scale economy) and knowledge (human capital) spillovers. Cities grow in size with human capital accumulation and in number if the exogenous national rate of population growth is high enough. There is observed real income inequality across cities, but not net of human capital costs. The human capital equilibrium levels or rates of accumulation are less than the efficient levels or rates, given individuals don't internalize spillover externalities.

Table 1

Growth in City Sizes and Numbers

	Growth in Average Urban Population sample: all 1950 metro areas)	Growth in Number of Urban Areas (sample: Relative 0.12 of mean)	Growth in Number of Metro Areas (sample in each decade: 50,000 urban population)	Number of Metro Areas Where % Intra-Decade Urban Population Loss Exceeds 5%
1900-10	39%	15%	31%	1
1910-20	29	9	24	1
1920-30	29	4	22	2
1930-40	6	1	1	4
1940-50	18	7	12	2
1900 level	143,000	115	72	
1950 level	417,000	162	162	

Table 2. Transition Matrix

Cell in t

Cell in		1	2	3	4	5	6	Decade Rate* of Entry
$t+1$	1	.873	.058	0	0	0	0	.0649
	2	.123	.797	.082	0	0	0	.0029
	3	.032	.101	.656	.083	0	0	0
	4	0	.043	.213	.750	.111	0	0
	5	0	0	.049	.167	.825	.031	0
	6	0	0	0	0	.063	.969	0

*The average decade rate of entry is $i * Z$. i is total entrants (47) over the sum of cities in base years 1900-1940 (693). For Z, 95.745% enter in cell 1 and the rest in cell 2.

Table 3

Predicted Distributions

Cell	(i) Cell cut-off point	(ii) 1900 actual distribution	(iii) Steady state	(iv) 1950 trans. from 1900	(v) 1950 actual	(vi) Steady State: absolute cut-off point definition	(vii) Steady State: transitions with no entry
f_1	.3	.452	.405	.422	.401	.503	.245
f_2	.5	.209	.217	.225	.253	.240	.216
f_3	.75	.096	.082	.079	.087	.087	.099
f_4	1	.044	.118	.101	.093	.076	.157
f_5	2	.104	.110	.092	.105	.055	.199
f_6	open	.096	.068	.081	.080	.040	.084

Table 4 - Summary of Clustering for City Types

Cluster group	No. of clusters	No. of cities	Range: Avg. cluster spec-lalization, (within group)	Range: Avg.* cluster city size, (1000's) % college +
1. Clothing and Food	8	48	.05 - .19	121 - 915 12.4 – 19.3
2. Wood Products (pulp, furniture, lumber)	4	14	.05 - .18	110 - 208 14.1 - 17.7
3. Electronics	5	20	.08 - .25	114 - 712 12.6 - 29.7
4. Heavy Manu. (primary metals, trans. equip., machinery)	8	48	.05 - .28	175 - 965 12.4 - 24.5
5. Oil and Chemicals	5	19	.05 - .18	145 - 246 15.5 - 18.3
6. Market Center	6	52	n.a.	1285 - 2673 21.5 - 27.2
7. Health and Hotel Services	9	72	.11 - .28	148 - 530 15.2 - 29.8
8. Other Services (insurance, business, education, trans.)	10	44	.06 - .22	154 - 1169 13.4 - 27.9

* Range excludes clusters which are isolette's

Table 5

Composition Change and Population Growth

	CC	CC	$\overline{CC}$
constant	.0248*	.0226*	.0249*
	(.0057)	(.0058)	(.0052)
ln (pop. 80)	-.0016*	-.0016*	-.0018*
	(.0005)	(.0004)	(.0004)
manu. share 80	.0126*	.0128*	.0112*
	(.0042)	(.0042)	(.0038)
pop. growth rate 80-90	-.0141*		-.0101**
	(.0063)		(.0057)
pop. growth rate squared	.0239**		.0204**
	(.0125)		(.0112)
dummy: city transitting down		.0229*	
		(.0087)	
dummy: city transitting up		.0050	
		(.0055)	
N	317	317	317
R^2	.11	.12	.12

Mean and s.d.: ln (pop. 80): (12.6, .997); manu share 80 (.276. .121); pop. growth (.113, .147); CC (.0070, .0083); $\overline{CC}$ (.0052, .0075).

* Significant at 5% level. Standard errors in parentheses.

** Significant at 10% level.

Appendix A

In this Appendix, we give more technical details on the steady-state solutions. We start with steady-state levels.

Growth Paths

Steady-State Levels. To derive the phase diagram in Figure A1, we know from equations (23) and (24)

$$\dot{h} = 0 \text{ when } c = Bh^{\epsilon} - gh \quad \text{and} \quad h = 0$$

$$\dot{c} = 0 \text{ when } h = \left(\tfrac{A}{\rho}\right)^{\frac{1}{1-\epsilon}} \qquad \text{and} \quad c = 0.$$

In the phase diagram the $\dot{c} = 0$ locus is a vertical line as illustrated. For the $\dot{h} = 0$ locus, we note

$$\begin{aligned} \tfrac{dc}{dh}\Big|_{\dot{h}=0} &= [\epsilon B]h^{(\epsilon-1)} - g \\ \tfrac{d^2c}{dh^2}\Big|_{\dot{h}=0} &= \epsilon(\epsilon-1)Bh^{\epsilon-2} < 0. \end{aligned}$$

$\dot{h} = 0$ has a maximum at

$$\overset{*}{h} = \left(\frac{B\epsilon}{g}\right)^{\frac{1}{1-\epsilon}} \tag{A1}$$

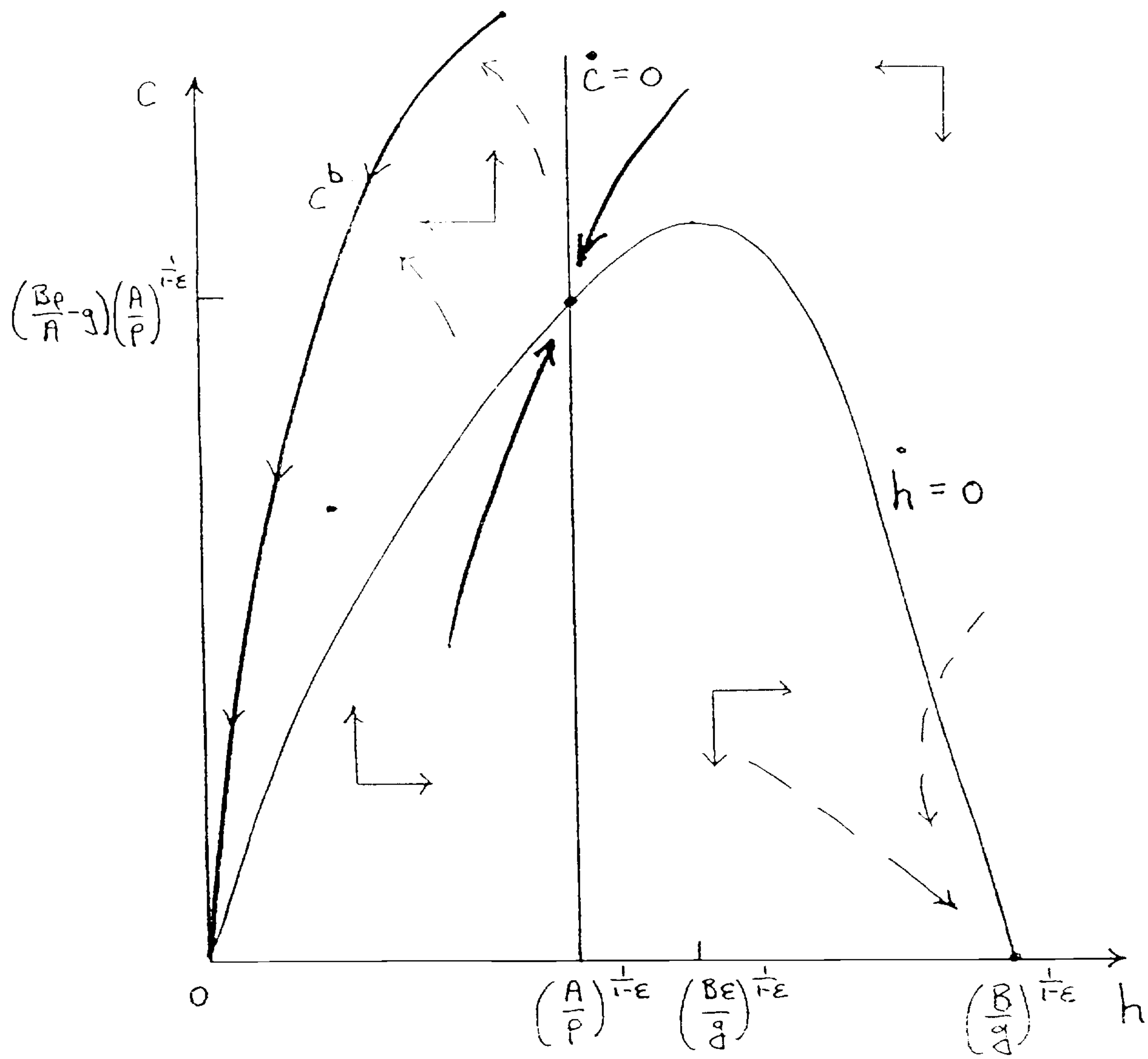
and $\dot{h} = 0$ intersects the horizontal axis at 0 and

$$h(c = 0,\ \dot{h} = 0) = \left(\frac{B}{g}\right)^{\frac{1}{1-\epsilon}}. \tag{A2}$$

Note $\overset{*}{h}$ in (A1) can be shown to be greater than the steady-state $\bar{h}$ in (27).

The motion in the system is shown by arrows, where to the right of $\dot{c} = 0$ (for high h) from (23) c is decreasing, while to the left c is increasing. For $\dot{h} = 0$, above $\dot{h} = 0$ (high c) from (24) h is decreasing while below h is increasing. The stable arm leading to $\bar{h}, \bar{c}$ is graphed. The point $\bar{h}, \bar{c}$ exhibits local saddle path stability given the Jacobian

$$\begin{vmatrix} \frac{\partial(\dot{h}/h)}{\partial h} & \frac{\partial(\dot{h}/h)}{\partial c} \\ \frac{\partial(\dot{c}/c)}{\partial h} & \frac{\partial(\dot{c}/c)}{\partial c} \end{vmatrix} < 0,$$



C
C^b
$\dot{C} = 0$
$\left(\frac{B\rho}{A} - g\right)\left(\frac{A}{\rho}\right)^{\frac{1}{1-\varepsilon}}$
$\dot{h} = 0$
O
$\left(\frac{A}{\rho}\right)^{\frac{1}{1-\varepsilon}}$
$\left(\frac{B\varepsilon}{g}\right)^{\frac{1}{1-\varepsilon}}$
$\left(\frac{B}{g}\right)^{\frac{1}{1-\varepsilon}}$
h

Figure A1

at $\overline{h}, \overline{c}$. Transversality is satisfied at $\overline{h}, \overline{c}$, given

$$\lim_{t\to\infty}[\lambda_1(t)H(t)] = \lim_{t\to\infty}[c(t)^{-\sigma}e^{-\rho t}h(t)e^{gt}] = \lim_{t\to\infty}[\overline{c}^{-\sigma}\overline{h}e^{-(\rho-g)t}],$$

for $\rho > g$, as assumed.

To rule out paths other than the stable arm, we note to the right of the $\dot{c} = 0$ locus, arms converge to $c = 0$, $h = h(c = 0, \dot{h} = 0)$ in equation (A.2). That point violates transversality, where $\lim_{t\to\infty}[\lambda_1(t)H(t)] = 0$ requires $\lim_{t\to\infty}\left[\frac{\dot{\lambda}_1}{\lambda_1} + \frac{\dot{h}}{h} + g\right] = 0$. Evaluating this limit, given (15a) and (24), we get

$$\lim_{t\to\infty}[(B-A)h^{\epsilon-1} - c/h].$$

At $c = 0$ and h in (A2), this equals $(B-A)\frac{g}{B}$. Given $B > A$ in (23) and (24), transverality is violated at $c = 0$, $h(c = 0, \dot{h} = 0)$.

What about paths above the stable arm? In all cases, they hit the constraint in (1c) that $\dot{H} \geq 0$, which defines a boundary value of c, c^b, where

$$c^b = Bh^\epsilon \tag{A3}$$

which lies above the $\dot{h} = 0$ locus (where $c = Bh^\epsilon - gh$). Once a path above the stable arm hits c^b in Figure A1, the path travels along c^b to origin. We note $\partial c^b/\partial h > 0$, $\partial^2 c^b/\partial h^2 < 0$, and $c^b > c(\dot{h} = 0)\forall h$. To show paths going to the origin violate transversality we must examine the problem in a different space, multiplier, h space, showing that the redefined transversality condition is violated as we approach the origin. The proof is in Black and Henderson (1997).

Steady-State Growth. For steady-state growth, there are no transition dynamics: the economy is always at the steady-state growth rate. What is that growth rate? For consumption per person, consumer optimization implies $c(t) = c_0 e^{(\frac{A-\rho}{\sigma})t}$. However in the text, the solution to equation (24) after substituting in the solution for $c(t)$ from (23) has the form

$$h(t) = Fe^{(B-g)t} + c_0(B - g - \frac{(A-\rho)}{\sigma})^{-1}e^{\frac{(A-\rho)}{\sigma}t} \tag{A4}$$

In order for $h(t)$ to grow at the rate $(\frac{A-\rho}{\sigma})$ in equation (25) F must equal zero. The transversality condition will require an $h(t)$ path where $F = 0$. Transversality requires

$$\lim_{t\to\infty}[\lambda_1(t)H(t)] = \lim_{t\to\infty}[c(t)^{-\sigma}e^{-\rho t}e^{gt}[Fe^{(B-g)t} + c_0(B-g-\frac{A-\rho}{\sigma})^{-1}e^{\frac{(A-\rho)}{\sigma}t}]$$

after substituting $H(t) = h(t)e^{gt}$ and for $\lambda_1(t)$ from (15a). Substituting for $c(t)$, yields

$$\lim_{t\to\infty}[c_0^{-\sigma}Fe^{(B-A)t} + c_0^{1-\sigma}(B-g-(\frac{A-\rho}{\sigma}))^{-1}e^{(\frac{1-\sigma}{\sigma}A+g-\frac{\rho}{\sigma})t}].$$

Given $B-A>0$ in (23) and (24), satisfaction of the transversality condition requires a path where $F=0$ and imposes a parameter constraint given in (25b) that

$$\left(\frac{1-\sigma}{\sigma}A+g-\frac{\rho}{\sigma}\right)<0.$$

Note this inequality, along with $B-A>0$, ensures in (A4) that $B-g-\frac{A-\rho}{\sigma} = (B-A) - (g-\frac{\rho}{\sigma}+\frac{1-\sigma}{\sigma}A) > 0$.

Efficiency

For the planner's problem, given $m_1n_1 = ze^{gt}$, total $\overline{X}_1$ for use in production of X_2 is $(zD_1n_1^{\delta_1}h_1^{\psi_1+\theta_1} - zbn_1^{1/2} - (1-z)bn_2^{1/2})\, e^{gt}$ where the first term is total X_1 output (equation 3) and the next two are national commuting costs. Then total output of X_2 equals $m_2n_2D_2n_2^{\delta_2}h_2^{\psi_2+\theta_2}\left(\frac{X_1}{m_2n_2}\right)^{1-\alpha}$ where $m_2n_2 = (1-z)e^{gt}$. The planner's problem is then

$$\max_{n_2,n_1,h_2,h_1} \mathcal{L} = \frac{c^{1-\sigma}-1}{1-\sigma}e^{-(\rho-g)t}$$

$$+ \quad \lambda_1e^{gt}[(1-z)^{\alpha}D_2n_2^{\delta_2}h_2^{\psi_2+\theta_2}(zD_1n_1^{\delta_1}h_1^{\psi_1+\theta_1} - zbn_1^{1/2} - (1-z)bn_2^{1/2})^{1-\alpha} - c]$$

$$+ \quad \lambda_2[H - ze^{gt}h_1 - (1-z)e^{gt}h_2].$$

The solution to this problem is stated below. An alternative way to proceed is to have a quasi-social planner, who allows national output markets to operate but controls h_1, h_2, n_1, n_2 and z, recognizing the effects on P. In this case, to solve for an efficient outcome, into the representative family's dynamic optimization problem in (1), we substitute for I_1 and I_2 from (9) and (13). We then substitute in $P = \{[Q_1(\alpha - 2\delta_2)h_1^{\epsilon_1} z]/[Q_2(1-\alpha+2\delta_2)h_2^{\epsilon_2}(1-z)]\}^{\alpha-2\delta_2}$, obtained by combining (17) with (9), (13), $e^{gt}z = m_1 n_1$, and $e^{gt}(1-z) = m_2 n_2$. Optimizing with respect to h_1, h_2, c, and z yields with rearrangement

$$\begin{aligned}
\overset{*}{h}_1 \,/\, \overset{*}{h}_2 &= \left[\frac{\epsilon_1(1-\epsilon_2)}{\epsilon_2(1-\epsilon_1)}\right] \\
\overset{*}{h}_i &= \frac{\epsilon_i}{1-\epsilon_i}\,\overset{*}{K}\,h \\
\overset{*}{z} &= \frac{(1-\epsilon_1)(1-\alpha+2\delta_2)}{(1-\epsilon_1)(1-\alpha+2\delta_2)+(1-\epsilon_2)(\alpha-2\delta_2)} \\
\overset{*}{I}_1 \,/\, \overset{*}{I}_2 &= \frac{(1-\epsilon_2)}{(1-\epsilon_1)} \\
\overset{*}{K} &= \frac{(1-\epsilon_1)(1-\alpha+2\delta_2)+(1-\epsilon_2)(\alpha-2\delta_2)}{\epsilon_1(1-\alpha+2\delta_2)+\epsilon_2(\alpha-2\delta_2)} \\
\overset{*}{P} &= \overset{*}{Q}\,h^{(\epsilon_1-\epsilon_2)(\alpha-2\delta_2)} \\
\overset{*}{Q} &= \left[\frac{\epsilon_1 Q_1(\frac{\epsilon_1}{1-\epsilon_1})^{\epsilon_1-1}}{\epsilon_2 Q_2(\frac{\epsilon_2}{1-\epsilon_2})^{\epsilon_2-1}}\,\overset{*}{K}^{\epsilon_1-\epsilon_2}\right]^{\alpha-2\delta_2}.
\end{aligned}$$

These equations, along with (17)-(20) give us the $h_1/h_2, \overset{*}{h}_1 \,/\, \overset{*}{h}_2$, I_1/I_2, and $\overset{*}{I}_1 \,/\, \overset{*}{I}_2$ comparison in Proposition (4) in the text, by inspection. By substitution of the new values of $\overset{*}{h}_i$ into (8), (12) and (18) we can compare $\overset{*}{n}_1 \,/\, \overset{*}{n}_2$ with n_1/n_2 and corresponding $\overset{*}{m}_1 \,/\, \overset{*}{m}_2$ with m_1/m_2 to obtain detailed expressions.

For the rest of Proposition (4), we combine first-order conditions corresponding to (15a) and (15e) with the new equation of motion to solve for steady-state growth and levels. For steady-state growth, when $\epsilon = 1$, we have

$$\overset{*}{\gamma}{}^{c} = \overset{*}{\gamma}{}^{h} = \frac{\overset{*}{A} - \rho}{\sigma} \tag{A5}$$

$$\overset{*}{A} \equiv \frac{\epsilon_1 Q_1}{\overset{*}{Q}}\left(\frac{\epsilon_1}{1-\epsilon_1}\right)^{\epsilon_1 - 1} \overset{*}{K}^{\epsilon_1 - 1}.$$

For steady-state levels, when $\epsilon < 1$, we have

$$\overset{*}{\bar{h}} = \left(\frac{\overset{*}{A}}{\rho}\right)^{\frac{1}{1-\epsilon}} \tag{A6}$$

$$\overset{*}{\bar{c}} = \left(\frac{\overset{*}{A}}{\rho}\right)^{\frac{1}{1-\epsilon}} \left(\frac{\overset{*}{B}\,\rho}{\overset{*}{A}} - g\right) \tag{A7}$$

$$\overset{*}{B} \equiv \overset{*}{A}\overset{*}{z}\,(1-\alpha+2\delta_2)^{-1}(1-\epsilon_1)^{-1}\,\overset{*}{K}\,.$$

In comparing the equilibrium and the optimum, for steady-state growth where $\epsilon = 1$, we know $\overset{*}{\gamma}{}^{c} > \gamma^c \; iff \; \overset{*}{A} > A$. Similarly for steady-state levels where $\epsilon < 1, \overset{*}{\bar{h}} > \bar{h}$ iff $\overset{*}{A} > A$. With substitutions we get

$$\tilde{A} = \tilde{\phi}_1^{\epsilon_1(1-\alpha+2\delta_2)}(1-\tilde{\phi}_1)^{(1-\epsilon_1)(1-\alpha+2\delta_2)}\tilde{\phi}_2^{\epsilon_2(\alpha-2\delta_2)}(1-\tilde{\phi}_2)^{(1-\epsilon_2)(\alpha-2\delta_2)}$$

$$Q_1^{1-\alpha+2\delta_2}Q_2^{\alpha-2\delta_2}\tilde{K}^{\epsilon-1} \tag{A8}$$

where

$$\tilde{K} = \left[\frac{(1-\tilde{\phi}_1)(1-\alpha+2\delta_2)+(1-\tilde{\phi}_2)(\alpha+2\delta_2)}{\tilde{\phi}_1(1-\alpha+\delta_2)+\tilde{\phi}_2(\alpha-2\delta_2)}\right]^{\epsilon-1}.$$

In (A9), for $\overset{*}{A}, \tilde{\phi}_i = \epsilon_i$ and for $A, \tilde{\phi}_i = \phi_i$. $\overset{*}{A} > A$, if $\epsilon_1 \geq \phi_1$ and $\epsilon_2 \geq \phi_2$, with strict inequality for one. This can be seen by noting that $\partial\tilde{A}/\partial\tilde{\phi}_i > 0$ iff $\epsilon_i > \phi_i$, (where $\frac{\partial\tilde{A}}{\partial\tilde{K}}\frac{\partial\tilde{K}}{\partial\tilde{\phi}_i} \geq 0$ by inspection).

Finally for $\bar{c}$ *vs.* $\overset{*}{\bar{c}}$ when $\epsilon < 1$, repeated substitution gives

$$\tilde{\bar{c}} = [\tilde{A}Q_2^{\alpha-2\delta_2}Q_1^{1-\alpha+2\delta_2}]^{1-\epsilon}\rho^{\frac{-\epsilon}{1-\epsilon}}\left[\frac{1}{(1-\tilde{\phi}_1)(1-\alpha+2\delta_2)+(1-\tilde{\phi}_2)(\alpha-2\delta_2)}\right].$$

$$[1-\frac{g}{\rho}(\tilde{\phi}_1(1-\alpha+2\delta_2)+\tilde{\phi}_2(\alpha-2\delta_2))]$$

where, for $\overset{*}{\bar{c}}$, $\tilde{\phi}_i = \epsilon_i$ and for $\bar{c}, \tilde{\phi}_i = \phi_i$. Differentiation tells us $\frac{\partial\tilde{\bar{c}}}{\partial\tilde{\phi}_i} > 0$ if $\rho > g$ and either or both $\epsilon_1 > \phi$, or $\epsilon_2 > \phi_2$. Thus, if $\rho > g$ and either or both $\epsilon_1 > \phi_1$ or $\epsilon_2 > \phi_2$, $\overset{*}{\bar{c}} > \bar{c}$.

Implementation of an Efficient Solution.

If developers can control the per person level of human capital in their cities, under the correct incentives, they will achieve an efficient outcome. Assume, for example, h_1 and h_2 are chosen by developers, who borrow in a national human capital market from capital owned by families, at a prevailing rental rate of r. In city type 1, the developer's problem is now

$$\max \frac{1}{2} bn_1^{3/2} - T_1 n_1 - rh_1 n_1 + \lambda(D_1 n_1^{\delta_1} h_1^{\psi_1+\theta_1} + T_1 - \frac{3}{2} bn_1^{1/2} - I_1).$$

In this problem $T_1 = D_1 \delta_1 n_1^{\delta_1} h_1^{\psi_1+\theta_1} - rh_1 = \frac{1}{2} bn_1^{1/2} - rh_1$, so while the Henry George Theorem still holds, T_1 is reduced by implied capital rental payments. Solving the problem for city types 1 and 2, we get the text expressions for n_1 and n_2 and

$$r = \epsilon_1 Q_1 h_1^{\epsilon_1 - 1} = \epsilon_2 Q_2 P^{\frac{1}{\alpha - 2\delta_2}} h_2^{\epsilon_2 - 1} \tag{A9}$$

$$I_1 = Q_1 h_1^{\epsilon_1}(1 - \epsilon_1) = I_2 = Q_2 P^{\frac{1}{\alpha - 2\delta_2}} h_2^{\epsilon_2}(1 - \epsilon_2). \tag{A10}$$

Note since capital is here a freely mobile input, chosen by developers, labor incomes in national markets must be equalized for families in allocating members. Now, in the family's optimization problem, $\dot{H} = I_1 P^{-1} z e^{gt} + I_2 P^{-1}(1-z) e^{gt} + P^{-1} rh e^{gt} - c e^{gt}$, where capital rent is paid in units of X_1 but new additions to capital come from X_2. In choosing z, in the family's problem to maximize $\frac{c^{1-\sigma}-1}{1-\sigma} e^{-(\rho-g)t}$ subject to (just) the equation of motion, we get $I_1 = I_2$. (Note now I's are adjusted for T's reflecting capital rental payments.)

Besides (A9) and (A10) we substitute into $m_1 X_1 = m_2 n_2 x_1 + m_1(bn_1^{3/2}) + m_2(bn_2^{3/2})$ for $ze^{gt} = m_1 n_1, (1-z)e^{gt} = m_2 n_2$, for X_1 and x_1 as before, and for P from (A10) to get $zh_1^{\epsilon_1} Q_1 = (1-z) Q_2 P^{\frac{1}{\alpha-2\delta_2}} h_2^{\epsilon_2} (\frac{1-\alpha+2\delta_2}{\alpha-2\delta_2})$. Combining this, (A9), (A10) and $h = zh_1 + (1-z)h_2$ gives the efficient outcomes listed above for $\overset{*}{h}_1$ / $\overset{*}{h}_2$ and $\overset{*}{z}$. Solving the family's optimization problem gives (A5)-(A7).

APPENDIX B -- CITY TYPES

Cluster Number	Cluster Name	No. of MSA's	Average Size (1000's); % College +	Share of Dominant Industry*		Comments
A) Health and Hotel Services						
1	Health services (80)	3	152 29.8	.278	(.115)	(Iowa City, IA, Rochester, MN)
2	Health services (80)	8	148 17.3	.190	(.115)	Former heavy industry (Alexandria, VA, Wichita Falls, TX)
3	Health services services (80) and mixed industry	20	311 21.4	.136	(.115)	Transforming industrial centers (Jackson, MS, Springfield, MO)
4	Health services (80) with electronics	5	160 28.2	.172	(.115)	State govt. &, former heavy machinery (Charlottesville, VA, Gainesville, FL)
5	Eating places, Health services	16	458 20.7	n.a.	n.a.	Govt., military (Bremerton, WA, Fort Walton, FL)
6	Hotels	2	530 15.2	.330	(.016)	(Las Vegas, NV, Cape May, NJ)
	Hotels (70) and recreation (79)	1	255 20.7	.192 .042	(.016) (.012)	(Reno, NV)
[illegible]	Eating places (58), some industry	4	159 21.9	.113	(.079)	Military (Fayetteville, NC, Lawton, OK)
[illegible]	Eating places (58) and hotels	13	293 21.9	.121	(.079)	State govt., military (Jacksonville, NC, Bryan, TX)

APPENDIX B -- CITY TYPES (Continued)

Cluster Number	Cluster Name	No. of MSA's	Average Size (1000's); % College +	Share of Dominant Industry*		Comments
B) Electronics						
10	Computers (87 + 35 + 36) (engineering, R&D, electronics)	3	712 27.7	.220	(.06)	(San José, CA, Huntsville, AL)
11	Electrical machinery (36)	2	114 12.6	.245	(.016)	(Madison County, IN Kokomo, IN)
12	Electronics (36)	4	159 29.7	.101	(.016)	(Binghamton, NY, Bloomington, IN)
13	Diverse machinery (electronics)	6	232 26.4	n.a.	n.a.	(Boulder, CO, Cedar Rapids, IA)
14	Instruments	5	291 18.2	.077	(.008)	(Rochester, NY, Sherman, TX)
C) Clothing and Food Processing						
15	Textiles (22)	3	121 12.5	.188	(.006)	(Danville, VA, Washington County, RI)
16	Textiles (22)	3	915 19.0	.069	(.006)	(Charlotte, NC, Greenville, SC)
17	Textiles & food processing (22 +20)	4	236 19.3	.105	(.025)	(Athens, GA, Columbus, GA)
18	Apparel (23) and mixed industry	10	370 14.9	.054	(.007)	Declining industrial centers (Brownsville, TX, Anniston, AL)
19	Food processing (20)	6	165 15.5	.106	(.019)	(Merced, CA, Sioux City, IA)

APPENDIX B -- CITY TYPES (Continued)

Cluster Number	Cluster Name	No. of MSA's	Average Size (1000's); % College +	Share of Dominant Industry*		Comments
20	Food processing, machinery, transport	2	156 12.4	n.a.	n.a.	(Fort Smith, AR, Joplin, MO)
21	Food processing (20), wholesale, agriculture, transport	11	293 17.6	.046	(.019)	(Visalia, CA, Fresno, CA)
22	Food processing (20), and diversified heavy machinery	9	248 14.5	.050	(.019)	Declining heavy industry (Gadsden, AL, Sheboygan, WI)
D) Heavy Manufacturing						
23	Machinery (35)	11	175 16.7	.080	(.020)	Declining machinery, steel centers (Peoria, IL, Waterloo, IA)
24	Transforming steel cities (33)	9	260 14.0	.049	(.010)	Former steel giants (Huntington, WV) Reading, PA)
25	Diverse heavy machinery	6	212 13.6	n.a.	n.a.	Transport, machinery, metals (Jackson, MI, Rockford, IL)
26	Primary metals	3	290 12.4	.165	(.010)	(Gary, IN, Steubenville, OH)
27	Transport equipment (37)	4	363 24.5	.149	(.016)	(Wichita, KS. Flint, MI)
28	Transport equipment (37)	13	965 19.3	.067	(.016)	Declining transport (Kenosha, WI, Saginaw, MI)

APPENDIX B -- CITY TYPES (Continued)

Cluster Number	Cluster Name	No. of MSA's	Average Size (1000's); % College +	Share of Dominant Industry*		Comments
29	Transport equipment (37) and lumber (24)	1	156 14.2	.155 .074	(.016) (.006)	(Elkhart, IN)
30	Transport equipment (37)	1	115 14.4	.284	(.016)	(Jackson, MI)
E) Oil and Chemicals						
31	Non-metallic minerals (32)	1	138 10.8	.141	(.006)	(Vineland, NJ)
32	Oil and gas (13), mining & water transport	2	145 18.0	.106	(.003)	(Houma, LA. Midland, TX)
33	Oil and gas (13), transport	5	231 18.3	.050	(.012)	(Lafayette, LA, Anchorage, AK)
34	Chemicals (28), petro. chemicals	9	246 15.5	.073	(.011)	(Johnson City, TN, Lake Charles, LA)
35	Chemicals (28)	2	171 18.1	.179	(.011)	(Brazooria, TX Richland, WA)
F) Wood Products						
36	Furniture (25)	1	222 16.1	.178	(.004)	(Hickory, NC)
37	Pulp and paper (26)	3	208 16.1	.076	(.008)	Diversifying (Green Bay, WI, Oshkosh, WI)
38	Pulp and paper (26), leather	4	110 14.1	.057	(.008)	Declining non-heavy industry (Pine Bluff, AR Bangor, ME)

APPENDIX B -- CITY TYPES (Continued)

Cluster Number	Cluster Name	No. of MSA's	Average Size (1000's); % College +	Share of Dominant Industry*		Comments
39	Lumber (24) and non-durable wholesale	6	185 17.7	.054	(.006)	N.W. Declining lumber cities (Salem, OR, Eugene, OR)
G) Market Centers						
40	Diverse services (health, educ., eng. & mgmt.)	7	2,471 25.8	n.a.	n.a.	North-East post industrial metro areas (Boston, MA Philadelphia, PA)
41	Financial services (60 + 62)	1	8,547 24.6	.086	(.010)	(New York, NY)
42	Mixed base metro areas (high tech., wholesale, transportation and business services)	9	2,673 26.0	n.a.	n.a.	Western (and S.W.) metro areas (Orange County, CA, Denver, CO)
43	New mixed base (high tech. eating places, eng. & mgmt.)	8	1,405 27.5	n.a.	n.a.	Western (and S.W.) cities (Phoenix, AZ, San Diego, CA)
44	Business services (73), transport. eating	7	1,403 24.0	.093	(.051)	Southern cities (Houston, TX, Tampa, FL)
45	Diverse manufacturing	20	1,285 21.5	n.a.	n.a.	Mid-west MSA's, (Cleveland, OH, Chicago, IL)
H) Service Centers						
46	Insurance (63)	2	627 27.7	.142	(.017)	(Bloomington, IL, Hartford, CT)
47	Insurance (63) and business services	5	369 23.3	.063	(.017)	(Des Moines, IA, Lincoln, NE)

APPENDIX B -- CITY TYPES (Continued)

Cluster Number	Cluster Name	No. of MSA's	Average Size (1000's); % College +	Share of Dominant Industry*		Comments
48	Business services (73)	2	154 13.7	.216	(.051)	(Nanatee County, FL, Kankakee, IL)
49	Air transport (45) and oil	1	57 17.3	.084	(.005)	(Enid, OK)
50	Trucking and transport services (42 + 47)	1	133 11.1	.163	(.020)	(Laredo, TX)
51	Private education services (82), manufacturing	2	295 27.9	.113	(.018)	(Provo, UT, Trenton, NJ)
52	Diverse services (business, insurance, education, health)	8	556 20.1	n.a.	n.a.	NE cities transforming from heavy industry (Springfield, MA, Syracuse, NY)
53	New Jersey transport, wholesale, chemicals	4	1,169 26.2	n.a.	n.a.	(Bergen, Somerset)
54	Diverse services and light manufacturing	15	796 19.2	n.a.	n.a.	Govt., and military (Miami, FL, Sacramento, CA)
55	Manufacturing NEC (3999), chemicals, diverse manufacturing	4	373 24.4	.075	(.010)	(Racine, WI, Wilmington, DE)

* A number in parentheses in column 2 represents an SIC code for which city-industry share, s_{ij}, is reported in column 5. In column (5), the number in parenthesis is the national share $\bar{s}_i$ of the industry.

References

Abdel-Rahman, H. and M. Fujita (1990), "Product Varieties. Marshallian Externalities and City Sizes," *Journal of Regional Science,* 2, 165-183.

Amemiya, T. (1985), *Advanced Econometrics,* Harvard University Press.

Becker, R. and J. V. Henderson (1996), "Intra-Industry Specialization and Urban Development," *The Economics of Cities,* J. Huriot and J. Thisse (eds.), forthcoming.

Benabou, R. (1993), "Workings of a City: Location, Education and Production," *Quarterly Journal of Economics,* 108, 619-652.

Benabou, R. (1996), "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review,* 86, 584-609.

Bergsman, J., P. Greenston and R. Healy (1972), "The Agglomeration Process in Urban Growth," *Urban Studies,*

Black, D. and J. V. Henderson (1997), "Urban Growth," Brown University Working Paper No. 97-1.

Bogue, D. (1953), *Population Growth in Standard Metropolitan Areas 1900-1950,* Scripps Foundation in Research in Population Problems.

Deo, S. and G. Duranton (1995), "Local Public Goods in a Production Economy," mimeo.

Dixit, A. and J. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity," *American Economic Review,* 67, 297-308.

Dobkins, L. and Y. Ioannides (1995), "Evolution of the U.S. Size Distribution of Cities," mimeo.

Eaton, J. and Z. Eckstein (1994), "Cities and Growth: Theory and Evidence from France and Japan," NBER Working Paper No. 4612.

Flatters, F., V. Henderson and P. Mieszkowski (1974), "Public Goods, Efficiency, and Regional Fiscal Equalization," *Journal of Public Economics,* 3, 99-112.

Fujita, M. and M. Ogawa (1982), "Multiple Equilibria and Structural Transition of Non-Monocentric Configurations," *Regional Science and Urban Economics,* 12, 161-96.

Fujita, M., P. Krugman, T. Mora (1995), "On the Evolution of Hierarchical Urban Systems," mimeo.

Garreau, J. (1991), *Edge City,* New York: Doubleday.

Helsley, R. and W. Strange (1990), "Matching and Agglomeration Economies in a System of Cities," *Journal of Urban Economics,* 189-212.

Helsley, R. and W. Strange (1993), "City Developers and Efficiency," mimeo.

Henderson, J.V. (1974), "The Size and Types of Cities," *American Economic Review,* 64, 640-656.

Henderson, J.V. (1988), *Urban Development: Theory Fact and Illusion,* Oxford University Press.

Henderson, J.V. and R. Becker (1996), "City Formation," Brown University, mimeo.

Kim, H.S. (1988), "Optimal and Equilibrium Land Use Pattern in a City," Brown University Ph.D. disseration.

Krugman, P. (1993), "On the Location and Number of Cities," *European Economic Review,* 37, 243-308.

Lucas, R.E., Jr. (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics,* 22, 3-42.

Quah, (1993), "Empirical Cross Section Dynamics and Economic Growth," *European Economic Review,* 37 (2/3), 426-434.

Rauch, J. (1993), "Does History Matter When It Only Matters a Little," *Quarterly Journal of Economics,* 108, 843-867.

Romer, P. (1986), "Increasing Returns and Long Run Growth," *Journal of Political Economy,* 94, 1002-37.

Romer, P. (1990), "Endogenous Technological Change," *Journal of Political Economy,* 98, 5, S71-S102.

Stiglitz, J.E. (1977), "The Theory of Local Public Goods," in *Economics of Public Services,* M.S. Feldstein and R.P. Inman (eds.), New York: MacMillan.

Upton, C. (1981), "An Equilibrium Model of City Sizes," *Journal of Urban Economics,* 10, 15-36.