Local Knowledge Spillovers and Inequality [*]

Duncan Black
London School of Economics [†]

February 1999

Abstract: Human capital is not evenly distributed across the metropolitan areas of the U.S. This paper models the endogenous determination of the distribution of human capital across metropolitan areas and the resulting implications for the issue of income inequality. If there is a local public benefit to private knowledge, then the degree to which individuals can capture externality effects will depend on their location. These externalities provide an incentive for sorting by ability type across different cities. Sorting augments productivity differences between workers of different types, leading to regional variation in both real and nominal wages. The wage gap between different worker types is increasing by the magnitude of local knowledge spillovers. A change in the degree of local knowledge spillovers affects the wage gap even if returns to individual ability remain constant. This wage gap is partially offset by cost of living differences between cities. Failure to properly account for this will lead to an overstatement of both the difference in real incomes between worker types. I present evidence that college graduates have become increasingly concentrated in recent decades, suggestive of a possible role of knowledge spillovers in rising income inequality.

Keywords: sorting, inequality, spillovers, human capital, cities

[*]A previous version of this paper circulated with a different title, "Local Human Capital Externalities, Educational Segregation, and Inequality." I would like to thank Roland Benabou, John Driscoll, Oded Galor, Herschel Grossman, Vernon Henderson, Moshe Justman, Harl Ryder, and David Weil for their help and suggestions. This paper has also benefited from participants of seminars at Brown, CU-Denver, Royal Holloway, Southampton, Sussex and USF as well as conference participants at the NBER Summer Institute and the CEPR/Arne Ryde Foundation "Conference on Trade and Location" at Lund University. All remaining errors are my own.
[†]Economics Department, London School of Economics, Houghton Street, London WC2A 2AE, D.Black@lse.ac.uk

1 Introduction

There is growing consensus that the social returns to education exceed the private ones. Human capital externalities arising from knowledge spillovers are thought to play a role in explaining global patterns of economic growth (Lucas 1988). Implicit in this is the notion that these spillovers do not benefit all locations equally. They are, to some degree, localized. In fact, direct evidence of the public benefit of human capital has been found by focusing on the localized, within country, variations in knowledge spillovers. Evidence has been found in such diverse settings as the localized nature of patent citations (Jaffe, Trajtenberg and Henderson 1993), learning spillovers in the adoption of new agricultural technologies in rural India (Foster and Rosenzweig 1995), and productivity benefits of local education levels in U.S. cities (Rauch 1993*b*). With the degree of these knowledge spillovers varying across geography, the productivity and income benefits that any individual can obtain will depend on their own location and resulting access to this information.

The premise of this paper rests on four pieces of empirical evidence. The first is that there are productivity-enhancing local knowledge spillovers, the magnitude of which are affected by the average human capital level of the population.[1] The second is that local human capital levels vary substantially across locations. The third is that these differences in local human capital levels influence the location choices of workers. Thus, the geographic distribution of human capital is endogenously determined. The fourth is that these spillovers appear to operate at the metropolitan area level.[2] Following from these, I present a model exploring how the geographic distribution of human capital across cities is determined and some implications of this for the analysis and understanding of income inequality.

Though emphasizing the role of spillovers in explaining cross-country growth and income differentials, Lucas (1988) does suggest that it is in the cities where evidence of productivity-enhancing knowledge spillovers may be found. Rauch (1993*b*) follows this suggestion and finds that average metropolitan area levels of educational attainment have a significant effect on worker productivity and wages after conditioning on individual characteristics.[3] In cities, the close proximity of firms facilitates interactions between workers in which knowledge is exchanged, with more highly educated workers transmitting more valuable information. Informal information exchanges that are reported to take place in the bars and restaurants of Silicon Valley are widely credited for some of that region's success.[4] Given that metropolitan areas are essentially labor market areas (Mills 1972), they are the appropriate geographic

[1] Throughout this paper, concepts of ability, knowledge, human capital, and educational attainment will be used interchangeably.

[2] Conceptually, metropolitan areas are defined as the economic, rather than political, city, though these terms will be use somewhat interchangeably throughout this paper.

[3] Specifically, Rauch constructs an estimate of the overall productivity effect and finds that a one-year increase in the average level of educational attainment in a Standard Metropolitan Statistical Area (SMSA) raises total factor productivity by 2.8%.

[4] See, for example, "A Survey of Silicon Valley," *The Economist*, 5/29/97.

area on which to focus.[5]

With worker mobility, one would expect the location decisions of workers to be influenced by these variations in local productivity levels.[6] Therefore, the geographic distribution of human capital would be endogenously determined by the location choices of heterogeneous workers. However, absent some additional sorting mechanism to allow higher ability workers to self-segregate, logically all workers cannot locate in the highest human capital areas. Evidence of a tendency to sort by ability type across cities is is provided by Borjas, Bronars and Trejo (1992) who demonstrate that the difference between one's own skill level and the average skill level of one's initial home region is an important factor in the propensity for inter-city migration.

It is illustrative to first examine the actual distribution of human capital across metropolitan areas. Figure 1 contains a histogram of the percentage of the population aged 25+ who are college educated relative to the national mean (unweighted by population) in 1990 for the 318 Primary Metropolitan Statistical Areas (PMSAs) in the 48 contiguous states.[7] In this sample, the mean percentage college educated across metropolitan areas is equal to 19.8 with a minimum of 9.5, a maximum of 44.0 and a standard deviation of 6.4. Clearly, there is some dispersion in the distribution, demonstrating that college educated individuals are not distributed evenly across cities.

The links between location and inequality arising from disparities in the provision of education across communities have been comprehensively explored in recent works by Benabou ((1993),(1996*b*), (1996*a*)) and Durlauf (1996). This literature has followed from basic insights in Tiebout (1956) and focused on the fact that with local financing of public education and local peer group effects in learning, sorting of the population will lead to unequal levels of educational output. Such sorting may be inefficient (de Bartolome 1990), can lead to dynastic poverty traps, and may lower long-run output levels, given economy-wide knowledge spillovers. Here, I depart from this focus on education and contribute this literature by examining the productivity effects of local knowledge spillovers. I abstract from the education issues here, not to diminish their importance, but simply to focus on a different mechanism. This paper is also related to to Kremer and Maskin (1996), who examine the sorting of heterogeneous workers across firms.

As the emphasis here is on city-wide spillovers, it is appropriate to use the theoretical tools already present in the literature on urban economics. I adapt a standard general equilibrium system of cities model (cf. Henderson (1974),(1988)). This general equilibrium approach system of cities approach used here is desirable because it allows for endogenous determination of city sizes and numbers and relies

[5]Considering further disaggregated geographic units, such as counties, would not be appropriate given substantial cross-county commuting.

[6]In the U.S., the work force appears to be highly mobile. Over the one year period 1992-1993, 6% of the population moved across county lines with 3% moving across state borders (from the U.S. Bureau of the Census, *Current Population Reports*, P20-485, as reported in the 1996 Statistical Abstract of the U.S.)

[7]The data is constructed from county-level decennial census information that is then aggregated up to the PMSA level.

on standard assumptions of agglomeration and congestion effects that allow for wage and cost of living differentials across cities of different sizes and population compositions. Given the extent of growth in city numbers and sizes over time as well as substantial consistent cross-sectional variation in city sizes (Black and Henderson 1998) this flexibility is important. In addition, theoretical results regarding the nature of equilibria in this type of model are often driven by the fixing of community sizes and numbers (Henderson 1985).

The first model presented below contains two types of individuals, differing in their ability levels, and a single occupation. Within this simple framework, it is possible to focus on the costs and incentives that lead to sorting by ability type across cities. Basic results regarding the existence, nature, and efficiency of equilibria are established.

The important results derived from this simple model are as follows. Individuals of different ability, who are imperfectly substitutable units of labor, will sort by type across cities. Economically identical individuals living in cities with different average levels of human capital will produce different levels of output. Thus, with segregation by type across different cities, higher ability workers will earn higher wages than less able workers due both to their own greater ability and to the productivity benefits of sorting. However, some of this wage gap will be offset by cost of living differences across cities. Cities with higher ability workers will be larger, a simple prediction that holds up empirically,[8] and have higher costs of living. For less able workers, the higher living costs will offset any wage gains they would obtain by relocating. Though their nominal wages would increase by relocating to cities with high ability residents, their real wages would decrease. This mechanism provides for equilibrium stratification by type.

Rising income inequality has been a notable trend in recent decades (see, for example, Juhn, Murphy and Pierce (1993)). This paper contributes to the understanding of this issue. One implication of this model is that failure to account for these cost of living differences across cities will overstate both the level of income inequality and the effects on this of any technological changes. In addition to this measurement issue, it also points to other potential causes for widening inequality. One standard explanation is that it is a consequence of an increase in the demand for skills resulting from some form of skill-biased technological change (Katz and Murphy 1992). The model presented here points to other, potentially complementary, causes. Given that the population is sorting by type in response to the incentives provided by local human capital externalities, small changes either in the underlying skill distribution or in the returns to skill will be magnified by spatial sorting. Small changes in these underlying parameters will lead to much larger effects on the income gap. Additionally, increases in the degree of local knowledge spillovers will increase the wage gap even if returns to individual ability levels

[8] In 1990, the correlation coefficient between the log of PMSA population and the fraction over 25 who are college educated is equal to .29. This holds even though some of the highest human capital cities, such as Iowa City, IA, are relatively small cities that are home to state universities, a government imposed distortion.

remain constant.

The basic model, sufficient to capture the important issues, does have one obviously unappealing prediction. In this framework, workers sort across different cities by ability type, and all equilibria are characterized by internally homogeneous cities. This result follows from the lack of any complementarity between worker types. In order to demonstrate that the qualitative results regarding sorting and its effects on income inequality hold up without this outcome, I augment the basic model by adding second occupation, producing a locally traded intermediate input good. This extended model is much less tractable than the basic one, but it can be shown that it preserves the qualitative results of the original model. This model exhibits *within occupation* sorting by ability type across cities. Real and nominal wage equality exists across ability types and within cities, and nominal wage inequality (only) exists within ability types. These results reaffirm the ideas present in the analysis of the basic model. Within occupations, knowledge spillovers lead to sorting by type across cities and augment wage inequality. Cost of living differences across cities exaggerate real inequality differences both between and within different worker types. These interesting results may shed some light on the puzzling increase in "residual inequality" in recent decades (Juhn et al. 1993).

To motivate this potential link between spillovers, sorting, and recent trends in the income distribution, later in the paper I present some evidence regarding the changing geographic concentration of human capital. I show that in recent decades there has been a trend towards increasing geographic concentration of human capital, mirroring increases in income inequality over the same period. Such increasing concentration suggests that the relative benefits that spillovers may provide to a select group of the population may be increasing. Ciccone, Peri and Almond (1999) show that the magnitude of city-wide externalities has increased in recent decades strongly. The combination of these two results - increased sorting coupled with higher benefits of sorting - suggests that variations in local spillovers may be playing a role in recent rising income inequality.

The remainder of the paper is organized as follows. In section 2, I outline the characteristics of the population, the technology of production, and the structure of cities in the basic model. Section 3 contains an analysis the equilibria of the basic model under two different mechanisms of city formation. In section 4 I discuss implications of the model for the issue of income inequality and present evidence demonstrating a trend of increasing geographic concentration of human capital. Section 5 presents the basic results for the augmented model with internally heterogeneous cities, and section 6 concludes.

2 The Basic Model

In this section the characteristics of the population, the technology of production, and the internal urban spatial structure are defined. Models in the literature related to this that have incorporated heterogeneous individuals into a system of cites framework have restricted attention to the cases in which different types of labor are complementary inputs of production (e.g. Henderson and Becker (1996),Abdel-Rahman (1996),(1997)), or the case where ex ante identical dynasties allocation individuals according to joint human capital investment and location decisions in an economy with different types of specialized cities (Black and Henderson 1999). Here it is assumed that the different types of workers are substitutable inputs possessing different efficiency units of labor. These are workers involved in similar productive tasks, with some workers simply being better than others.

2.1 Population

The population consists of N agents who are freely mobile and can migrate costlessly. There are two types of individuals in the economy. Proportion z of the population are high ability individuals and proportion $(1-z)$ are low ability. High ability individuals possess a_h efficiency units of labor and low ability individuals possess a_ℓ units, with $a_h > a_\ell > 0$.[9]

2.2 Production

There is a single good produced in the economy and firms consist of one worker.[10]. All workers provide their labor inelastically and receive what they produce. The wage for individual i working in city j is:

$$W_{ij} = D(\bar{a}_j^{\gamma} n_j^{\delta}) a_i \tag{1}$$

In (1), D is a technological constant, the same for all cities. In city j, $\bar{a}_j = (z_j a_h + (1-z_j)a_\ell)$ is the average level of human capital, n_j is the total city population, z_j the proportion of high ability workers, and a_i is the ability level of individual i. The expressions in parentheses represent the two sources of externalities in this model. First, δ is the elasticity of worker output with respect to city size, reflecting pure scale economies. Second, γ is the elasticity of worker output with respect to the average ability level in the city. Assume $0 < \delta < \frac{1}{2}$.[11]

I assume that scale economies in this model result from information spillovers (Fujita and Ogawa 1982). Trade secrets are passed between different worker-firms, with the overall volume of information increasing in the number of firms. The quality of information exchanged is enhanced by the ability level of the agents involved, which is measured by the average level of ability in the city.[12]

Workers are employed in identical occupations, performing identical tasks. However, some workers are simply better than others. As this is a static model, the source of this difference is not specified.

[9]This is a mapping from underlying ability level to efficiency units of labor given current technology levels. Later in the paper, the effects of technological changes that affect this relationship are considered.

[10]This is assumption is made to avoid any potential coordination problem between local developer, worker and firm decisions.

[11]See below for the reason for this parameter restriction.

[12]Using instead the total stock of human capital rather than the average would not affect the analysis. In this case, n_j would simply be raised to the power $\delta + \gamma$.

It could arise due to some underlying differences in innate talent across workers or from various social environment factors received before working age, such as parental and peer group effects or differences in local education quality.

2.3 Urban Spatial Structure

Given the existence of external scale economies, it is necessary to have some other force restricting city size. Offsetting benefits of population agglomeration are congestion costs. I use a simple monocentric circular city model, standard in the urban literature, in which increasing commuting costs are the basis for scale diseconomies.[13] An intuitive description of the model is that the boundary of the city is pushed farther and farther outward as its population increases, leading to longer commuting times for those at the city's edge since all workers must commute to the center of the city in order to work. The differing commuting times for those living at different distances from the city center are compensated for by corresponding differences in rent levels. Workers living near the center pay high rents and low commuting costs and workers at the city edge pay low rents and high commuting costs; equilibrium requires that the sum of rent and commuting costs is equal for all individuals. Both average land rents and average commuting costs are increasing in city size. Details of the formal derivation of this standard model are footnoted below,[14] but the important results, expressions for total commuting costs and land rents as a function of city size, can be found in equations (2) and (3) respectively.

$$\text{total commuting costs} = bn^{\frac{3}{2}} \tag{2}$$

$$\text{total land rents} = \frac{1}{2}bn^{\frac{3}{2}} \tag{3}$$

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

[13]A standard objection to this type of model is the existence of metropolitan areas with multiple employment sub centers and "edge cities". This phenomenon is an interesting one in its own right, and has been analyzed in works such as Fujita and Ogawa (1982) and Henderson and Mitra (1997). This stylized model is used as a simple microfoundation of urban congestion.

[14]Assume all production occurs at a point in the center of a city, the Central Business District (CBD). Workers live in fixed lots of size one surrounding the CBD in concentric circles and pay a cost τ per unit distance u to commute to work. Given identical inelastic demand for land and no housing, commuting costs give rise to a rent gradient $R(u)$ such that in equilibrium all individuals pay identical rent plus commuting costs - $R(u) + \tau u$ must be the equal for all u. Normalizing the alternative land use value to 0, rent at the edge of the city $R(u_1) = 0$, where u_1 denotes the distance from the CBD to the edge of the city. Thus, $R(u) + \tau u = \tau u_1$, or $R(u) = \tau(u_1 - u)$. Total land rents are equal to $\int_0^{u_1} 2\pi u R(u) = \frac{1}{2}\pi\tau u_1^3$. Total commuting costs are equal to $\int_0^{u_1} 2\pi u(\tau u) = \frac{2}{3}\pi\tau u_1^3$. Given total population $n = \pi u_1^2$, substituting provides $u_1 = \pi^{-1/2}n^{1/2}$. Substituting provides (2) and (3).

3 Equilibria in the Basic Model

It is necessary still to specify the process of city formation. In the literature, there are two standard approaches to this problem. One is to assume that cities are formed in a competitive market by real estate developers (cf. Helsley and Strange (1990), Rauch (1993a)). These developers establish cities by attracting residents through their choice of city characteristics, recognizing the externality effects. The other approach, favored in such works as Krugman (1992), is to assume that the economy is "self-organizing", that cities form through the movements of atomistic agents. For the basic model in this paper, the equilibrium spatial allocations are analyzed under both mechanisms - through the actions of land developers and by self-organization. Equilibria given city formation by large agents, land developers, are considered first followed by examination of equilibria in the self-organizing economy. For the former, when conditions for existence hold (they may not), there is a unique equilibrium that may or may not be efficient. For the latter, a continuum of potential equilibria, generally inefficient, exists. Other than this efficiency issue, the basic qualitative results on sorting do not depend on the specification of the mechanism of city formation.

3.1 Land Developers

Throughout this subsection I assume that cities are formed in a competitive market by developers who set city size, collect all land rents, and offer lump-sum subsidies for worker-firms to locate in their cities.[15] There are an unexhausted number of identical sites on which to develop cities. Developers control only one site and must provide its residents with the market rate of compensation. Free-entry in this competitive market drives developer profits to zero.

I assume that developers cannot observe worker quality. Or, equivalently, that institutions are such that they are unable to offer discriminatory contracts to different ability types.[16] In this, the problem facing developers is similar to that facing insurance firms in Rothschild and Stiglitz (1976). Developers offer contracts consisting of city sizes n and subsidies T based on knowledge of a_h, a_ℓ and z, the ability levels and proportions of each type of worker in the economy. Since firm output is increasing in the proportion of high ability workers in the city, wages received by workers depend on the composition of the local labor force. Through the terms of the contract, developers can attempt to screen workers by type and establish segregated cities. Alternatively, developers can offer terms designed to attract both types of workers and establish mixed cities. Self-selection by real income maximizing workers in this competitive market determines the equilibrium set of contracts and cities.

[15] Garreau (1991) reports on the extent that modern "edge cities" are formed through the actions of such private development agents, though there are less recent historical examples as well. Henderson and Mitra (1997) report that these new cities cities are generally formed by a single development company.

[16] Such direct discrimination would likely face legal constraints.

The requirements for a spatial equilibrium allocation of workers are as follows:[17]

1. No developer can offer a contract that earns negative expected profit in equilibrium.

2. There must be no city size and subsidy level outside of the equilibrium set that a developer can offer which can attract workers away from existing cities and earn non-negative expected profit if announced.

3. There must not be an incentive for any worker to switch cities.

There are three possible equilibrium outcomes. The first possibility is a *natural separation equilibrium* in which there are two types of segregated cities with each type being efficiently sized.[18] In a natural separation equilibrium, city sizes are set optimally and correspond to an alternative specification in which developers have the ability to observe and directly exclude workers by type as well as to the solution of a planner allocating workers in order to maximize total output net of commuting costs. The second possibility, an *inefficient separation equilibrium* is also an equilibrium with segregated cities. However, in an inefficient separation equilibrium high ability cities are set at inefficient sizes, reducing incomes for high ability individuals; the level of compensation for low ability workers is unaffected. The third possibility is that there is no equilibrium spatial allocation of workers. As in Rothschild and Stiglitz's insurance model, there are no mixed city equilibria in this model. This is demonstrated below.

Consider the formal problem of a land developer attempting to form segregated high ability cities by setting the terms of the contract to screen out low ability workers. Subscripting h to denote high ability cities and individuals, the problem is:

$$\max_{n_h, T_h} \Pi_h = \frac{1}{2} b n_h^{\frac{3}{2}} - T_h n_h \tag{4}$$

$$\text{s.t.} \quad W_h + T_h - \frac{3}{2} b n_h^{\frac{1}{2}} = I_h \tag{5}$$

$$W_{\ell h} + T_h - \frac{3}{2} b n_h^{\frac{1}{2}} \leq I_\ell \tag{6}$$

Developers choose city size n_h and per person subsidy T_h in order to maximize total land rents received (from (3)) minus total subsidies paid. Given the competitive nature of land development markets, in order to attract workers these are chosen subject to the constraint in (5) which states that high ability workers receive real incomes equal to their opportunity cost I_h. Real incomes are equal to wages received by high ability individuals W_h plus transfers T_h minus total rent and commuting expenditures ($\frac{3}{2} b n_h^{\frac{3}{2}}$ from (2)+(3)). The second constraint in (6) is the low ability self-selection constraint which states that

[17] Helsley and Strange (1990) demonstrate that this type of problem can be set up more formally as a staged game.

[18] The number of each type of city depends on the total populations of each type of worker. Any integer problems are ignored and I assume that the population is large enough so that there exists a "large" number of cities of each type in equilibrium, necessary for the assumption of a competitive market.

wages earned by a low ability individual working in an otherwise all high ability city, $W_{\ell h}$, plus transfers minus living costs in that city must be less than or equal to the opportunity cost of low ability workers.

3.1.1 Natural Separation Equilibrium

In a natural separation equilibrium, parameters are such that the choice of n_h in (4) by a developer subject to the constraint in (5) is sufficient to deter entry by low ability individuals - the self-selection constraint for low ability workers in (6) is non-binding in this type of equilibrium. I first solve the developer's problem ignoring this constraint, and then determine the conditions under which it is not violated by the solution. First, solve for T_h in (5) and substitute into (4). Since this is a segregated high ability city, $z = 1$ and therefore $W_h = Dn_h^{\delta} a_h^{\gamma+1}$. Rewriting:[19]

$$\max_{n_h} \Pi_h = Dn_h^{\delta+1} a_h^{\gamma+1} - bn_h^{\frac{3}{2}} - I_h n_h \tag{7}$$

The zero-profit condition, resulting from free-entry into the land development market, requires that total worker real incomes are equal to the city's aggregate output net of commuting costs. Solving provides the following results:[20]

$$\begin{aligned} n_h &= [2b^{-1}\delta D a_h^{\gamma+1}]^{\frac{2}{1-2\delta}} \\ W_h &= Ba_h^{\frac{\gamma+1}{1-2\delta}} \\ I_h &= (1-2\delta)Ba_h^{\frac{\gamma+1}{1-2\delta}} \end{aligned} \tag{8}$$

where $B \equiv (2b^{-1}\delta D^{\frac{1}{2\delta}})^{\frac{2\delta}{1-2\delta}}$.

Developers transfer all land rents to workers, a version of the "Henry George Theorem",[21] and consequently their real incomes are equivalent to their wages minus the average level of commuting costs, defined as each city's cost of living. From the equations in (8) we can see that the ratio of real incomes to wages is equal to $(1-2\delta)$.

Solving the analogous problem for developers establishing segregated low ability cities provides the following results, subscripting ℓ to indicate a low ability type city: $n_\ell = [2b^{-1}\delta D a_\ell^{\gamma+1}]^{\frac{2}{1-2\delta}}$, $W_\ell = Ba_\ell^{\frac{\gamma+1}{1-2\delta}}$, $I_\ell = (1-2\delta)Ba_\ell^{\frac{\gamma+1}{1-2\delta}}$.

[19] Note that this reformulated problem is effectively one of choosing city size in order to maximize total city output net of commuting costs and incomes paid to workers.

[20] Taking the first order condition provides $(\delta+1)Dn_h^{\delta}a_h^{\gamma+1} - \frac{1}{2}bn_h^{\frac{3}{2}} = I_h$. Substituting this into the zero-profit condition for land developers give us $\delta Dn_h^{\delta+1}a_h^{\gamma+1} = \frac{1}{2}bn_h^{\frac{3}{2}}$. This provides n_h and substitution into the first order condition provides the rest. The second order condition is satisfied as long as $\delta < 1/2$, placing a restriction on δ. If $\delta > 1/2$, scale benefits always offset congestion effects and there will be a single city. The existence of multiple cities is required in order for a competitive land development market to exist.

[21] City sizes are set so that the difference between the social and marginal products of a firm locating in a city are equal to total land rents, $\delta Dn^{\delta+1}a_h^{\gamma+1} = \frac{1}{2}bn_h^{\frac{3}{2}}$, a statement of the "Henry George Theorem" (Flatters, Henderson and Mieszkowski (1974),Stiglitz (1977)).

Note that the ratio of city sizes, n_h/n_ℓ, is equal to $(\frac{a_h}{a_\ell})^{\frac{2}{1-2\delta}} > 1$. Consequently, the ratio of average commuting cost expenditures, the cost of living paid by all workers, between the two types of cities is equal to $(\frac{a_h}{a_\ell})^{\frac{1}{1-2\delta}} > 1$. This increased cost of living associated with larger cities is the factor that potentially deters low ability individuals from wanting to live in high ability cities. Low ability workers face a trade-off. On one hand, by locating in high ability cities they can increase their wages due to the higher level of average human capital in the city. On the other hand, high ability cities are larger and thus have a higher cost of living. If the increase in living costs more than offset the increase in wages, then the above allocation is an equilibrium.

Proposition 1. *If the difference between the ability levels of the two types of workers is high enough, a natural separation equilibrium exists and all workers live in homogeneous cities of sizes n_h and n_ℓ.*

Proof. There must be no other contracts other than $\{n_h, T_h\}$ and $\{n_\ell, T_\ell\}$ that can attract workers away from existing cities and earn non-negative profit for a developer. Additionally, the self-selection constraints for workers must be satisfied. To demonstrate that the former condition holds first note that given segregation, city sizes n_h and n_ℓ correspond to the solutions to the problems of choosing city sizes to maximize per person real income in each type of city.[22] Therefore, no segregated cities can be established that would provide higher real incomes for either type. Next note that as $\frac{\partial I_i}{\partial z_j} = \gamma(a_h - a_l)D[z_j a_h + (1-z)a_\ell]^{\gamma-1} n_j^\delta a_i > 0$, for any city size n_j, real incomes for both types of workers are increasing in z_j, the proportion of high ability workers in the city. Thus, no mixed city exists that can offer greater compensation than I_h for high ability workers. Though some values of z_j and n_j may provide greater real incomes for low ability workers, no contract of this type can be offered that can attract high ability workers away from segregated cities of size n_h.

Having established that there never exists an incentive for high ability workers to move to a city with low ability workers, it is still necessary to show the conditions under which the low ability worker self-selection constraint in (6) is satisfied given this spatial allocation. A low ability worker in an otherwise high ability city of size n_j receives wages $W_{\ell h} = D n_h^\delta a_h^\gamma a_\ell$ and after receiving transfers pays living costs of $b n_h^{\frac{1}{2}}$. Substituting in for n_h and I_ℓ into (6) provides:

$$B a_h^{\frac{\gamma+2\delta}{1-2\delta}} a_\ell - 2\delta B a_h^{\frac{\gamma+1}{1-2\delta}} \leq (1-2\delta) B a_l^{\frac{\gamma+1}{1-2\delta}} \qquad \text{low ability self-selection constraint} \tag{9}$$

Define $I_{\ell h}$ as the LHS of (9), the income that low ability individuals receive by locating in a high ability city. When $a_h = a_\ell$, $I_{\ell h} = I_\ell$. The partial derivative of $I_{\ell h} - I_h$ with respect to a_h is:

$$\frac{\partial I_{\ell h}}{\partial a_h} = \left(\frac{\gamma+2\delta}{1-2\delta}\right) a_h^{\frac{\gamma+4\delta-1}{1-2\delta}} a_\ell - 2\delta \left(\frac{\gamma+1}{1-2\delta}\right) a_h^{\frac{\gamma+2\delta}{1-2\delta}} \tag{10}$$

When $a_h = a_\ell$, this expression is positive, with wage changes dominating commuting cost changes. However, $I_{\ell h}$ reaches a maximum at $\frac{a_h}{a_\ell} = \left[\frac{\gamma+2\delta}{2\delta\gamma+2\delta}\right] > 1$ and is monotonically decreasing after that,

[22] These sizes correspond to the solutions to $\max_{n_h} D n_h^\delta a_h^{\gamma+1} - b n_h^{\frac{1}{2}}$ and $\max_{n_\ell} D n_\ell^\delta a_\ell^{\gamma+1} - b n_\ell^{\frac{1}{2}}$.

reaching 0 when $\frac{a_h}{a_\ell} = \frac{1}{2\delta} > 1$. For any given a_ℓ, eventually $I_{\ell h}$ must fall below I_ℓ somewhere on the interval where $\left[\frac{\gamma+2\delta}{2\delta\gamma+2\delta}\right] < \frac{\overset{*}{a}_h}{a_\ell} < \frac{1}{2\delta}$ and thus the self-selection constraint holds for all $a_h > \overset{*}{a}_h$, where $\overset{*}{a}_h$ solves (9) with equality. □

Thus, Proposition 1 is established.

Proposition 2. *An economy consisting of* $m_h = \frac{zN}{n_h}$ *high ability cities of size* n_h *each and* $m_\ell = \frac{(1-z)N}{n_\ell}$ *low ability cities of size* n_ℓ *is efficient, producing the maximum possible output net of commuting costs.*

Proof. To prove this, it is sufficient to demonstrate that starting from a spatial allocation of workers corresponding to a natural separation equilibrium, there is no way to reallocate workers and increase net output in the economy. Consider a planner reallocating a subset of the population N_p, a proportion z_p of which are high ability workers (with z_p not necessarily equal to z). I will show that for all values of N_p and z_p, the efficient allocation of workers involves segregated cities of size n_h and n_ℓ. The planner can either optimally assign these workers to segregated cities,[23] providing a total net output of $(1-2\delta)BN_p[z_p a_h + (1-z_p)a_\ell]^{\frac{\gamma+1}{1-2\delta}}$. Or, a planner can allocate these workers to the mixed city solution that maximizes output by solving the following problem, choosing the number of mixed cities m_p and number of workers in each city n_p, given N_p and z_p.

$$\max_{m_p, n_p} \quad m_p D[z_p a_h + (1-z_p)a_\ell]^{\gamma+1} n_p^{\delta+1} - m_p b n_p^{\frac{3}{2}} \tag{11}$$

$$\text{s.t.} \quad m_p n_p = N_p \tag{12}$$

Substituting the population constraint in (12) into (11), taking first order conditions and solving provides a total net output of $(1-2\delta)BN_p[z_p a_h + (1-z_p)a_\ell]^{\frac{\gamma+1}{1-2\delta}}$. The level of output that is obtained from the mixed solution is less than that obtained from the separating one when:

$$[z_p a_h + (1-z_p)a_\ell]^{\frac{\gamma+1}{1-2\delta}} \quad < \quad z_p a_h^{\frac{\gamma+1}{1-2\delta}} + (1-z_p)a_\ell^{\frac{\gamma+1}{1-2\delta}}$$

This is always true given convexity ($\frac{\gamma+1}{1-2\delta} > 1$). Since this is true *for all* N_p and z_p, for any splinter of the population composed of any proportions of ability types, separation always produces greater net output; the natural separation equilibrium is equivalent to the planner's solution. Note that the ratio of the two terms above represents the efficiency loss obtained from mixing. □

[23] The problem:

$$\begin{aligned} \max_{m_\ell, m_h, n_\ell, n_h} \quad & m_h D n^{\delta+1} a_h^{\gamma+1} + m_\ell D n^{\delta+1} a_\ell^{\gamma+1} - m_h b n_h^{\frac{3}{2}} - m_\ell b n_\ell^{\frac{3}{2}} \\ \text{s.t.} \quad & m_h n_h = zN \\ & m_\ell n_\ell = (1-z)N \end{aligned}$$

where m_i is the number of cities of type i. The city sizes that solve this problem are equivalent to those established in a natural separation equilibrium.

The efficiency of segregation results from the lack of any complementarity between worker types. In a model closely related to this one, Henderson and Becker (1996) demonstrate that when there are different occupations, such as entrepreneurs and workers, that are complementary inputs to production, then a symmetrical allocation of mixed cities is both efficient and the only equilibrium allocation. This is echoed in Berglas (1976) who shows in a model of club good provision that if workers are essential complementary inputs in production, then heterogeneous communities are optimal. Brueckner (1994) extends this work to allow for the possibility of non-essential complementarity between types and demonstrates that homogeneity may or may not be optimal depending on the degree of complementarity.

Note that in mixed cities, different city sizes will maximize the net incomes for different worker types,[24] and the above comparison is derived from the city size that maximizes output given a symmetric allocation of mixed cities. The ratios of incomes in mixed cities to segregated cities for each type of worker given such an allocation are:

$$\frac{I_{\ell m}}{I_\ell} = \frac{[z_p a_h + (1-z_p)a_\ell]^{\frac{\gamma+2\delta}{1-2\delta}}}{a_\ell^{\frac{\gamma+2\delta}{1-2\delta}}} > 1 \qquad \frac{I_{hm}}{I_h} = \frac{[z_p a_h + (1-z_p)a_\ell]^{\frac{\gamma+2\delta}{1-2\delta}}}{a_h^{\frac{\gamma+2\delta}{1-2\delta}}} < 1$$

Low ability workers gain and high ability workers lose relative to the efficient segregated allocations.

3.1.2 Inefficient Separation

When $a_h < \overset{*}{a}_h$ the difference between ability levels is such that the cost of living in efficiently sized high ability cities does not the deter entry of low ability individuals. Remember that a developer's only instruments are choice of n and T, city size and subsidy. Ability types are unobservable, or institutions are such that developers are unable to directly exclude workers based on ability. In addition, an equilibrium in a competitive land development market requires that all land rents be transferred back to the city residents. Given these constraints, two options are potentially available. The first option is that developers can attempt to form mixed cities. In this case, developers choose n and T based on the expectation that the composition for the local population is the same as the national population. That is, for city j, $E(z_j) = z$. The other option for developers is to establish inefficiently sized cities in order to deter entry by low ability workers.

Examining the latter case first, consider the reduced form problem of a developer choosing $\hat{n}_h$ to deter entry of low ability workers:

$$\max_{\hat{n}_h} \qquad D\hat{n}_h^{\delta+1}a_h^{\gamma+1} - b\hat{n}_h^{\frac{3}{2}} - \hat{n}_h\hat{I}_h \tag{13}$$

$$\text{s.t.} \qquad D\hat{n}_h^{\delta}a_h^{\gamma}a_\ell - b\hat{n}_h^{\frac{1}{2}} \leq B(1-2\delta)a_\ell^{\frac{\gamma+1}{1-2\delta}} \tag{14}$$

[24]Becker and Henderson (1996) explore the trade-off between the welfare of workers and entrepreneurs who are complementary inputs in production for city sizes.

When the constraint in (14) (which is derived from (6) after substituting in for $W_{\ell h}$ and $T_h = \frac{1}{2}bn_h^{\frac{1}{2}}$) is non-binding, this solution is $\hat{n}_h = n_h$. When the constraint binds there are two values of $\hat{n}_h$ that satisfy this condition, one less than and one greater than n_h; cities are either too large or too small. In order to attract workers, developers will offer the contract that maximizes per worker income. The difference between incomes potentially earned by both types of workers in a segregated city is equal to $\hat{I}_h - \hat{I}_{\ell h} = D\hat{n}_h^{\delta}a_h^{\gamma}(a_h - a_\ell)$. Since this is increasing in $\hat{n}_h$, for a given level of low ability income, namely $\hat{I}_{\ell h} = I_\ell$, it is the larger value of $\hat{n}_h > n_h$ which satisfies (14) that maximizes high ability workers' incomes.

This establishes that when $a_h < \overset{*}{a}_h$, the segregated allocation that maximizes the incomes of high ability workers contains high ability cities of size $\hat{n}_h$, where $\hat{n}_h$ is the larger of the two values which satisfy $D\hat{n}_h^{\delta}a_h^{\gamma}a_\ell - b\hat{n}_h^{\frac{1}{2}} \leq B(1-2\delta)a_\ell^{\frac{\gamma+1}{1-2\delta}}$. No explicit solution for this can be found, so the resulting incomes for high ability workers $\hat{I}_h$ can only be defined implicitly as a function of $\hat{n}_h$.

Is this an equilibrium? High ability workers must live in inefficiently sized cities in order to segregate themselves from low ability workers. Incomes may be greater for high ability workers in mixed cities than in cities designed to exclude low ability workers. The price paid for segregation may be too great. This is the case if z is "large", if there are not many low ability workers in the economy. Then, the cost to high ability workers of letting a few low ability workers into the city is less than the cost of excluding them. This leads to the following proposition:

Proposition 3. *When $a_h < \overset{*}{a}_h$ and the proportion of low ability workers $(1-z)$ large enough, a separating equilibrium exists that is characterized by efficiently sized low ability cities and inefficiently over-sized high ability cities.*

The basic proof is in the appendix. One portion of the proof is established above, namely that there exists an $\hat{n}_h$ such that the low ability self-selection constraints are satisfied. What remains is to derive the conditions under which there does not exist a city size and subsidy that a developer can offer that can attract both types of workers, providing incomes greater than I_ℓ and $\hat{I}_h$ for each type respectively.

Proposition 4. *When $a_h < \overset{*}{a}_h$ and the proportion of low ability workers is sufficiently small, there exists no equilibrium spatial allocation of workers.*

The proof is in the appendix. Proposition 4 rules out the possibility of an equilibrium with mixed cities. The argument that establishes this follows from Rothschild and Stiglitz (1976). Given that developers are offering a set of contracts to establish mixed cities, it is always the case that opportunity exists for a developer to enter the market and earn a profit. By offering a subsidy level less than the average rent level, a developer can earn positive profit and attract high ability workers away from existing cities while deterring entry by low ability workers. However, this profit-earning contract will itself violate the conditions for a competitive equilibrium.

3.2 Self-Organization

In the absence of large land development agents, a wider range of equilibria are possible. Requirements for a Nash (free mobility) equilibria in a self-organizing economy are that no individual has an incentive to move and that workers must at least earn their endowments, normalized to zero. These requirements admit a continuum of city sizes into the set of possible equilibria. Additionally, depending on the parameter values, mixed and segregated cities may both exist simultaneously in the economy.

However, given free mobility of the population the requirement that these equilibria are locally stable, impervious to small deviations in the spatial allocation of the population, is desirable.. One approach to doing this is to examine perturbations to the population of one city in a partial equilibrium framework, assuming that national market compensation of the factors of production are unaffected by these deviations. Define equilibrium national market compensation for high and low ability workers as $\bar{I}_h$ and $\bar{I}_\ell$ respectively, corresponding to city populations of $\bar{n}_h$ and $\bar{n}_\ell$ (where for a given mixed city, total city size $\bar{n} = \bar{n}_h + \bar{n}_\ell$ and $\frac{n_h}{n_\ell} = \frac{z}{1-z}$). Assume that by perturbing the populations in a city that compensation rates in that city adjust instantaneously, leaving national compensation rates unaffected. Assume a simple dynamic adjustment process for populations in a city such that $\dot{n}_h = d(I_h - \bar{I}_h)$ and $\dot{n}_\ell = d(I_\ell - \bar{I}_\ell)$ where d is the speed of worker migration. Workers respond to differentials in local versus national compensation rates by moving in or out of the city. An equilibrium allocation is stable if worker movements are such that the original equilibrium is restored after a small deviation.

Proposition 5. *In a self-organizing economy mixed equilibria are unstable.*

The proof is in the appendix.

Having ruled out equilibria with mixed cities, the remaining possible spatial allocation is of course segregated cities. Unlike the regime with land developers, in a self-organizing economy a continuum of robust Nash equilibria are possible. With segregated equilibria, city sizes tend to be inefficiently large.

For a segregated allocation to satisfy the conditions of a Nash equilibrium, neither type of worker must have any incentive to move, either to a city of his own type or to another type of city. Additionally, workers must earn an income greater than their endowments, normalized to zero.

For individuals of type i, a Nash equilibrium requires that $\frac{\partial \tilde{I}_i}{\partial \tilde{n}_i} < 0$, given self-organizing city sizes and incomes $\tilde{n}_i$, $\tilde{I}_i$ respectively. City sizes must be such that no worker in a segregated city of his type has an incentive to move to another city of that type. This places lower bounds on city sizes for each type, $\tilde{n}_i^{min} = [2b^{-1}\delta D a_i^{\gamma+1}]^{\frac{2}{1-2\delta}}$.[25] This lower bound corresponds to the equilibrium natural separation city sizes given land developers. Additionally, incomes must be greater than the endowment of zero. This imposes an upper bound on city sizes of each type $\tilde{n}_i^{max} = [b^{-1} D a_i^{\gamma+1}]^{\frac{2}{1-2\delta}}$.[26]

[25] $\frac{\partial \tilde{I}_i}{\partial \tilde{n}_i} = \delta D \tilde{n}_i^{\delta-1} a_h^{\gamma+1} - \frac{1}{2} b \tilde{n}_i$.
[26] $\tilde{n}_i^{min} < \tilde{n}_i^{max}$ given the requirement that $\delta < \frac{1}{2}$.

The final requirement for a Nash equilibrium is that no worker has an incentive to move to a city of a different type. This requires:

$$\begin{aligned} D\tilde{n}_h^{\delta} a_h^{\gamma} a_\ell - b\tilde{n}_h^{\frac{1}{2}} &\leq D\tilde{n}_\ell^{\delta} a_\ell^{\gamma+1} - b\tilde{n}_\ell^{\frac{1}{2}} \\ D\tilde{n}_\ell^{\delta} a_\ell^{\gamma} a_h - b\tilde{n}_\ell^{\frac{1}{2}} &\leq D\tilde{n}_h^{\delta} a_h^{\gamma+1} - b\tilde{n}_h^{\frac{1}{2}} \end{aligned}$$

The local stability of any allocation satisfying the Nash equilibrium requirements is established in the appendix.

Proposition 6. *A segregated self-organizing equilibrium always exists.*

Proof. Setting $\tilde{n}_\ell = n_\ell^{min}$ and $\tilde{n}_h = \hat{n}_h$ satisfies the self-selection constraints as demonstrated in the proof of Proposition (3) and fall within the range of city sizes that provides no incentive for individuals of a given ability type to move to a city of that same type. □

In general, a continuum of potential equilibria exist. Equilibrium city sizes must be such that workers do not have incentive to move either to a city of their own type or to another type of city. The former requires simply that $\tilde{n}_i \in [\tilde{n}_i^{min}, \tilde{n}_i^{max}]$ for a person of type i. The latter requires that in addition the self-selection constraints are satisfied. This potentially shrinks the upper bound on low ability city sizes and increases the lower bound on high ability city sizes. High ability cities must be "big" relative to low ability cities in order to deter entry by low ability workers. The upper bound on high ability city sizes potentially forces the upper bound on low ability cities to be smaller than $\tilde{n}_\ell^{max}$ in order to satisfy the self-selection constraints.

Under self-organization there are a continuum of potential equilibrium city sizes. Both types of cities are in general too large. Since equilibria have to satisfy the self-selection constraints, high ability cities will tend to be "more inefficient" than low ability cities, though in general both will be inefficiently large.

4 Implications for Income Inequality

As we have seen, spatial sorting by ability type augments income differences between workers of different ability level. This motivates an analysis of how this spatial sorting affects the income distribution and how changes in technology affect the degree of true and measured inequality within this framework.

4.1 The Causes of Widening Inequality

There is no consensus regarding the underlying causes of widening income inequality. An appeal to the usual suspect - technological change, or specifically in this case skill-biased technological change benefiting high skilled workers - has not been unequivocally supported by solid empirical evidence (DiNardo and Pischke 1997). One puzzling aspect of this phenomenon is that in addition to inequality rising between different groups of individuals with identical observable characteristics, the degree of inequality within these groups has also been rising. This model has explored how stratification of the population by ability type across different cities can augment productivity differences across observably identical workers in the employed in identical occupations. Such sorting can therefore affect this "within group" inequality.

An alternative explanation for widening inequality suggested by this model is that it is not simply technological changes affecting returns individual skills that could affect the wage gap but also technological changes enhancing the benefits of interaction and increasing the magnitude of human capital externalities of the type found in this paper. In the modern service economy, where both inputs and outputs of production are to some degree simply information, it is perhaps the case that knowledge spillovers have an increasingly powerful effect on productivity. [27] In fact, Ciccone et al. (1999) finds both that city-wide public returns to human capital both exceed the private ones and that they may have increased as much as 50% from 1980 to 1990. To demonstrate, the ratio of incomes given a natural separation equilibrium in this model is:

$$\frac{I_h}{I_\ell} = \left(\frac{a_h}{a_\ell}\right)^{\frac{\gamma+1}{1-2\delta}}$$

Note that as $\frac{\gamma+1}{1-2\delta} > 1$, this ratio is convex in the ability ratio. Given sorting, scale economies and knowledge spillovers magnify any type of skill-biased technological changes.[28] Small changes in the returns to individual ability will lead to much larger changes in the real income ratio. This ratio is also increasing in γ, the elasticity of output with respect to the average level of ability in the city.[29]. Thus, even without any changes in returns to individual ability, an increase in the magnitude of human capital externalities will widen the income gap between different ability types.

4.2 An Issue of Measurement

The model raises an important issue concerning the measurement of income and income inequality. What is generally observed by an econometrician is $W^* = W + T$, total disposable income, while the appropriate measure for utility comparison is real income, I. This is the amount available for consumption, after

[27] As discussed in Gaspar and Glaeser (1996), there exists the notion that improvements in telecommunications technology may decrease the need for face-to-face interactions and thus remove the incentive for geographic concentration or sorting. However, the authors demonstrate that depending on the degree of complementarity between face-to-face and distant communications, the demand for face-to-face meetings might actually increase with improvements in telecommunications technologies, which in the context of this model would increase the magnitude of these externalities and thus the benefits to high ability individuals of spatial sorting.

[28] $\frac{d}{d\left(\frac{a_h}{a_\ell}\right)} = \frac{\gamma+1}{1-2\delta}\left(\frac{a_h}{a_\ell}\right)^{\frac{\gamma+2\delta}{1-2\delta}}$.

[29] $\frac{d}{d\gamma} = \left(\frac{1}{1-2\delta}\right)\left(\frac{a_h}{a_\ell}\right)^{\frac{(\gamma+1)}{(1-2\delta)}} \ln\left(\frac{a_h}{a_\ell}\right)$.

paying necessary subsistence expenditures. Comparing measures of the "wage gap" using measures of real income (I), wages (W), and disposable income (W^*) given the existence of a natural separation equilibrium we have:[30]

$$\begin{aligned} \Delta I_{h\ell} \equiv I_h - I_\ell &= (1-2\delta)B[a_h^{\frac{\gamma+1}{1-2\delta}} - a_\ell^{\frac{\gamma+1}{1-2\delta}}] \\ \Delta W_{h\ell} \equiv W_h - W_\ell &= B[a_h^{\frac{\gamma+1}{1-2\delta}} - a_\ell^{\frac{\gamma+1}{1-2\delta}}] \\ \Delta \overset{*}{W}_{h\ell} \equiv \overset{*}{W}_h - \overset{*}{W}_\ell &= (1+\delta)B[a_h^{\frac{\gamma+1}{1-2\delta}} - a_\ell^{\frac{\gamma+1}{1-2\delta}}] \end{aligned} \tag{15}$$

The ratio between the degrees of measured to true inequality can be written as:

$$\frac{\Delta \overset{*}{W}_{h\ell}}{\Delta I_{h\ell}} = \frac{1+\delta}{1-2\delta} \tag{16}$$

This demonstrates that failing to account for cost of living overstates the true degree of income inequality.[31] This comparison is made assuming the existence of a natural separation equilibrium. If the allocation is inefficient, using disposable incomes to measure wage inequality will overestimate the gap by an even greater amount. Inefficient separation implies that high ability cities are oversized, leading to reduced real incomes for high ability workers compared to their incomes under efficient separation. Both wages and subsidies are increasing in city size, thus for larger cities disposable income is higher while real income is lower in an inefficient separation.

That there are cost of living differences across cities is not a new idea, but given sorting by ability level, different groups are affected differently by living costs. Of course, in reality there are a variety of factors such as local amenities and fiscal conditions that may influence local price levels (Gyuorko and Tracy 1989). Thus, one would not advocate simply using cost of living indices as simple income deflators when examining wage inequality. In addition to being associated with higher productivity cities, variations in local living costs may also reflect variations in local amenities, among other things. However, the importance of attempting to account for living costs in some fashion is clear.[32]

In addition to this contemporaneous overstatement of the degree of income inequality, technological changes affecting the elasticity of output with respect to external scale economies, δ, will change this degree of mismeasurement. The degree of bias will change with the degree of scale economies. Thus, with technological change, comparisons of the degree of inequality across years will be biased accordingly.[33]

[30]If an efficient equilibrium does not exist, then this problem is even greater.

[31]In reality, cost of living differences across cities are non-trivial. Using a national base of 100, estimated indices for sample of cities for 1995 are as follows: Little Rock, AK - 87.0, Pensacola, FL - 93.8, Phoenix, AZ - 100.8, Denver, CO - 104.3, Portland, OR 109.1, Washington, DC - 124.6, Boston, MA - 139.2, and Manhattan - 221.1 (Source: ACCRA as published in the *1996 Statistical Abstract of the United States*).

[32]See Gyourko and Tracy (1991) and Glaeser (1998) for further discussion of these issues.

[33]The absolute degree of mismeasurement, as measured by $\Delta \overset{*}{W}_{h\ell} - \Delta I_{h\ell} = 3\delta B[a_h^{\frac{\gamma+1}{1-2\delta}} - a_\ell^{\frac{\gamma+1}{1-2\delta}}] > 0$ will be affected by other changes such as skill-neutral or skill-biased technological change (changes in D or the effective skill gap ($\frac{a_h}{a_\ell}$) respectively), or by changes in the importance of local knowledge spillovers, as measured by γ.

4.3 Trends in the Geographic Distribution of Human Capital

Like any relatively simple model, the predictions regarding the distribution of human capital are too sharp. The assumptions in the model ignore complications that would arise by considering other realistic features such as costs of moving, the age distribution of the population, heterogeneous preferences for climate and other locational attributes, and many more items that also may influence the location choices of households other than the reasons focused on in this paper.[34] However, given the basic ideas in the paper - that local knowledge spillovers influence location decisions, that these decisions in turn dictate the spatial distribution of human capital, which in turn have an influence on nominal and real wage and inequality across cities and worker types - it seems appropriate to ask how the spatial distribution has been evolving. If more highly educated workers are becoming more concentrated, then this would be suggestive of a contributing factor to rising income inequality.

To examine the changing spatial distribution of human capital, I calculate an index of dissimilarity (Duncan and Duncan 1955), that is more commonly used to measure the degree of racial segregation (for a recent application, see Cutler and Glaeser (1997)). I apply it here to examine how the level of segregation between college graduates and non-graduates has changed over time. The index takes the form $D.I. = \sum_j \frac{100}{2}|\frac{G_j}{G} - \frac{NG_j}{NG}|$ where G_j and NG_j are the number of graduates and non-graduates in each geographic unit j, respectively, and G and NG are the total number of each summed across all j. Dividing by 2 and scaling up by 100 provides this index with a very useful interpretation - its value is the percentage of college graduates who would have to move in order for the proportion of graduates in each j to match the population proportions. A value of 100 implies that there is complete segregation and a value of 0 implies that the population is perfectly integrated.

Table 1 presents calculations of this index for the years 1940-1990 excluding 1960, for which I did not have data, across different samples and specifications of the geographic unit.[35] Column 1 uses uses data covering the 318 PMSAs in existence in 1990.[36] Using this sample, over the twenty year period from 1950 to 1970 there is a 6.4% increase in the value of the index, demonstrating a slow trend of increasing educational segregation. From 1970 to 1990, decades of rising income inequality, the change is striking, with the index increasing in value by 28.4%. Given the concurrence of increasing educational segregation across cities with rising income inequality over the same time period (see for example Juhn et al. (1993)), it is tempting to link these trends.[37]

[34] One such complication, the introductory of a complementary input good, is addressed in the following section.

[35] For qualitative comparison of the magnitudes of this index, the value of the index for black/white racial segregation across all metropolitan areas in 1990 is 33.1 and across all metropolitan counties is 21.8 (Iaia 1997).

[36] All data is taken from county-level decennial census information over the period 1940-1990. Given changes in political boundaries over time, it is necessary to correct for this by merging counties that have split since 1940 and recombine counties that have since merged. These redefined counties are then aggregated up to the city level based on 1990 PMSA definitions. This decreases the total number of metropolitan counties by 13, but affects only one PMSA definition.

[37] Specifications that were tried include excluding any number of the largest or smallest cities or counties by both rank

Table 1: Index of Dissimilarity for College Graduates and Non-Graduates

	(1)	(2)	(3)
Year	**318 PMSAs**	**Top 100 PMSAs**	**PMSAs with urban population > 50K**
1940	10.2	9.0	9.7
1950	10.9	9.6	10.4
1970	11.6	10.3	11.5
1980	13.0	11.7	13.0
1990	14.9	13.5	14.9

5 Extension With Heterogeneous Cities

This section contains an extension of the basic model that exhibits internally heterogeneous cities in equilibrium. The goal of this extension is to demonstrate that by relaxing strong assumption of having only one occupation the basic model, the basic qualitative result of sorting can be preserved while having a more appealing outcome. This model strains the limits of tractability, so I restrict analysis to one possible equilibrium configuration under city formation by land developers, which is analogous to the natural separation equilibrium in the previous model.

Assume now that there are two occupations. Workers can either be employed in the production of the final consumption good or a locally traded intermediate input good, non-transportable across cities.[38] All workers, regardless of ability type, can become employed in either occupation. In addition, all workers are equally adept in the production of the intermediate input good, while ability differences only affect productivity for those employed in the final consumption good.

5.1 Production

Assume that intermediate input workers produce on unit of this good, regardless of ability type. Output for an final good producer of ability of type i working in city j is:

$$y_{ij} = D[\bar{a}_j^{\gamma} n_j^{\delta}] u_{ij}^{\alpha} a_i \tag{17}$$

where in this case z_j is the proportion of final good producing workers n_j that are high ability and u_i is the amount of intermediate input used by firm i. Denote the number of final good producing workers in city j as $q_j = n_j u_{ij}$.

and population size.

[38]An equivalent formulation would be to have final goods producers hire the intermediate good producing workers directly.

In this specification, scale and ability externalities are intra-industry, accruing only from workers producing the final good. The intermediate input good is produced in cities only because it is assumed to be non-transportable across cities. One can think of this as encompassing a variety of locally traded services, such as restaurant and hotel services, trash removal, or local business services.

5.2 An Equilibrium with Efficient City Sizes

This more complicated model strains the limits of tractability. I examine one possible equilibrium configuration. In this case, there are two types of cities. Given complementarity of worker types (and the non-tradeability of the intermediate input good), all cities contain both workers in both occupations. One type of city will contain all low ability individuals. The other type will contain high ability final good producers and low ability intermediate input producers.

This configuration requires assuming both that the population is "large", and that the proportion of the population that are low ability workers is not "too small". The first assumption is made to ensure that there are multiple cities and the second guarantees that there are enough low ability workers available to fill all of the available intermediate input occupations in high ability cities. In addition, as the heading of this section suggests, only efficient allocations are considered, corresponding to the case of a natural separation equilibrium in the basic model. The conditions for this are derived, and then are assumed to hold. Other potential equilibria (or non-existence) could arise if these assumptions do not hold. These other possibilities are analogous to the possible configurations of the basic model.

I solve first for the characteristics of the all low ability cities, subscripted ℓ as before. Final good producing workers in these cities choose a quantity u of intermediate input in order to maximize profit:

$$\max_{u_\ell} \pi_\ell \quad = \quad D[n_\ell^\delta a_\ell^\gamma] u_\ell^\alpha a_\ell - P_\ell u_\ell \tag{18}$$

Final output producing firms choose an amount u_ℓ of the intermediate input, priced at P_ℓ (the price of the intermediate input in general varies across cities), in order to maximize profit. First order conditions require:

$$\alpha D n_\ell^\delta u_\ell^{\alpha-1} a_\ell^{\gamma+1} = P_\ell \tag{19}$$

As intermediate input workers produce one (normalized) unit of these good, P_ℓ is also equal to their wages from firm operations.

Given this, a land developer faces the following rent maximization problem, choosing final good workers n_ℓ and intermediate input workers q_ℓ:

$$
\begin{aligned}
\max_{n_\ell, q_\ell, T_{\ell n}, T_{\ell q}} \Pi &= \frac{1}{2} b(n_\ell + q_\ell)^{\frac{3}{2}} - T_{\ell n} n_\ell - T_{\ell q} q_\ell && (20) \\
\text{s.t.} \quad & (1-\alpha) D n_\ell^\delta u_\ell^\alpha a_\ell^{\gamma+1} + T_{\ell n} - \frac{3}{2} b(n_\ell + q_\ell)^{\frac{1}{2}} = I_\ell && (21) \\
& P_\ell + T_{\ell q} - \frac{3}{2} b(n_\ell + q_\ell)^{\frac{1}{2}} = I_{\ell q} && (22) \\
& n_\ell u_\ell = q_\ell && (23)
\end{aligned}
$$

where $T_{\ell n}$ and $T_{\ell q}$ are the land developer's choice of subsidies for each type. The constraint in (21) states that firm profits for final good producing workers must be equal to their opportunity cost, (22) states the same for intermediate input workers, and (23) equates supply and demand of the intermediate input within the city.

Substituting (21) and (22) (noting that $P_\ell = \alpha D n_\ell^\delta u_\ell^{\alpha-1} a_\ell^{\gamma+1}$ from (19) and $q_\ell = n_\ell u_\ell$ from (23)) provides a reformulated maximization problem for land developers:

$$\max_{n_\ell, q_\ell} D n_\ell^\epsilon q_\ell^\alpha a_\ell^{\gamma+1} - b(n_\ell + q_\ell)^{\frac{3}{2}} - I_\ell n_\ell - I_q q_\ell \tag{24}$$

where $\epsilon \equiv \delta + 1 - \alpha$. First order conditions are then $\epsilon D n_\ell^{\epsilon-1} q_\ell^\alpha a_\ell^{\gamma+1} - \frac{3}{2} b(n_\ell + q_\ell)^{\frac{1}{2}} = I_\ell$ and $\alpha D n_\ell^\epsilon q_\ell^{\alpha-1} a_\ell^{\gamma+1} - \frac{3}{2} b(n_\ell + q_\ell)^{\frac{1}{2}} = I_q$ with the 0-profit condition for land developers requiring $\delta D n_\ell^\epsilon q_\ell^\alpha a_\ell^{\gamma+1} = \frac{1}{2} b(n_\ell + q_\ell)^{\frac{3}{2}}$. Substitution of first order conditions into (21) and (22) provides $T_{\ell n} = \delta D n_{\ell q}^{\epsilon-1} q_\ell^\alpha a_\ell^{\gamma+1}$ and $T_q = 0$.[39]

An analogous exercise can now be performed for the other type of city, containing high ability final good producers and low ability intermediate input workers. Results are as above, simply subscripting h instead of ℓ as needed. Equilibrium requires that the self-selection constraints for all workers are satisfied. There can be no incentive for any worker to switch occupations or cities. These are derived in detail in the appendix, but they are summarized here (defining $v_i = \frac{n_i}{q_i}$):

$$
\begin{aligned}
I_\ell = I_{\ell q} \quad &\rightarrow \quad v_\ell = \frac{\epsilon}{\alpha} && (25) \\
I_{qh} = I_\ell \quad &\rightarrow \quad \frac{B v_h^{\frac{\epsilon}{1-2\delta}} a_h^{\frac{\gamma+1}{1-2\delta}}}{(v_h+1)^{\frac{1+\delta}{1-2\delta}}} [\alpha(v_h + 1) - 3\delta] = I_\ell && (26) \\
I_{qh} \geq I_{\ell h} \quad &\rightarrow \quad v_h \geq \frac{(1-\alpha) a_h^{\frac{\gamma(1-\alpha)-\alpha}{1-\alpha}} a_\ell^{\frac{1}{1-\alpha}} + \delta a_h^{\gamma+1}}{\alpha a_h^{\gamma+1}} < \frac{\epsilon}{\alpha} && (27) \\
I_h \geq I_{qh} \quad &\rightarrow \quad v_h \leq \frac{\epsilon}{\alpha} && (28)
\end{aligned}
$$

In words, (25) states that workers in homogeneous low ability cities must earn identical real incomes, regardless of type. Constraint (26) states that low ability workers must earn identical incomes across all cities.[40] Constraint (27) states that low ability workers must earn a greater real income than they

[39] All land rents are transferred to final good producers, as only they contribute to local economies of scale.

[40] After substitutions, the income earned by low ability workers everywhere is $I_\ell = \frac{(1-2\delta)\epsilon^{\frac{\epsilon}{1-2\delta}} \alpha^{\frac{\alpha}{1-2\delta}} B a_\ell^{\frac{\gamma+1}{1-2\delta}}}{(\epsilon+\alpha)^{\frac{\epsilon+\alpha}{1-2\delta}}}$.

would earn by working as final goods producers in otherwise high ability cities. Finally, (28) states that high ability workers earn greater net income by working as final good rather than intermediate input produces

These are all straightforward except for (26), which has no explicit solution, complicating the analysis. To have an equilibrium with efficiently sized cities, the value of v_h that solves this must fall within the range dictated by (27) and (28), or:

$$v_h \in \left[\frac{(1-\alpha) a_h^{\frac{\gamma(1-\alpha)-\alpha}{1-\alpha}} a_\ell^{\frac{1}{1-\alpha}} + \delta a_h^{\gamma+1}}{\alpha a_h^{\gamma+1}}, \frac{\epsilon}{\alpha} \right] \tag{29}$$

5.3 Results from Extended Model

Though difficult to analyze as directly, the equilibrium in this extended model exhibits many of the same characteristics as in the basic model. There is real income inequality between high and low ability workers in the working in different cities in the final goods occupation. This income gap is overstated by cost of living differences across cities. Technological changes will affect this income gap similarly as before. In addition, there is nominal wage inequality only between low ability individuals working in different occupations and cities. Low ability workers in heterogeneous cities are paid more in wages than in homogeneous low ability cities, however this is precisely offset by cost of living differences across the two types of cities.

6 Conclusions

This paper has explored how the spatial distribution of human capital is determined and the role that spatial sorting and local human capital externalities may have on the widening of income inequality in recent decades. Workers of different ability types sort across different cities and local knowledge spillovers augment individual ability differences. This leads to inequality across cities between workers engaged in identical occupations.

Increases in the magnitude of local knowledge spillovers will increase the wage gap between workers of different abilities, even if individual ability levels are held constant. In addition, the effects of skill-biased technological change are augmented by human capital externalities given spatial sorting by type. However, some of the effects of these technological changes are offset by corresponding changes in cost of living differences across cities. Both the degree of and changes in real income inequality may be overstated if these cost of living differentials are not properly accounted for.

College graduates in the United States have become increasingly concentrated across U.S. cities in recent decades. This trend is concurrent with recent increases in the degree of income inequality. With knowledge spillovers, concentrations of human capital can influence the income gap. Thus, this may be a contributing factor to the rise in income inequality.

The static models presented here have obvious extensions in a richer, dynamic framework. The role of moving costs in determining the equilibrium distribution of workers is of interest. In addition, the effect of sorting on human capital investment is an area for future exploration. With economy-wide as well as local spillovers, though sorting is statically efficient, its disincentives on investment could potentially lead to reductions in long-run output levels.

7 Appendix

7.1 Proof of Proposition 3

First, it is necessary to show that $\hat{n}_h$ satisfies the self-selection constraints for high ability workers, that they would never have incentive to move a segregated low ability city of size n_ℓ. This requires:

$$D\hat{n}_h^{\delta} a_h^{\gamma+1} - b\hat{n}_h^{\frac{1}{2}} \geq B a_\ell^{\frac{\gamma+2\delta}{1-2\delta}} a_h - 2\delta B a_\ell^{\frac{\gamma+1}{1-2\delta}} \tag{30}$$

Remember that $\hat{n}_h$ is chosen to satisfy $D\hat{n}_h^{\delta} a_h^{\gamma} a_\ell - b\hat{n}_h^{\frac{1}{2}} = (1-2\delta) B a_\ell^{\frac{\gamma+1}{1-2\delta}}$. Subtracting this from (30) gives the condition that $D\hat{n}_h^{\delta} a_h^{\gamma}(a_h - a_\ell) \geq B a_\ell^{\frac{\gamma+2\delta}{1-2\delta}}(a_h - a_\ell)$, or:

$$\hat{n}_h \geq [2b^{-1}\delta D]^{\frac{2}{1-2\delta}} a_\ell^{\frac{\gamma+2\delta}{\delta(1-2\delta)}} a_h^{-\frac{\gamma}{\delta}} \tag{31}$$

Since $\hat{n}_h > n_h$, if we can show that the RHS of (31) is less than n_h, then the self-selection constraint is satisfied. After simplification, this requires simply that $a_\ell < a_h$, and thus the constraint is satisfied for all relevant parameter values.

This may not be an equilibrium. The possibility exists that a developer can offer a contract that will attract both types of individuals to a mixed city. Assume that when a developer chooses a city size in anticipation of both types of workers arriving in her city, the expected proportion of each worker type that will arrive is simply equal to population proportions. Thus, for city j, $E(z_j) = z$. A developer anticipating z will set city size in order to maximize the income of high ability individuals.[41] For this to attract both types of workers, the contract must be such that $I_{hm} > \hat{I}_h$ and $I_{\ell m} > I_\ell$. The problem:

$$\max_{n_m} D[za_h + (1-z)a_\ell]^{\gamma} n_m^{\delta+1} a_h - b n_m^{\frac{3}{2}} - z I_{hm} n_m - (1-z) I_{\ell m} n_\ell \tag{32}$$

Solving:

$$\begin{aligned} n_m &= [2b^{-1}\delta D(za_h + (1-z)a_\ell)^{\gamma} a_h]^{\frac{2}{1-2\delta}} \\ I_{hm} &= (1-2\delta) B[za_h + (1-z)a_\ell]^{\frac{\gamma}{1-2\delta}} a_h^{\frac{1}{1-2\delta}} \\ I_{\ell m} &= B[za_h + (1-z)a_\ell]^{\frac{\gamma}{1-2\delta}} a_h^{\frac{2\delta}{1-2\delta}} a_\ell - 2\delta B[za_h + (1-z)a_\ell]^{\frac{\gamma}{1-2\delta}} a_h^{\frac{1}{1-2\delta}} \end{aligned} \tag{33}$$

For low ability workers to prefer this allocation and be attracted to a mixed allocation requires $\hat{I}_{\ell m} > I_\ell$ or:

$$z > \left[\frac{(1-2\delta) a_\ell^{\frac{\gamma+1}{1-2\delta}}}{a_h^{\frac{2\delta}{1-2\delta}} a_\ell - 2\delta a_h^{\frac{1}{1-2\delta}}} \right]^{\frac{1-2\delta}{\gamma}} (a_h - a_\ell)^{-1} - a_\ell \tag{34}$$

If the proportion of low ability workers in the population is small, as $z \to 1$, $I_{hm} \to I_h$. Thus, there exists values of z and other parameters such that $I_{hm} > \hat{I}_h$ (given $\hat{I}_h < I_h$).[42].

[41] An allocation that has developers maximizing total worker incomes in a city ($zD[za_h + (1-z)a_\ell]^{\gamma} n_m^{\delta+1} a_h + (1-z)D[za_h + (1-z)a_\ell]^{\gamma} n_m^{\delta+1} a_\ell - b n_m^{\frac{3}{2}}$) will definitely not be an equilibrium allocation. Given this, there will always be opportunity for a developer to offer slightly larger city sizes than the n_m that would solve this problem and attract high ability people away from the city. Only a city size chosen to maximize high ability workers' incomes will potentially be an equilibrium.

[42] Note that a necessary condition for the non-existence of a natural separation equilibrium is that $\frac{a_h}{a_\ell} < \frac{1}{2\delta}$. When this is the case, $a_h^{\frac{2\delta}{1-2\delta}} a_\ell < 2\delta a_h^{\frac{1}{1-2\delta}}$, preserving the inequality sign when solving for the condition in (34) for which values of z satisfy $I_{\ell m} > I_\ell$.

7.2 Proof of Proposition 4

To demonstrate the non-existence of a mixed-city equilibrium, it is necessary to show that there exists a contract that developers can offer earning non-negative profit that can attract workers away from any such mixed allocation. Specifically, in this case, given any allocation of mixed cities, developers will be available to offer a profit-earning contract by offering a subsidy T^* that is less than the average rent level. This lesser subsidy will deter entry by low ability workers by providing compensation of less than $I_{\ell m}$ and attract high ability workers, providing them with expected compensation of greater than I_{hm}. From footnote 41, we see that n_m from above is the only city size which will potentially exist in a mixed equilibrium allocation. However, this will not be an equilibrium when there exists a T^* which satisfies the following conditions:

$$Dn_m^{\delta}a_h^{\gamma+1} - \frac{3}{2}bn_m^{\frac{1}{2}} + T^* \geq I_{hm} \tag{35}$$

$$Dn_m^{\delta}a_h^{\gamma}a_\ell - \frac{3}{2}bn_m^{\frac{1}{2}} + T^* \leq I_{\ell m} \tag{36}$$

Solving (36) with equality for T^*, obtaining the largest possible subsidy level that can deter entry by low ability workers, and substituting into (35) along with values for $I_{hm} - I_\ell m$ provides:

$$Dn_m^{\delta}a_h^{\gamma}(a_h - a_\ell) \geq Dn_m^{\delta}[za_h + (1-z)a_\ell]^{\gamma}(a_h - a_\ell) \tag{37}$$

This is always true.

7.3 Proof of Proposition 5

Consider an allocation of cities each of size $\tilde{n}$ and containing $z\tilde{n} = \tilde{n}_h$ high ability and $(1-z)\tilde{n} = \tilde{n}_\ell$ low ability workers.[43] Assume that all land rents are distributed equally to the residents of the city. This can be implemented by assuming that all city residents are Arrow-Debreu shareholders in a local land development corporation that owns the land. This assumption is not critical for any results, it simply makes the solutions more comparable to those with land developers. Writing real incomes for high and low ability workers as functions of the numbers of each type in the city,:

$$\begin{aligned}\bar{I}_h &= D[\bar{n}_h a_h + \bar{n}_\ell a_\ell]^{\gamma}(\bar{n}_h + \bar{n}_\ell)^{\delta-\gamma}a_h - b(\bar{n}_h + \bar{n}_\ell)^{\frac{1}{2}}\\ \bar{I}_\ell &= D[\bar{n}_h a_h + \bar{n}_\ell a_\ell]^{\gamma}(\bar{n}_h + \bar{n}_\ell)^{\delta-\gamma}a_\ell - b(\bar{n}_h + \bar{n}_\ell)^{\frac{1}{2}}\end{aligned}$$

The first requirement for an equilibrium is that city sizes are such that worker do not prefer to merely consume their endowments. Normalizing this value to 0, this requirement puts a limit on maximum city sizes for high and low ability individuals, denoted n_{max}^h and n_{max}^ℓ respectively:

$$n_{max}^h = \left[2b^{-1}D[za_h + (1-z)a_\ell]^{\gamma}a_h\right]^{\frac{2}{1-2\delta}} \tag{38}$$

$$n_{max}^\ell = \left[2b^{-1}D[za_h + (1-z)a_\ell]^{\gamma}a_\ell\right]^{\frac{2}{1-2\delta}} \tag{39}$$

Since $n_{max}^\ell < n_{max}^h$, city sizes are limited to populations of $n_{max} = n_{max}^\ell = \left[2b^{-1}D[za_h + (1-z)a_\ell]^{\gamma}a_\ell\right]^{\frac{2}{1-2\delta}}$.

Secondly, an allocation is a Nash equilibrium as long as $\left.\frac{\partial I_h}{\partial n_h}\right|_{\bar{n}_\ell} < 0$ and $\left.\frac{\partial I_\ell}{\partial n_\ell}\right|_{\bar{n}_h} < 0$. This provides a minimum bound on city sizes, requiring that $\bar{n} > n_{min}$ where:

$$n_{min} = \left[2b^{-1}\gamma D[za_h + (1-z)a_\ell]^{\gamma-1}a_h^2 + 2(\delta-\gamma)b^{-1}D[za_h + (1-z)a_\ell]^{\gamma}\right]^{\frac{2}{1-2\delta}} \tag{40}$$

[43]Note that any equilibrium will require all cities of a given type to be identically sized and proportioned in order for compensation rates to both types of individuals to be identical.

Thus, any city size $n \in [n_{min}, n_{max}]$ is a Nash equilibrium (if $n_{max} < n_{min}$, no such equilibrium will exist). To examine local stability of this type of equilibrium, take a first order Taylor series expansion around $\bar{n}_h$ and $\bar{n}_\ell$ providing:

$$\begin{bmatrix} \dot{n}_h \\ \dot{n}_\ell \end{bmatrix} = \begin{bmatrix} \frac{\partial I_h(\bar{n}_h,\bar{n}_\ell)}{\partial n_h} & \frac{\partial I_h(\bar{n}_h,\bar{n}_\ell)}{\partial n_\ell} \\ \frac{\partial I_\ell(\bar{n}_h,\bar{n}_\ell)}{\partial n_h} & \frac{\partial I_h(\bar{n}_h,\bar{n}_\ell)}{\partial n_\ell} \end{bmatrix} \begin{bmatrix} n_h - \bar{n}_h \\ n_\ell - \bar{n}_\ell \end{bmatrix}$$

A sufficient condition for this dynamic system to be a saddle point is that the determinant of the Jacobian matrix is negative, which after some simplification can be found to be equal to $-\frac{1}{2}b\bar{n}^{-\frac{1}{2}}\gamma D[za_h + (1-z)a_\ell]^{\gamma-1}\bar{n}^{\delta-1}(a_h - a_\ell)^2 < 0$, where $\bar{n} = \bar{n}_h + \bar{n}_\ell$. Thus, except for perturbations along a knife-edge path, this equilibrium is locally unstable.

7.4 Local Stability of Segregated Self-Organization Equilibria

Incomes in segregated cities of size $\bar{n}_h$ and $\bar{n}_\ell$ for high and low ability workers are given by:

$$I_h = D\bar{n}_h^\delta a_h^{\gamma+1} - b\bar{n}_h^{\frac{1}{2}} \quad (41)$$

$$I_\ell = D\bar{n}_\ell^\delta a_\ell^{\gamma+1} - b\bar{n}_\ell^{\frac{1}{2}} \quad (42)$$

To determine local stability, take a Taylor-series expansion as follows, evaluating it at $\bar{n}_\ell = 0$ in high ability cities and $\bar{n}_h = 0$ in low ability cities::

$$\begin{bmatrix} \dot{n}_h \\ \dot{n}_\ell \end{bmatrix} = \begin{bmatrix} \left.\frac{\partial I_h(\bar{n}_h,\bar{n}_\ell)}{\partial n_h}\right|_{\bar{n}_\ell=0} & \left.\frac{\partial I_h(\bar{n}_h,\bar{n}_\ell)}{\partial n_\ell}\right|_{\bar{n}_\ell=0} \\ \left.\frac{\partial I_\ell(\bar{n}_h,\bar{n}_\ell)}{\partial n_h}\right|_{\bar{n}_h=0} & \left.\frac{\partial I_h(\bar{n}_h,\bar{n}_\ell)}{\partial n_\ell}\right|_{\bar{n}_h=0} \end{bmatrix} \begin{bmatrix} n_h - \bar{n}_h \\ n_\ell - \bar{n}_\ell \end{bmatrix}$$

For any Nash equilibrium allocation, the trace of the Jacobian matrix is negative. Evaluating the determinant, by simplifying it can be shown that the determinant is positive when the following holds:

$$\frac{1}{2}b\bar{n}_h^{\frac{1}{2}-\delta}a_\ell^\delta a_\ell^\gamma(a_h - a_\ell) + \frac{1}{2}b\bar{n}_\ell^{\frac{1}{2}-\delta}a_h^\gamma a_\ell^\gamma(a_h - a_\ell) > Da_h^\gamma a_\ell^\gamma[(\gamma a_\ell + (\delta-\gamma)a_h)(\gamma a_h + (\delta-\gamma)a_\ell) - \delta^2 a_h a_\ell] \quad (43)$$

By substituting in the minimum possible Nash equilibrium values for $\bar{n}_h$ and $\bar{n}_\ell$ this reduces to:

$$a_h + a_\ell > \gamma(\delta-\gamma)(a_h - a_\ell) \quad (44)$$

which always holds.

7.5 Self-Selection Constraints in Extended Model

For low ability cities, with low ability workers engaged in both occupations, equilibrium requires $I_q = I_\ell$, implying:

$$\epsilon D n_\ell^{\epsilon-1} q_\ell^\alpha a_\ell^{\gamma+1} = \alpha D n_\ell^\epsilon q_\ell^{\alpha-1} a_\ell^{\gamma+1} \quad (45)$$

or:

$$v_\ell = \frac{\epsilon}{\alpha} \quad (46)$$

Low ability individuals working as intermediate input producers in these high ability cities must earn the same income as workers in low ability cities. Thus, $I_q = I_\ell$ or:

$$\alpha D n_h^\epsilon q_h^{\alpha-1} a_h^{\gamma+1} - \frac{3}{2}b(n_h + q_h)^{\frac{1}{2}} = I_\ell \quad (47)$$

substituting in from the zero-profit condition $\delta D n_h^{\epsilon} q_h^{\alpha} a_h^{\gamma+1} = \frac{1}{2} b(n_h + q_h)^{\frac{3}{2}}$ provides:

$$\frac{B v_h^{\frac{\epsilon}{1-2\delta}} a_h^{\frac{\gamma+1}{1-2\delta}}}{(v_h+1)^{\frac{1+\delta}{1-2\delta}}}[\alpha(v_h+1) - 3\delta] = I_\ell \tag{48}$$

In the "high ability" cities, high ability final good producers must not have an incentive to work as intermediate input producers, requiring that $I_h \geq I_q$, or:

$$\epsilon D n_h^{\epsilon-1} q_h^{\alpha} a_h^{\gamma+1} \geq \alpha D n_h^{\epsilon} q_h^{\epsilon-1} a_h^{\gamma+1} \tag{49}$$

implying $\frac{n_h}{q_h} \leq \frac{\epsilon}{\alpha}$.

Finally, low ability individuals must not have an incentive to work as final goods producers workers in high ability cities. A low ability worker thinking of operating a firm in a high ability city will take the price of the intermediate good as given and face the following maximization problem:

$$\max_{\tilde{u}} \pi \quad = \quad D n_h^{\delta} \tilde{u}^{\alpha} a_h^{\gamma} a_\ell - P\tilde{u} \tag{50}$$

We know that $P_h = \alpha D n_h^{\epsilon} q_h^{\alpha-1} a_h^{\gamma+1}$. Substituting this into the first order condition for (50) provides $\alpha D n_h^{\delta} \tilde{u}^{\alpha-1} a_h^{\gamma} a_\ell = \alpha D n_h^{\epsilon} q_h^{\alpha-1} a_h^{\gamma+1}$. This implies that $\tilde{u}$, the amount of intermediate input employed in production of the final good by a low ability final good producer worker in an otherwise all high ability final good producer city is:

$$\tilde{u} = \frac{q_h}{n_h}\left[\frac{a_\ell}{a_h}\right]^{\frac{1}{1-\alpha}} \tag{51}$$

Substitutions provide an expression for $I_{\ell h} = D n_h^{\delta} \tilde{u}^{\alpha} a_h^{\gamma} a_\ell - P\tilde{u} + T_h - \frac{3}{2} b(n_h + q_h)^{\frac{1}{2}}$:

$$I_{\ell h} = (1-\alpha) D n_h^{\epsilon-1} a_h^{\frac{\gamma(1-\alpha)-\alpha}{1-\alpha}} a_\ell^{\frac{1}{1-\alpha}} q_h^{\alpha} + \delta D n_h^{\epsilon-1} a_h^{\gamma+1} q_h^{\alpha} - \frac{3}{2} b(n_h + q_h)^{\frac{1}{2}} \tag{52}$$

Equilibrium requires that $I_{qh} \geq I_{\ell h}$, or:

$$\frac{n_h}{q_h} \geq \frac{(1-\alpha) a_h^{\frac{\gamma(1-\alpha)-\alpha}{1-\alpha}} a_\ell^{\frac{1}{1-\alpha}} + \delta a_h^{\gamma+1}}{\alpha a_h^{\gamma+1}} < \frac{\epsilon}{\alpha} \tag{53}$$

References

Abdel-Rahman, H. M. (1996), "Income Disparity, Time Allocation and Social Welfare". mimeo, University of New Orelans.

Abdel-Rahman, H. M. (1997), "Training Cost and Skill Distribution in a System of Cities". mimeo, University of New Orelans.

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

Benabou, R. (1993), "Workings of a City: Location, Education, and Production", *Quarterly Journal of Economics* **108**(3), 619–652.

Benabou, R. (1996*a*), "Equity and Efficiency in Human Capital Investment: The Local Connection", *Review of Economic Studies* **63**(2), 237–264.

Benabou, R. (1996*b*), "Heterogeneity, Stratification, and Growth: Macroeconomic Implcations of Community Structure and School Finance", *American Economic Review* **86**(3), 584–609.

Berglas, E. (1976), "Distribution of Tastes and Skills and the Provision of Local Public Goods", *Journal of Public Economics* **6**, 409–423.

Black, D. and Henderson, V. (1998), "Urban Evolution in the USA". Brown University Working Paper No. 98-21.

Black, D. and Henderson, V. (1999), "A Theory of Urban Growth", *Journal of Political Economy* **107**(2), 252–284.

Borjas, G. J., Bronars, S. G. and Trejo, S. J. (1992), "Self-Selection and Internal Migration in the United States", *Journal of Urban Economics* **32**, 159–185.

Brueckner, J. K. (1994), "Tastes, Skills, and Local Public Goods", *Journal of Urban Economics* **35**, 201–220.

Ciccone, A., Peri, G. and Almond, D. (1999), "Identifying Human Capital Externalities: Theory and Evidence from U.S. Cities". mimeo.

Cutler, D. M. and Glaeser, E. L. (1997), "Are Ghettos Good or Bad?", *Quarterly Journal of Economics* **112**(3).

de Bartolome, C. A. (1990), "Equilibrium and Inefficiency in a Community Model with Peer Group Effects", *Journal of Political Economy* **98**(1).

DiNardo, J. E. and Pischke, J.-S. (1997), "The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?", *Quarterly Journal of Economics* **112**, 291–304.

Duncan, O. and Duncan, B. (1955), "A Methodological Analysis of Segregation Indexes", *American Sociological Review* **20**, 210–217.

Durlauf, S. (1996), "A Theory of Persistent Income Inequality", *Journal of Economic Growth* **1**, 75–93.

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

Foster, A. D. and Rosenzweig, M. R. (1995), "Learning by Doing and Learning from Others: Human capital and Technical Change in Agricultre", *Journal of Political Economy* .

Fujita, M. and Ogawa, M. (1982), "Multiple Equilibria and Structural Transition of Non-Monocentric Configurations", *Regional Science and Urban Economics* **12**, 161–196.

Garreau, J. (1991), *Edge City: Life on the New Frontier*, Doubleday.

Gaspar, J. and Glaeser, E. L. (1996), "Information Technology and the Future of Cities". NBER Working Paper 5562.

Glaeser, E. (1998), "Should Transfer Payments be Indexed to Local Price Levels?", *Regional Science and Urban Economics* **28**(1).

Gyourko, J. and Tracy, J. (1991), "The Structure of Local Public Finance and the Quality of Life", *Journal of Political Economy* **99**(4).

Gyuorko, J. and Tracy, J. (1989), "The Importance of Local Fiscal Conditions in Analyzing Local Labor Markets", *Journal of Political Economy* **97**.

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

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

Henderson, J. V. (1985), "The Tiebout Model: Bring Back the Entrepreneurs", *Journal of Political Economy* **93**(2).

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

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

Henderson, V. and Mitra, A. (1997), "The New Urban Landscape: Developers and Edge Cities", *Regional Science and Urban Econonomics* .

Iaia, D. (1997), "Historical Patterns of National Segregation". mimeo, Brown University.

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics* **108**, 577–598.

Juhn, C., Murphy, K. M. and Pierce, B. (1993), "Wage Inequality and the Rise in Returns to Skill", *Journal of Political Economy* **101**(3).

Katz, L. and Murphy, K. (1992), "Changes in Relative Wages, 1963-1987: Supply and Demand Factors", *Quarterly Journal of Economics* .

Kremer, M. and Maskin, E. (1996), "Wage Inequality and Segregatin by Skill". NBER Working Paper 5718.

Krugman, P. (1992), "A Dynamic Spatial Model". NBER Working Paper 4219.

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

Mills, E. S. (1972), *Studies in the Structure of the Urban Economy*, Johns Hopkins Press.

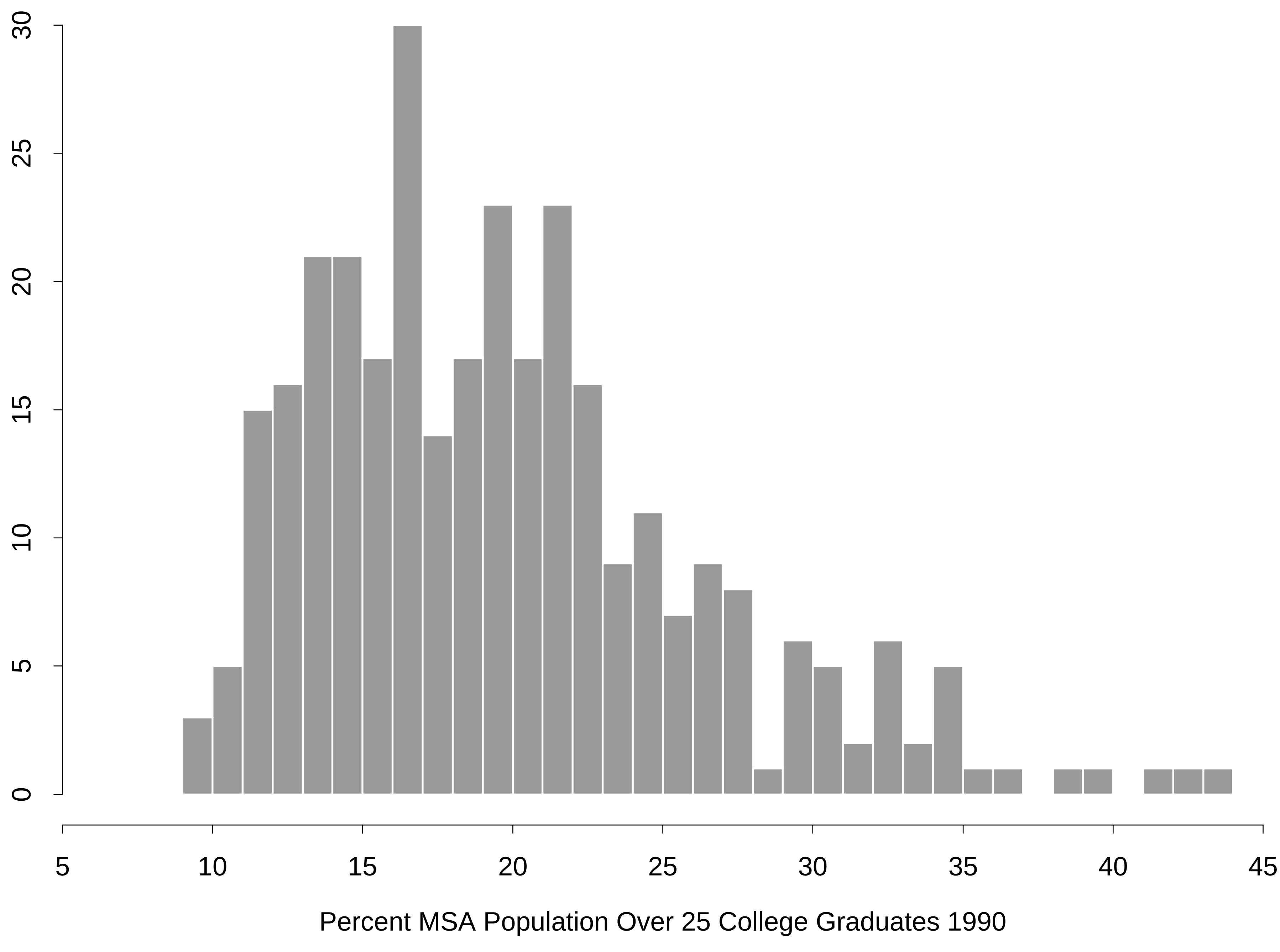
Rauch, J. E. (1993*a*), "Does History Matter Only When It Matters Little? The Case of City-Industry Location", *Quarterly Journal of Economics* **108**(3), 843–867.

Rauch, J. E. (1993*b*), "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities", *Journal of Urban Economics* **34**, 380–400.

Rothschild, M. and Stiglitz, J. (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *Quarterly Journal of Economics* pp. 629–649.

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

Tiebout, C. M. (1956), "A Pure Theory of Local Public Expenditures", *Journal of Political Economy* **64**, 416–424.


30
25
20
15
10
5
0
5
10
15
20
25
30
35
40
45
Percent MSA Population Over 25 College Graduates 1990