Aspirations, Segregation and Occupational Choice

Dilip Mookherjee
Boston University

Stefan Napel
University of Bayreuth
and

Debraj Ray
New York University

November 2008

Abstract

This paper examines steady states of an overlapping generations economy with a given distribution of household locations over a one-dimensional interval. Parents decide whether or not to educate their children. Educational decisions are affected by location: there are local complementarities in investment incentives stemming from aspirations formation, learning spillovers or local public goods. At the same time, economy-wide wages endogenously adjust to bring factor supplies into line with demand. The model therefore combines local social interaction with global market interaction. The paper studies steady-state configurations of skill acquisition, both with and without segregation, and the model is used to compare macroeconomic and welfare properties of segregated and unsegregated steady states.

We are grateful to two anonymous referees for their constructive comments. Mookherjee and Ray acknowledge funding from the National Science Foundation under grant nos. SES-0617874 and SES-0241070. Napel thanks the Institute for Economic Development, Boston University for hosting his visit for half a year, during which parts of this research project were carried out.

1. Introduction

This paper studies patterns of spatial segregation in a model of human capital accumulation. Accumulation incentives are subject to local complementarities, arising either from peer effects in aspiration formation, learning spillovers or the provision of local public goods. Neighborhoods with a greater proportion of skilled individuals are not just richer because of higher earnings, they also motivate parents to invest more in the human capital of their children. Having more accomplished neighbors raises the standards to which parents aspire for their children, enables their children to learn more from neighbors, and allows individuals to take advantage of better schools and facilities in the neighborhood (either because of parental involvement in school management or contributions to school resources). In this manner the pattern of segregation is reinforced over time.

Existing models of occupational choice and inequality persistence (going back to Loury 1981; Ray 1990; Banerjee and Newman 1993; Galor and Zeira 1993) mostly rely on capital market imperfections: unskilled parents are poorer and face higher costs of financing their children's education. Such models abstract from geographical aspects of inequality, which are a cause of concern for policy-makers and citizens in developed and developing countries alike. These issues include inner-city decay, slums coexisting side-by-side with affluent gated communities, urban versus rural areas, and inequality across regions. Moreover, as emphasized by certain authors such as Akerlof (1997), notions of distance can be applied to social dimensions as well as physical geography: social networks based on ethnicity, race or caste can play a role analogous to those of one's physical neighbors.[1]

One way of thinking about geographical inequality is to see it as a consequence of capital market imperfections. Such imperfections can result in persistence of inequality across occupations, which in turn could give rise to geographic inequality owing to the tendency for unskilled and skilled agents to cluster geographically (e.g., owing to variations in housing prices, local congestion or simply the preference of skilled agents to locate near other skilled people). Under this view, observed geographical patterns are the result of economic inequality rooted in capital market imperfections, combined with the spatial mobility of agents.

In contrast, this paper examines whether the reverse is possible, i.e., whether geography can itself be a primary determinant of human capital incentives and economic inequality, in the absence of capital market imperfections and agent mobility. Under this alternative view, different patterns of segregation could be associated with different macroeconomic outcomes, because variations in these patterns directly influence the decision to invest. In such a model, it is natural to investigate the association between persistent geographic inequality (or segregation, for short) and macroeconomic and welfare outcomes. This is why we deliberately abstract from capital-market imperfections or the spatial mobility of economic agents. This distinguishes our analysis from a number of previous models of

[1]Our model can be interpreted along the lines of either physical or social geography, provided locations can vary continuously over a single dimension (such as a color spectrum or caste which admits fine grades of distinction).

segregation that results from mobility, e.g., Schelling (1978), Bénabou (1993) or more recently Pancs and Vriend (2007).[2]

In the end, the extent to which agents are able to move (the cost of mobility) is an empirical question on which we do not comment. We simply study the implications of immobility, under the presumption that this assumption is just as plausible as perfect mobility (depending on the situation), and in many cases, more plausible. For instance, less developed countries exhibit substantially less spatial mobility than their developed counterparts. Bardhan (2002) argues that the assumption of zero mobility costs underlying the Tiebout model of local public goods is empirically implausible for most developing countries.[3]

Our model, then, has the following features. There is a large number of families in which parents make a binary skill investment in their children. Wages for skilled and unskilled labor settle to equate supply and demand for these factors on an economy-wide labor market. This is a standard model of occupational choice, and it can easily incorporate capital market imperfections. In order to focus on geographical effects *per se*, for the most part we consider a version where parental utility for consumption is linear, so rich and poor parents at the same location have the same incentives.[4]

In this standard formulation we introduce sources of local complementarity of parental investment incentives. One possible source could be local peer effects in parental aspirations for their children. Such aspirations can be thought of as targets or benchmarks for the economic status of their children. The higher the benchmark, the greater the parental motivation to invest in the child (controlling for wage differentials). Aspirations, in turn, are based on the current or past achievements of one's neighbors, located within some given spatial or social window. Alternative sources of local complementarity could be local effects on the (personal) cost of educational investments, such as access to quality schools or opportunities to learn from one's neighbors.

Our formulation of social preferences is related to but is not the same as models of identity (Akerlof and Kranton 2000; 2005) or conformity (Bisin et al. 2006), where well-being depends negatively on the distance between one's own *actions* and those of one's neighbors. Not only are preferences in those models defined directly over actions, but the main concern is for conformity. In contrast, agents in the aspirations-based version of our model are happiest

[2]In any case, if mobility costs were really zero, geographic segregation or housing price differences (or local congestion or school quality) would not be associated with welfare inequality among agents with the same wealth. Differences in housing costs or congestion would fully reflect differences in school quality in equilibrium. Any utility differences across neighborhoods would generate mobility and get "arbitraged away". In Bénabou's model, this is precisely the nature of the equilibrium. He assumes the stock of housing in each neighborhood to be given, so different patterns of segregation have implications for rents earned by passive landlords, and inequality between landlords and tenants.

[3]On the other hand, spatial mobility may be replaced by other forms of clustering which can have significant implications for human capital accumulation and wealth inequality across families. An excellent example of this is marital sorting (see Fernández and Rogerson 2001 and Fernández 2003).

[4]Our dynamic formulation with overlapping generations is motivated by our interest in models of human capital in which parents make investment decisions on behalf of their children. An analogous static model (where a given generation of agents make their own investment decisions) would deliver similar macroeconomic outcomes but substantially distinct welfare properties, in the absence of the intra-family parent-child externality.

if they can differentiate themselves favorably from their peers. Nevertheless, given the complementarity property it induces, the results of our model concerning existence and characterization would be similar if we were to use a conformity-based formulation instead.

In a steady state, the investment decisions of all agents *and* their aspirations are endogenously determined, and so are market wages. Many features of our model will be driven by the interaction of the market (economy-wide) externality with the social (local) externality. This is another feature that differentiates our model from most of the existing literature on segregation.[5]

We focus attention on two polar classes of steady state equilibrium allocations with stationary investment decisions of all households. One class involves complete lack of segregation: every location contains an identical mix of skilled and unskilled households. Such equilibria generate a pattern of aspirations that exhibit no geographical variation. The other class of equilibria are associated with alternating zones of skilled and unskilled neighborhoods (and corresponding geographically varying earnings and aspirations), each of which is wide enough that there exist households whose neighborhoods consist of only a single occupation. We refer to these as purely unsegregated and purely segregated equilibria.

Both classes of equilibria are shown to exist in general. Detailed restrictions on equilibrium patterns of segregation are derived: e.g., with a unimodal density of households, there can be at most three distinct clusters. More generally, with multiple local modes (identified with "cities"), either skilled clusters tend to concentrate in cities ("city-skilled" equilibria) or the reverse is true ("city-unskilled" equilibria). If both co-exist, city-skilled equilibria tend to exhibit a higher fraction of skill in the economy as a whole (i.e., are associated with higher per capita income and lower inequality in earnings between occupations).

The key results pertain to macroeconomic comparisons of purely segregated and unsegregated equilibria. If the technology is sufficiently skill-biased (and the equilibrium fraction of the economy that is skilled exceeds one-half), unsegregated equilibria are associated with a higher skill ratio in the economy as a whole. The converse is true if the economy is insufficiently skill-biased that the majority of the population is unskilled.

The preceding results are implications only of the assumption that there are local complementarities in investment, and that the size of local neighborhoods is small relative to the size of the economy. Welfare comparisons between segregated and unsegregated equilibria depend additionally on assumptions concerning how local aspirations affect the *level* of utility, over and above their effects on the marginal payoff to investment. Therefore different sources of local interactions will have varied effects: in some specifications, the higher achievement of one's peers may lower utility (as in the case of "competitive" aspirations); in others they

[5]After a first draft of this paper was completed, we became aware of independent work by Ghiglino and Goyal (2008), which studies local effects in the consumption of status goods within a static model. In particular, they emphasize the effect of exogenous social networks — specifically, the positioning of high-wealth, high-status consumers in that network — on equilibrium prices, allocations and welfare. Therefore, like us, Ghiglino and Goyal also combine local interactions with endogenous market-clearing prices. But there are many differences in questions, emphasis and modeling. For instance, segregation patterns are (to some extent) endogenous in this paper and so is wealth inequality. Moreover, Ghiglino and Goyal deal with a different form of social interdependence; namely, competitive consumption à la Veblen-Frank. In future work, it would be interesting to explore its effects in our setting (see footnote 23 below).

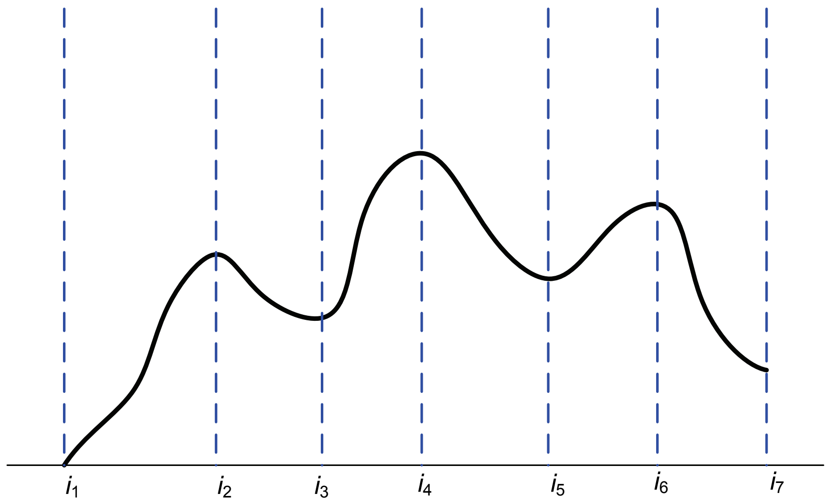

FIGURE 1. A DISTRIBUTION OF HOUSEHOLDS

may raise utility (as in the case of complementarities driven by local public goods). This variation complicates welfare comparisons between segregated and unsegregated equilibria. Additional ambiguities arise owing to the trade-off between higher inter-neighborhood inequality and lower intra-neighborhood inequality in segregated equilibria.

Section 2 introduces a model of occupational choice, one that allows both for the existence of capital-market imperfections and for socially-determined parental aspirations. Definitions of segregated and unsegregated steady state equilibria (or equilibria for short) are provided. Section 3 shows that segregated and unsegregated equilibria both exist in general. Section 4 specializes to the case of linearity of parental utility in own consumption, where the capital market imperfection is effectively absent, and derives specific spatial properties of segregated and unsegregated equilibria in this setting. Section 5 then compares macroeconomic properties of the two main classes of equilibria, while Section 6 compares their welfare effects. Finally, Section 7 summarizes the main results and discusses issues that need to be addressed in future work.

2. THE MODEL

2.1. **Locations and Skills.** There is a continuum of households indexed $h \in \mathcal{H}$ of unit measure, with given locations on an interval $I = [\underline{\iota}, \bar{\iota}]$ of the real line.[6] The location of household h in this interval is denoted $i(h)$, a function mapping $\mathcal{H}$ into I. This induces a distribution of households over I, described by a continuous density function f on I. We assume that f is strictly positive everywhere in the interior of I, that it is nowhere flat, and that it has a finite number of turning points. These assumptions imply, in particular, that there are finitely many locations $i_1, \ldots, i_{K+1}$, with $\underline{\iota} = i_1 < \cdots < i_{K+1} = \bar{\iota}$ such that f is strictly monotone on each interval $[i_k, i_{k+1}]$ and alternately increasing and decreasing across

[6]For concreteness we shall be thinking of this as a bounded interval, though no particular result in the paper depends on this feature.

consecutive intervals I_k and I_{k+1}. We shall refer to these "pieces" of f as (increasing or decreasing) *stretches*.

For notational convenience, we define $f(i) = 0$ elsewhere on $\mathbb{R}$ (the extension need not be continuous at the edges of I).[7]

Figure 1 illustrates a typical distribution f, as well as its stretches.

There is a sequence of generations $t = 1, 2, \ldots$. Each household is represented by a single adult in each generation, who is either *skilled* or *unskilled*, and earns a corresponding wage on an economy-wide labor market. This agent then decides whether or not to invest in the education of its child, which will determine whether the latter will be skilled or not in the next generation. This investment decision is the key endogenous variable in the model. The indicator function $\mathbf{1}(h)$ denotes whether household h is skilled, and $\beta(i)$ denotes the fraction of households at location i that are skilled. The overall skill ratio in the economy as a whole is therefore given by

$$\lambda = \int_I \beta(i) f(i) di.$$

2.2. **Wages.** The going skill ratio λ determines skilled and unskilled wages, denoted by $\bar{w}(\lambda)$ and $\underline{w}(\lambda)$, in the society at large.[8] This represents the economy-wide market interaction among households, arising from imperfect substitutability between skilled and unskilled labor in the technology, and the assumption that the labor market is integrated throughout the economy. Specifically, we presume that these wages are the marginal products of a production technology, described by a continuously differentiable, constant-returns-to-scale, strictly quasiconcave, Inada production function T defined on skilled and unskilled labor:

$$\bar{w}(\lambda) = T_1(\lambda, 1-\lambda) \text{ and } \underline{w}(\lambda) = T_2(\lambda, 1-\lambda),$$

where subscript j, $j = 1, 2$, denotes the derivative with respect to the jth input.[9]

[7]Given this, there will be no need to treat endpoints differently in the notation below for left and right neighborhoods of any given location: e.g., the right endpoint will have a right neighborhood which is entirely uninhabited.

[8]It is simplest to think of a setting where each adult supplies one unit of labor inelastically, so that earnings are entirely defined by the wage rates and the level of skill of the agent. It is easy to extend the model to a context where labor supply is endogenously affected by wages. For instance, in the "no income effects" case that we consider for most of this paper, the investment decision of each household depends only on the location of the household, and is independent of its current skill or income (conditional on location). Endogenous labor supply then has no effect on investment decisions of each household, controlling for its location and the incomes of its neighbors (which affect its aspirations). Given the absence of income effects, labor supply is non-decreasing in the wage rate, implying that skilled agents always earn higher incomes than unskilled agents in equilibrium.

[9]We make the constant returns assumption in order to ensure there are only two occupations with positive earnings. In the presence of decreasing returns, entrepreneurship would represent a third distinct occupation, a complication we wish to avoid in this paper. We defer to future research the question of how to extend the theory to more than two occupations.

It follows that $\bar{w}(\lambda)$ and $-\underline{w}(\lambda)$ are continuous and strictly decreasing, that the end-point conditions $\lim_{\lambda\downarrow 0}\bar{w}(\lambda) = \infty$ and $\lim_{\lambda\downarrow 0}\underline{w}(\lambda) = 0$ are satisfied, and that there is some value $\bar{\lambda} \in (0,1)$ with $\bar{w}(\bar{\lambda}) = \underline{w}(\bar{\lambda})$.[10]

2.3. **Parental Motivations and Skill Choices.** In line with dynamic models of occupational choice, parent and child within any household in any given generation are linked by intergenerational altruism. We normalize so that to create an unskilled descendant costs nothing, while a skilled child costs $x > 0$. This is an exogenous price for education or training that is incurred by the parent. The utility of a parent is the sum of two components.

The first is a direct utility component which depends on current consumption. For a parent in household h with going wage w, whose investment decision is represented by the indicator function $\mathbf{1}'$, we write this as $u\left(w - \mathbf{1}'(h)x\right)$.

The second component reflects parental altruism (or "pride" in the child's future economic status), represented by an indirect utility v. This depends on the earnings that the child is expected to acquire as an adult. The central premise of this paper is that v depends also on the achievements of geographical neighbors, via peer effects, local learning spillovers or resource externalities. For concreteness, we presume that the function v is affected by parental "aspirations", which are in turn determined by the distribution of wages in the parent's local neighborhood.[11] In what follows, our restrictions on the indirect utility function will make clear the notion of aspirations we have in mind.

Letting the scalar variable a represent parental aspirations, the indirect utility component depends on descendant wages w' as well as aspirations: $v(w', a)$. Thus overall parental utility is given by

$$u\left(w - \mathbf{1}'(h)x\right) + v(w', a).$$

We make two key assumptions concerning the nature of local externalities affecting parental investment incentives. The first concerns the way that aspirations are formed. The relevant influences might be "private" — some function of *that* household's wage history, or they might be "social" — some average of wages of the household's neighbors. In this paper we focus on the latter. We assume that the aspiration $a = a(i)$ for every household at location i is an average of wages earned by all neighbors in an ϵ-neighborhood centered at i.[12] For any

[10]Our assumptions imply that the wage differential to the right of $\bar{\lambda}$ is negative, but behavior in this region is unimportant as long as the wage differential does not turn strictly positive again. For instance, if skilled individuals can do unskilled jobs, it might make sense to assume that the wage differential is exactly 0 to the right of $\bar{\lambda}$.

[11]We are also assuming that parental altruism is paternalistic, i.e., that the utility of their children is not the key concern; instead it is the income that they will achieve, in relation to parental aspirations. Parental aspirations are an inherently paternalistic phenomenon, so this is a natural way to formulate parental motivations. However, our formulation imposes the restriction that parents do not care directly about their grandchildren or subsequent descendants. We suspect that similar results will obtain in a model with a dynastic (non-paternalistic) bequest motive as well, where local interactions arise not due to parental aspirations but to learning spillovers or contributions of neighbors to local public goods. This remains to be checked in future research.

[12]Weighting wages by distance could remove the (formally inconsequential) discontinuity of perception at $i \pm \epsilon$. We leave the investigation of more sophisticated aspiration formation for future research.

interval $N = [j, j'] \subseteq \mathbb{R}$, define $F(N) \equiv \int_j^{j'} f(i)di$; then

$$a(i) = \frac{1}{F(N)} \int_N w(j) f(j) dj, \quad (1)$$

where N is the interval $[i - \epsilon, i + \epsilon]$,[13] and $w(j)$ is average wage at location j:

$$w(j) = \beta(j)\bar{w} + [1 - \beta(j)]\underline{w}.$$

The second restriction is on the way that parental incentives depend on aspirations, as represented by properties of the v function. We assume that v is continuous, increasing and unbounded in its first argument. The important restriction is complementarity: we assume that for any pair of wages $\underline{w} < \bar{w}$, the difference

$$v(\bar{w}, a) - v(\underline{w}, a)$$

is increasing in a. In short, higher neighborhood wages always increase the marginal incentive to invest in one's own child. A particular version of the v function is where it is a (strictly concave and increasing) function solely of the difference between the earnings of one's child and the parent's aspiration.

This is a strong assumption: one can imagine situations in which it is not satisfied. For instance, Ray (1998, Chapter 3; 2006) has argued that extremely high aspirations could be detrimental to investment, simply because it may be very difficult to catch up. That argument cannot be fully incorporated into the current model because we work only with two skill levels, so that a single educational investment does, indeed, permit full catchup. Nevertheless, the complementarity assumption is a strong one because it rules out "frustration". In future work it will be interesting to explore the consequences of dropping this assumption, in a context with several occupational choices.

On the direct utility function u, we impose standard assumptions: that it is strictly increasing and concave. For simpler exposition of households' decision problems, we suppose further that u is defined over both positive and negative consumptions. Indeed, we shall primarily focus on the linear case where $u(c) = c$, so there are no income effects *per se*. The motivation for this, as explained in the Introduction, is to abstract from sources of history-dependence based on capital market imperfections, which cause the utility costs of investment to depend on current earnings of the household. Nevertheless in the remainder of this and in the following Section, we shall continue to retain the assumption that u is a concave function, to describe properties of the model that do not depend on particular assumptions concerning capital market imperfections.

2.4. **Segregated and Unsegregated Steady State Allocations.** An allocation in this economy is a specification of an investment decision for every household in any given generation. We shall focus on allocations with two sets of properties. The first is a *steady state* property: investment decisions for each household are stationary, i.e., unchanging across generations. The second pertains to the spatial structure of investments, which are either purely unsegregated or purely segregated, as we explain next.

[13]Recall that we have extended f to all of $\mathbb{R}$ so there is no ambiguity in this definition at the edges of I; households near to or at the edge see few or no individuals to one side.

A steady state allocation specifies investments and skills for every household, which determines an occupational distribution for the economy as a whole, as represented by λ, the fraction of households that are skilled. In turn, this determines skilled and unskilled wages, and therefore the earnings of all households. In turn this determines the aspirations of households at different locations.

An allocation is (purely) *unsegregated* if aspirations do not change with location: $a(i)$ is a constant for all $i \in I$. An example of such an allocation is one in which skilled individuals are distributed uniformly across all locations — in every subinterval of I, no matter how small, the proportion of skilled individuals is the same.

An allocation is (purely) *segregated* if the distribution of skills $\mathbf{1}$ takes on values of 0 and 1 over successive intervals, with each interval at least of size 2ϵ. This last requirement is related to the qualification "purely". It insists that the extent of the segregation be at least as large so that each interval of (un)skilled agents contains at least one household whose aspirations are determined by peers with the same skill type only.[14] Thus in a purely segregated allocation the indicator function $\mathbf{1}$ carves up I into a succession of skilled and unskilled intervals, which are sizeable relative to individual cognitive windows.

Figure 2 illustrates a purely segregated allocation where the "cuts" $c_1, \ldots c_4$ divide up the society into successive segments of skilled and unskilled households.

There may, of course, be steady state allocations which are neither purely segregated nor purely unsegregated. For instance, there could be allocations where skills are segregated into successive intervals of unskilled and skilled, some of which have width smaller than 2ϵ. If all intervals are narrow in this sense, skilled and unskilled households coexist in the local neighborhood of *every* individual. In that sense the allocation exhibits lack of segregation. Yet aspirations need not be constant across locations, and all households at any given location are either skilled or unskilled — in that sense the allocation exhibits segregation. Clearly such allocations lie somewhere in-between the polar types of purely segregated and purely unsegregated allocations. In this paper we ignore such intermediate types of allocations, as they involve a number of delicate technical issues.

2.5. **Steady State Equilibrium Allocations.** We now describe equilibrium conditions on steady state allocations.

A *steady state equilibrium* (or just *equilibrium*, for short) is a stationary allocation, i.e., a stationary distribution $\mathbf{1}$ of skills on households, an aggregate skill ratio λ, and wages for skilled and unskilled labor ($\bar{w}$ and $\underline{w}$), with the following properties:

(i) wages are consistent with the aggregate skill ratio: $\bar{w} = \bar{w}(\lambda)$ and $\underline{w} = \underline{w}(\lambda)$;

(ii) the aggregate skill ratio is consistent with the distribution of skills: $\lambda = \int_I \beta(i) f(i) di$, where $\beta(i)$ is the average of $\mathbf{1}$ for all households located at i;

[14] Our definition of 'pure segregation' may be somewhat too suggestive: the measure of agents who indeed have only skilled or only unskilled neighbors may be small or even zero (when intervals have exactly size 2ϵ).

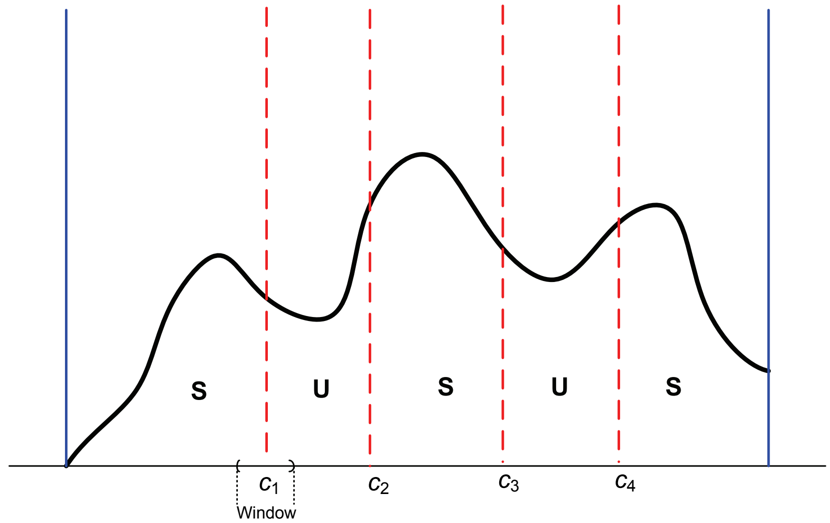

Figure 2. A Purely Segregated Equilibrium

(iii) skill choices are time-stationary and consistent with aspirations: for each h located at i, $\mathbf{1}'(h) = \mathbf{1}(h)$ solves the problem

$$\max_{\mathbf{1}'(h)} u\left(w^h - \mathbf{1}'(h)x\right) + v\left(\mathbf{1}'(h)\overline{w} + [1 - \mathbf{1}'(h)]\underline{w}, a^h\right),$$

where $w^h = \mathbf{1}(h)\overline{w} + [1 - \mathbf{1}(h)]\underline{w}$, and $a^h = a(i)$ solves equation (1).

Thus in a steady state equilibrium, each household is selecting an optimal investment decision, given wage rates and the distribution of earnings. And the latter are consistent with the occupational distribution generated by the aggregation of investment decisions of households. Moreover, the decisions of all households and all wages are stationary across generations.

The rest of the paper will be devoted to analysis of steady state equilibrium allocations that are either purely unsegregated or purely segregated, addressing both existence issues and characterization of their specific properties.

3. Existence of Unsegregated and Segregated Equilibria

In this section, we show that both segregated and unsegregated equilibria exist in general. Let's review the main steps that guarantee existence. First, for any skill ratio λ and any aspiration a, define

$$\Psi(\lambda, a) \equiv v\left(\overline{w}(\lambda), a\right) - v\left(\underline{w}(\lambda), a\right).$$

By the complementarity assumption, Ψ is increasing in a and by our assumptions on the wage functions, it is declining in λ.

In an unsegregated equilibrium with aggregate skill ratio λ, it must be the case that

$$a = \lambda\overline{w}(\lambda) + (1 - \lambda)\underline{w}(\lambda), \tag{2}$$

while the following condition ensures that skilled parents will choose to invest in their children's education, while unskilled parents do not, so that we do have a steady state:

$$u(\bar{w}(\lambda)) - u(\bar{w}(\lambda) - x) \leq \Psi(\lambda, a) \leq u\left(\underline{w}(\lambda)\right) - u\left(\underline{w}(\lambda) - x\right). \tag{3}$$

Indeed, conditions (2) and (3) are both necessary and sufficient for λ to be the outcome of an unsegregated equilibrium.

Proposition **1.** *An unsegregated equilibrium exists.*

Proof. To show that (2) and (3) must hold for some λ, define $a(\lambda)$ by the right hand side of (2), and then define $\Phi(\lambda) \equiv \Psi(\lambda, a(\lambda))$. It is easy to see that Φ is continuous for all $\lambda \in (0, 1)$.

By Euler's theorem, $a(\lambda) = T(\lambda, 1 - \lambda)$, so it is bounded in λ. Moreover, v is unbounded in w, so it follows from the end-point conditions on $\bar{w}(\lambda)$ that $\Phi(\lambda) \to \infty$ as $\lambda \to 0$. On the other hand, it is easy to see that $\Phi(\lambda) \to 0$ as $\lambda \to \bar{\lambda}$. At the same time, $u(\bar{w}(\lambda)) - u(\bar{w}(\lambda) - x)$ is bounded, strictly positive and continuous on $(0, \bar{\lambda}]$.

Putting all these observations together, we must conclude that there exists $\lambda \in (0, \bar{\lambda})$ such that

$$u(\bar{w}(\lambda)) - u(\bar{w}(\lambda) - x) = \Phi(\lambda). \tag{4}$$

Because u is concave and $\lambda \leq \bar{\lambda}$, we know that

$$u(\bar{w}(\lambda)) - u(\bar{w}(\lambda) - x) \leq u\left(\underline{w}(\lambda)\right) - u\left(\underline{w}(\lambda) - x\right). \tag{5}$$

It is now easy to see that (4) and (5) jointly imply that (2) and (3) are satisfied, and we are done. □

We remark in passing that the complementarity assumption on v is not needed for Proposition 1.

At the same time, the model also admits purely segregated equilibria. Such equilibria are described by a different set of conditions. We study, first, what might be called *single-cut* equilibria, in which I is divided into two intervals, with skilled individuals on (say) the right, and unskilled individuals on the left. Whilst the existence of a single-cut equilibrium, of course, implies existence of a purely segregated equilibrium in the general sense, we will later also provide sufficient conditions for equilibria with more than one cut to exist (Section 4.3).

To describe single-cut equilibria, define for any c in the interior of I, the closed intervals $R(c)$ and $L(c)$ to the right and left of c in the obvious way. See Figure 3 for an illustration. Now suppose that all individuals in the relative interior of $R(c)$ are skilled, and all those in the relative interior of $L(c)$ are unskilled. (That leaves open just the measure-0 issue of what households exactly at c do, something we'll settle later.) Define a function ρ on I by

$$\rho(c) \equiv \frac{F(N^+)}{F(N)}, \tag{6}$$

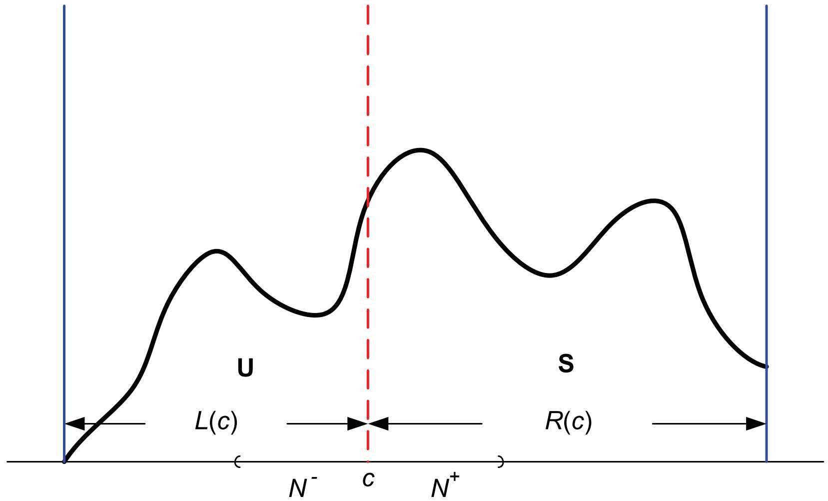

FIGURE 3. A SINGLE-CUT EQUILIBRIUM

where $N^+ \equiv [c, c+\epsilon]$ and $N = [c-\epsilon, c+\epsilon]$. Because $f(c) > 0$ in the interior of I, ρ is well-defined.

Given our spatial arrangement of skilled and unskilled labor, it is easy enough to see that $\rho(c)$ may be interpreted as the proportion of skilled individuals perceived by c, provided that the cut lies at c. Now individuals at location j just to the right of c see a skill proportion that converges to $\rho(c)$ as $j \to c$, so by a trivial continuity argument, a necessary condition for the cut at c to be a steady state is that

$$u\left(\overline{w}(\lambda(c))\right) - u\left(\overline{w}(\lambda(c)) - x\right) \leq \Psi\left(\lambda(c), a(c)\right),$$

where $\lambda(c) = F(R(c))$ is the aggregate proportion of skilled labor generated by the single cut at c, and

(7) $$a(c) = \rho(c)\overline{w}\left(\lambda(c)\right) + [1-\rho(c)]\underline{w}\left(\lambda(c)\right).$$

Similarly, using unskilled individuals just to the left of c, we must conclude that a second necessary condition for the cut at c to be a steady state is

$$u\left(\underline{w}(\lambda(c))\right) - u\left(\underline{w}(\lambda(c)) - x\right) \geq \Psi\left(\lambda(c), a(c)\right).$$

Combining these two inequalities, we obtain a necessary condition for the cut at c to generate a steady state:

(8) $$u\left(\overline{w}(\lambda(c))\right) - u\left(\overline{w}(\lambda(c)) - x\right) \leq \Psi\left(\lambda(c), a(c)\right) \leq u\left(\underline{w}(\lambda(c))\right) - u\left(\underline{w}(\lambda(c)) - x\right).$$

(Now assign households at c to be skilled or unskilled depending on which of these inequalities hold strictly. If both hold with equality, it doesn't matter.)

By the complementarity condition, (7) and (8) must be *sufficient* as well. The reason is that as we move away from c to the right (resp. left), the proportion of skilled people must rise (resp. fall), so that aspirations rise (resp. fall). If incentives are correct at the cut they must therefore be correct in the interior. We conclude that any cut that satisfies (7) must generate a segregated steady state.

Proposition **2.** *A (single-cut) segregated equilibrium exists.*

Proof. Define $\zeta(c) \equiv \Psi(\lambda(c), a(c))$. It is easy to see that ζ is continuous.

Let I be the interval $[\underline{\iota}, \bar{\iota}]$. Note that as $c \uparrow \bar{\iota}$, $\lambda(c) \downarrow 0$, so that $\bar{w}(\lambda(c)) \uparrow \infty$ and $\underline{w}(\lambda(c)) \downarrow 0$. By complementarity, we see that

$$\zeta(c) \equiv \Psi(\lambda(c), a(c)) \geq \Psi(\lambda(c), 0) \to \infty.$$

Now define i^* by the condition that $\lambda(i^*) = \bar{\lambda}$. Again, it is easy to see that as $c \downarrow i^*$, $\zeta(c) \to 0$. At the same time, $u(\bar{w}(\lambda(c))) - u(\bar{w}(\lambda(c)) - x)$ is bounded, continuous and positive on $[i^*, \bar{\iota})$.

We must therefore conclude that there exists $c \in (i^*, \bar{\iota})$ such that

$$u(\bar{w}(\lambda(c))) - u(\bar{w}(\lambda(c)) - x) = \zeta(c). \tag{9}$$

Because u is concave and $c \geq i^*$, we know that

$$u(\bar{w}(\lambda(c))) - u(\bar{w}(\lambda(c)) - x) \leq u\left(\underline{w}(\lambda(c))\right) - u\left(\underline{w}(\lambda(c)) - x\right). \tag{10}$$

It is now easy to see that (9) and (10) jointly imply that (7) and (8) are satisfied, and we are done.

$\square$

Steady state equilibria are driven by three features of the model. First, there is the nature of aspirations and how they affect the incentive to educate a child. This is summarized in the function v. Second, there is the general equilibrium effect: the fact that aggregate skill ratios affect wages. These two features lead to a particular theory of social interactions which are mediated by market prices. Finally, skilled dynasties have different steady state wealths and therefore different (utility) costs of educating their children. These varying costs are reflected in the steady-state equilibrium inequalities such as (3) and (8).

This last feature is generally a shorthand for imperfect or altogether missing capital markets (see Loury 1981, Ray 1990, and Mookherjee and Ray 2003), and it is well-known that the absence of such markets often leads to a multiplicity of steady states (Banerjee and Newman 1993, Galor and Zeira 1993). This model is no exception. For instance, whenever the inequality in (3) holds strictly for some λ, there is a continuum of unsegregated steady states, and whenever the inequality in (8) holds strictly for some $c \in I$, there is a continuum of single-cut segregated steady states. A slight modification of the proofs of Propositions 1 and 2 tells us that these strict inequalities will indeed hold (for some λ and some c), provided that the direct utility function u is *strictly* concave.

The possible existence of a continuum of steady states in our and many other models of occupational choice is a formal expression of extreme history-dependence. It is a consequence of the assumptions that capital markets are imperfect, and that the set of occupations is sparse. An interval of steady states creates scope even for small, temporary policy interventions to have a permanent beneficial effect on human capital and per capita income. However, such strong history dependence is not very robust. For example, as demonstrated by Mookherjee and Ray (2003), the multiplicity of equilibria disappears with a rich enough set of occupational choices. And when agents differ randomly in their educational talents, there (generically) can exist only a finite number of steady states which involve social mobility (see Mookherjee and Napel 2007 and Napel and Schneider 2008). We incorporate none of these possible extensions in the current model, for the simple reason that we wish to abstract from capital market imperfections and focus instead on the role of geographic location.[15] When the utility function is linear, we shall show in the subsequent sections, the set of purely segregated and unsegregated equilibria do not generally form a continuum.

4. Equilibrium Patterns for Linear Utility

4.1. **Unsegregated Equilibrium.** First consider purely unsegregated equilibria with linear direct utility. Invoking (3) for this special case, we see that an equilibrium skill ratio λ is fully characterized by the *equality*

$$\Psi(\lambda, a) = x,$$

where a, it will be remembered, equals $\lambda\bar{w}(\lambda) + (1-\lambda)\underline{w}(\lambda)$, which equals $T(\lambda, 1-\lambda)$ in turn. As we have already seen, steady states exist, but there may well be many of them, despite the linearity of utility. Suppose that λ_1 is one such steady state. Consider the thought experiment of increasing λ_1 to λ. The direct effect of this is to lower the value of Ψ (this is because the wage differential narrows, lowering the incentive for skill acquisition). At the same time, $T(\lambda, 1-\lambda)$ may well go up, raising a. By complementarity, this increases the incentive for skill acquisition. The net effect is ambiguous. If, indeed, Ψ goes up as a result, we can be sure that there will exist yet another steady state $\lambda_2 > \lambda > \lambda_1$.[16]

The potential multiplicity here is, however, different from what we observe with strictly concave u. Here, steady states are generically isolated and finite. In the strictly concave case, a continuum of steady states invariably exists.

4.2. **Segregated Equilibrium.** Now we turn to purely segregated equilibria. Any such equilibrium generates a finite collection of *cuts*, which we represent by the ordered set $C = \{c_1, \ldots, c_m\} \subset I$. Define $c_0 \equiv \underline{\iota}$ and $c_{m+1} \equiv \bar{\iota}$; then pure segregation implies that $c_{k+1} - c_k > 2\epsilon$ for all $k = 0, \ldots, m$. (Recall Figure 2.) Moreover, within the consecutive intervals of I generated by the cuts in C, there are alternating zones of skilled and unskilled labor.

One useful implication of pure segregation is that a person located at a cut $c \in C$ sees only skilled people on one side and only unskilled people on the other. Her perceived ratio of skilled individuals will therefore depend entirely on the value $\rho(c)$, where ρ is defined in (6). More specifically, her perceived ratio will equal $\rho(c)$ if the zone to the right of her is populated by the skilled, and it will equal $1 - \rho(c)$ if the zone to the left of her is populated

[15] Carneiro and Heckman (2002) suggest that capital market imperfections do not actually impose serious constraints for educational investments, at least in developed countries like the US. But see Heckman and Krueger (2003) for several dissenting views.

[16] Lemma 1, used in another context, contains a formal statement of this assertion.

by the skilled. To write this in a compact way, define a function $\chi(c)$ on C that takes the value 1 if a skilled interval lies to the right of c, and 0 otherwise. Then an individual at c must perceive the local skill ratio

$$\sigma(c) \equiv \chi(c)\rho(c) + [1 - \chi(c)][1 - \rho(c)],$$

so that if the *aggregate* skill ratio is λ, an individual positioned at the cut c must have the aspiration

$$a(c) = \sigma(c)\bar{w}(\lambda) + [1 - \sigma(c)]\underline{w}(\lambda)\,. \tag{11}$$

Now, the same argument that we used for a single-cut equilibrium shows that at *any* cut c in a purely segregated equilibrium with aggregate skill ratio λ,

$$u(\bar{w}(\lambda)) - u(\bar{w}(\lambda) - x) \leq \Psi(\lambda, a(c)) \leq u\left(\underline{w}(\lambda)\right) - u\left(\underline{w}(\lambda) - x\right).$$

Invoking the assumption that u is linear, this inequality reduces to the condition

$$\Psi(\lambda, a(c)) = x. \tag{12}$$

Finally, we pin down λ, which is simply the aggregate mass of all skilled intervals:

$$\lambda = \sum_{k=0}^{m} F([c_k, c_{k+1}])\chi(c_k). \tag{13}$$

By the complementarity assumption, conditions (11)–(13) completely characterize all purely segregated equilibria.

In order to avoid uninformative case distinctions, we place one further restriction on purely segregated equilibria: we ask that their cuts c have the property that none of them are located at a turning point of f, and both $c+\epsilon$ and $c-\epsilon$ lie on the same stretch of f as c does. It is easy to verify that this is a generic property of segregated equilibria provided we require it for all ϵ small enough.[17] We call such cuts and the corresponding equilibria *regular*. Note that our definitions of pure segregation and of regularity are compatible with the coexistence of many skilled and unskilled segments on a given stretch (ignoring incentives).

We now state our basic proposition for purely segregated regular equilibria under linear direct utility.

Proposition **3.** *In any purely segregated regular equilibrium, each stretch of f can contain at most one cut.*

Proof. Suppose, on the contrary, that there are two consecutive cuts, call them c and c', along some interval of I on which f is strictly monotone. We claim that $\sigma(c) \neq \sigma(c')$. To see this, note that $\chi(c) \neq \chi(c')$, simply because c and c' are consecutive. Therefore $|\sigma(c) - \sigma(c')| = |\rho(c) + \rho(c') - 1|$. However, because the equilibrium is regular, along the same

[17] Suppose that there is a sequence of ϵ-windows converging to zero and a corresponding sequence of purely segregated equilibria with a nonregular cut in each of them. Then, using the fact that each cut must contain an indifferent person, who must therefore collectively have the same aspirations as one another, it is possible to show that *all* cuts converge to local peaks or troughs as $\epsilon \downarrow 0$. This means, in turn, that aggregate λ can have only one of a finite possible set of values (use (13)), and it also means that limit aspirations for indifferent individuals have at most a finite number of values as well. Combining, we see that $\Psi(\lambda, a(c))$ converges to one of a finite number of possible values as ϵ vanishes, and therefore equation (12) will generically fail.

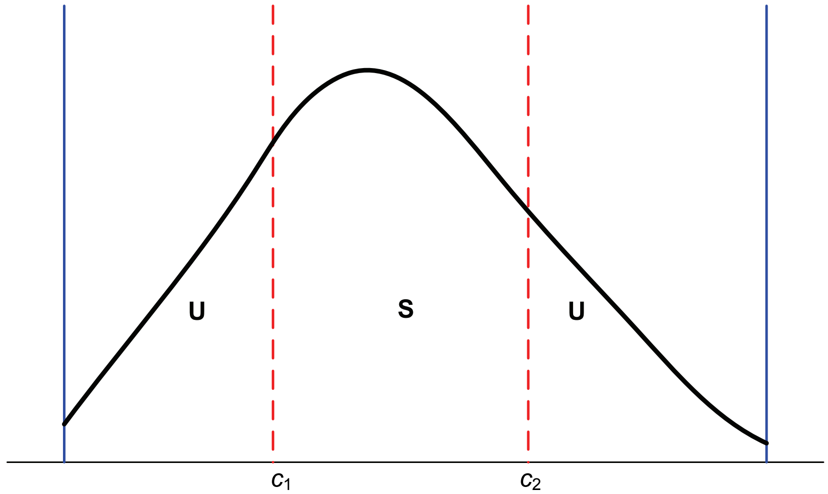

FIGURE 4. A TWO-CUT EQUILIBRIUM WITH SKILLED LABOR OCCUPYING THE MODE

stretch either $\rho(c)$ and $\rho(c')$ are both greater than $1/2$, or they are both less than $1/2$. This proves the claim.

Because $\bar{w}(\lambda) > \underline{w}(\lambda)$ in any equilibrium, we must conclude from (11) that $a(c) \neq a(c')$. Therefore, by the claim, at least one of c or c' must fail (12), which shows that both cannot be equilibrium cuts. This is a contradiction. $\square$

Proposition 3 has strong implications for special cases of the model. For instance, it severely restricts equilibrium outcomes when the distribution of population across locations is unimodal.

COROLLARY **1.** *If f is unimodal, then a purely segregated regular equilibrium can involve at most two cutoffs, and if there are two, they must be on either side of the mode.*

Proof. A unimodal f has no more than two stretches. Apply Proposition 3. $\square$

The restriction depends on the assumption that f is nonconstant almost everywhere. It is easy to see that Corollary 1 is false when f is uniform on I: in that case there are purely segregated equilibria with one, two, or several cuts (provided that ϵ is small enough), and no particular spatial pattern of segregation emerges. In the “strictly” unimodal case, however, we see that purely segregated equilibria must assume a very simple form. Moreover, in the two-cut case one of the two skill groups must occupy the highly populated “center zone” around the mode, while the other skill group occupies the low-density “periphery”. Figure 4 illustrates a two-cut equilibrium in the unimodal case.

A similar pattern obtains for multimodal distributions:

COROLLARY **2.** *If f is multimodal with n local modes, then a purely segregated regular equilibrium involves at most $2n$ cutoffs, and consecutive cutoffs must be located on stretches of f with slopes of opposite sign.*

Proof. An f with n local modes has no more than $2n$ stretches. Apply Proposition 3. $\square$

The fact that consecutive cutoffs must be located on stretches of f with slopes of opposite sign allows us to generalize the center-periphery pattern in the unimodal case. Suppose that a multi-cut equilibrium begins with a segment of skilled individuals, and exhibits its first cut on a downward stretch. Then there must be a local mode to the left of the cut (a "local center") occupied by S. But this one fact now *necessitates* that every succeeding unskilled segment must wrap around at least one trough (a "local periphery"), and moreover, that every succeeding skilled area again wrap around at least one more local center. On the other hand, if the first cut occurs on an upward stretch, then every unskilled zone must contain at least one local center, and every skilled zone must contain at least one local periphery.

The structure of local interactions, coupled with the market-clearing nature of prices, jointly impose this global structure on the spatial outcomes. In particular, the fact that prices clear markets, together with the assumption of linear direct utility, allow us to infer the existence of an "indifferent" agent, for whom aspirations and skill premia exactly balance out so that the acquisition of skills is an exact toss-up.[18] This indifferent agent imposes a lot of structure on spatial outcomes. The central idea used to obtain this structure is the fact that an indifferent agent in two different locations must locally see the same mix of skilled and unskilled individuals. (In the particular model we study here, this means that two neighboring cuts must lie on stretches of f with opposite slope.)

4.3. **Existence of Multi-Cut Equilibria.** It is important to note that (unlike single-cut equilibria), multi-cut purely segregated equilibria do not always exist. There is a good reason for this: the same structural discipline that pins down the spatial patterns of an equilibrium may on occasion (but not always!) be too restrictive to permit existence. To see this, imagine a unimodal distribution in which the upward stretch has markedly different slope from the downward stretch. Now the indifferent agents on either side of a two-cut equilibrium must have similar local perceptions; i.e., they must see the same mix of skilled and unskilled individuals in the ϵ-window around them. In other words, their locations must be chosen so that the local slopes of f around them (suitably weighted by height) are the same. However, the phrase "markedly different slopes" above means that in order to achieve the desired equality, we lose flexibility over the *aggregate* proportions of skilled to unskilled labor. In short, the proportions needed to assure equality of local skill ratios may be inconsistent with those needed to create market-clearing at the macroeconomic level.

This argument is informal and imprecise, but the following example tells us that it works.

Example **1.** *Let* $I \equiv [\underline{\iota}, \bar{\iota}] = \left[-1/\alpha, -1/\gamma \cdot \ln\left(1 - \gamma(1 - \frac{1}{2\alpha})\right)\right]$ *for* $\alpha \gg 1$*, and* $\gamma > 0$ *sufficiently small so that* $\alpha \geq 4\gamma$ *and* $\bar{\iota}$ *is large compared to* $1/\alpha$*. Consider the unimodal density function*

$$f(i) = \begin{cases} 1 + \alpha i, & i < 0 \\ e^{-\gamma i}, & i \geq 0. \end{cases}$$

[18]The existence of indifferent agents also presumes that the steady state is interior in skill acquisition. The Inada condition guarantees that skill premia are extremely large if no one acquires skills. Likewise, if everyone acquires skills, all skill premia vanish. These two assertions guarantee interiority.

On the linearly increasing part of f, $\rho(i)$ falls monotonically from $\rho(\underline{\iota}) = 1$ to

$$\rho(0) = \frac{F([0,\epsilon])}{F([-\epsilon,0]) + F([0,\epsilon])}.$$

Over the exponentially decreasing stretch, ρ falls continuously from $\rho(0)$ to $\rho(\epsilon) = (1-e^{\epsilon})/(e^{-\epsilon}-e^{\epsilon})$, stays constant at this level from ϵ up to $\bar{\iota}-\epsilon$, and then drops continuously again to $\rho(\bar{\iota}) = 0$. Cuts at any c and $c' \geq c+\epsilon$ must satisfy $\rho(c) = 1-\rho(c')$ in order to involve identical aspirations. So if $\rho(c) > 1-\rho(\epsilon)$ for every $c \in [\underline{\iota}, 0]$, then c' necessarily must be located towards the very right, namely, in the interval $[\bar{\iota}-\epsilon, \bar{\iota}]$; and the corresponding aggregate skill ratio is either at least $\lambda = \bar{\iota}-\epsilon$ (agents left of c are unskilled) or at most $\lambda' = 1-\bar{\iota}+\epsilon$ (agents left of c are skilled). A segregated equilibrium can then only exist if the given combination of technology T and preferences v admits such high or low skill ratios; many combinations do not, and for these no segregated equilibrium can exist.

It remains to confirm that indeed $\rho(c) > 1-\rho(\epsilon)$ for every $c \in [\underline{\iota}, 0]$ given the above density f. Since $\rho(i)$ is strictly decreasing for $i < 0$, it suffices to check that

$$\rho(0) = \frac{F([0,\epsilon])}{F([-\epsilon,0]) + F([0,\epsilon])} > \frac{F([0,\epsilon])}{F([0,\epsilon]) + F([\epsilon,2\epsilon])} = 1-\rho(\epsilon),$$

or

$$F([-\epsilon,0]) = \epsilon(1-\alpha\epsilon/2) < F([\epsilon,2\epsilon]) = [e^{-\gamma\epsilon} - e^{-2\gamma\epsilon}]/\gamma.$$

Because $e^{-\gamma i}$ is convex,

$$-\frac{\epsilon}{\gamma}\left.\frac{\partial e^{-\gamma i}}{\partial i}\right|_{i=2\epsilon} = \epsilon e^{-2\gamma\epsilon} = \epsilon\left(1 + \frac{-2\gamma\epsilon}{1!} + \frac{(-2\gamma\epsilon)^2}{2!} + \frac{(-2\gamma\epsilon)^3}{3!} + \ldots\right) < F([\epsilon,2\epsilon]).$$

Given that $|-2\gamma\epsilon| < 1$ for small ϵ, a sufficient condition for $\rho(0) > 1-\rho(\epsilon)$ is therefore

$$1-\alpha\epsilon/2 \leq 1-2\gamma\epsilon$$

or simply $\alpha \geq 4\gamma$, as we have assumed above.

That said, it is possible to provide sufficient conditions for the existence of multi-cut equilibria. We have, for instance:

Proposition **4.** *Suppose that f is unimodal and symmetric. Then a purely segregated regular two-cut equilibrium exists, provided ϵ is small enough.*

Proof. We prove existence of a two-cut equilibrium with the unskilled in the center and the skilled on the two peripheries. Without loss of generality the support of f is $[0,1]$ and the unique mode is located at $1/2$. For any cut $c \in [1/2, 1]$ it is natural to imagine its "mirror cut" on $[0, 1/2]$, which is the point $1-c$. Remembering that skilled labor is being placed on the sides, we may define

$$\lambda(c) \equiv F([c,1]) + F([0,1-c])$$

to be the aggregate amount of skilled labor associated with c. It is obvious that there exists $\underline{c} > 1/2$ such that $\lambda(\underline{c}) = \bar{\lambda}$, at which point wages are equalized. Clearly, an equilibrium c must exceed $\underline{c}$. Let B be the set $[\underline{c}, 1]$. Define $\underline{\epsilon} \equiv \underline{c} - 1/2$, and consider any $\epsilon < \underline{\epsilon}$. Notice that any equilibrium with cuts $\{1-c, c\}$, where $c \in B$, must be purely segregated and regular.

It remains to show that there is such an equilibrium. To do so, define

$$a(c) \equiv \rho(c)\bar{w}\left(\lambda(c)\right) + [1-\rho(c)]\underline{w}\left(\lambda(c)\right).$$

and note that it is continuous on B. Moreover,

$$\Psi\left(\lambda(\underline{c}), a(\underline{c})\right) < x,$$

simply because there is no wage differential when $c = \underline{c}$, while

$$\Psi\left(\lambda(c), a(c)\right) > x \text{ for all } c \text{ close to } 1,$$

because $\lambda(c) \downarrow 0$ as $c \uparrow 1$, so that $\bar{w}\left(\lambda(c)\right) \uparrow \infty$.

It follows that there exists $c \in B$ such that $\Psi\left(\lambda(c), a(c)\right) = x$. It is now easy to verify that $\{1-c, c\}$ forms a double-cut equilibrium. □

5. Comparisons Across Equilibria

Our model also makes possible sharp comparisons across equilibria. In this section we focus primarily on macroeconomic comparisons between unsegregated and segregated equilibria, and also between different geographic patterns of segregation. The next section will be devoted to welfare comparisons.

5.1. **City-Skilled Versus City-Unskilled Equilibria.** First we look at purely segregated regular multi-cut equilibria. Recall from the discussion following Corollary 2 that we've unearthed a particular spatial pattern of such equilibria. We begin by tightening this discussion. Define a segregated equilibrium to be *city-skilled* if there is some equilibrium cut which divides a skilled local mode from an unskilled local trough. If, on the other hand, there is some equilibrium cut that divides an unskilled local mode from a skilled local trough, call the equilibrium *city-unskilled*.

Note that these definitions refer to local properties (of *some* equilibrium cut in the whole set of cuts) so that in principle a segregated equilibrium could be both city-skilled *and* city-unskilled. But this cannot happen:

Proposition **5.** *A purely segregated regular equilibrium must be city-skilled or city-unskilled, and it can never be both.*

Proof. If a purely segregated equilibrium is neither city-skilled nor city-unskilled, *all* its cuts must lie on local peaks and troughs, which regularity rules out.

Suppose a purely segregated regular equilibrium is both city-skilled (the cut c verifies this) and city-unskilled (the cut c' verifies this). At c, it must be the case that $\sigma(c) > 1/2$, because individuals at c must see more skilled individuals than unskilled (we use regularity here). The opposite is true at c'. But then individuals at c and c' cannot have the same aspirations, which means that they cannot both be indifferent, a contradiction. □

Can city-skilled and city-unskilled equilibria coexist in the same model? There is no reason why not. For instance, we have:

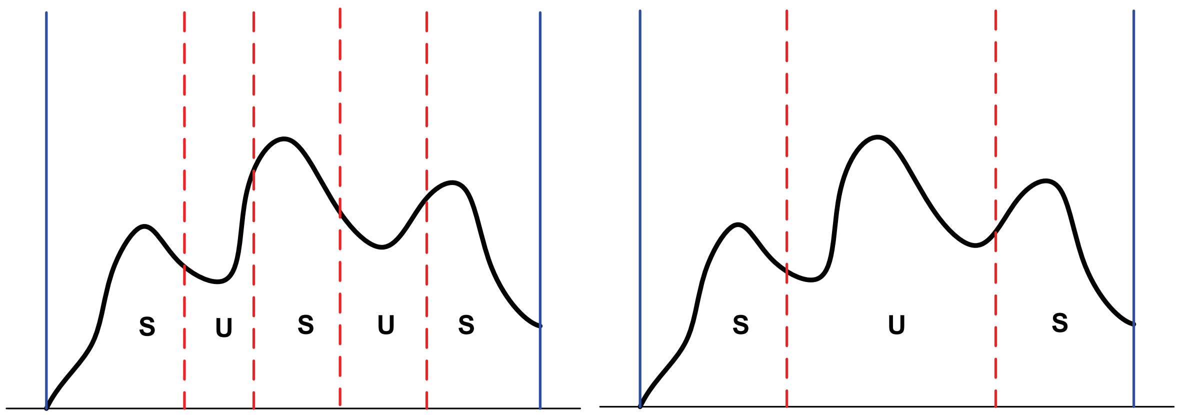

FIGURE 5. TWO CITY-SKILLED EQUILIBRIA

PROPOSITION **6.** *Suppose that f is unimodal and symmetric. Then both city-skilled and city-unskilled two-cut equilibria exist, provided ϵ is small enough.*

The proof of this proposition follows exactly the same lines as Proposition 4, and is therefore omitted.

What do city-skilled equilibria look like? By definition, *some* local mode (a "city") is occupied by skilled individuals. But this necessitates that *every* skilled interval must contain a local mode: in every skilled segment of a city-skilled equilibrium, there is a city. Exactly the opposite is true of city-unskilled equilibrium: every unskilled segment must contain a local city. In short, a city-unskilled equilibria exhibit "inner cities" in every interval of unskilled labor.

We add, however, that city-skilled equilibria don't rule out unskilled inner cities, nor do city-unskilled equilibria exclude the possibility of skilled urban areas. Figure 5 describes two city-skilled equilibria. In the first, there are no unskilled households around a local mode. In the second, there are unskilled local modes.

Of some interest is the fact that we can compare city-skilled and city-unskilled equilibria in terms of the aggregate skills they generate.

PROPOSITION **7.** *A city-skilled equilibrium must generate more skilled labor in the aggregate than any city-unskilled equilibrium.*

Proof. Consider a city-skilled equilibrium with aggregate skills λ, and consider a cut c that separates a skilled local mode from an unskilled local trough. It is easy to see that $\sigma(c) > 1/2$. In a similar way, there is a cut c' for a city-unskilled equilibrium (displaying aggregate skills λ') with $\sigma(c') < 1/2$. Now suppose, contrary to our assertion, that $\lambda \leq \lambda'$. Then $\bar{w}(\lambda) \geq \bar{w}(\lambda')$ and $\underline{w}(\lambda) \leq \underline{w}(\lambda')$, so that if a and a' are the aspirations at cuts c and c' under the two equilibria,

$$a = \sigma(c)\bar{w}(\lambda) + [1 - \sigma(c)]\underline{w}(\lambda) > \sigma(c')\bar{w}(\lambda') + [1 - \sigma(c')]\underline{w}(\lambda') = a'.$$

By complementarity of v and the wage and aspiration comparisons above, we must conclude that

$$\Psi(\lambda, a) > \Psi(\lambda', a'),$$

which contradicts condition (12) for at least one of the presumed equilibria. $\square$

5.2. **Segregated Versus Unsegregated Equilibria.** Unsegregated and segregated equilibria can also be compared. The comparison is not unambiguous, however. We conduct the analysis for small window sizes. The following proposition fully describes the limit outcomes of purely segregated equilibria as the window length ϵ shrinks to zero.

Proposition 8. *Let ϵ converge to 0, and $C(\epsilon)$ be a corresponding sequence of purely segregated equilibrium cut sets. If $\lambda(\epsilon)$ is the aggregate skill generated by $C(\epsilon)$, then $\lambda \to \lambda^*$, where λ^* solves*

$$\Psi\left(\lambda, \frac{\bar{w}(\lambda) + \underline{w}(\lambda)}{2}\right) = x. \tag{14}$$

Proof. First we claim that there exists a selection $c(\epsilon) \in C(\epsilon)$ for every ϵ, such that $c(\epsilon) \to c^*$, with $\underline{\iota} < c^* < \bar{\iota}$. If this were false, then all such selections have limit points that are either $\underline{\iota}$ or $\bar{\iota}$. It is easy to see that such a property implies the convergence of $\lambda(\epsilon)$ to either 0 or 1. Now the equilibrium λ can never exceed $\bar{\lambda}$ (which solves $\bar{w}(\bar{\lambda}) = \underline{w}(\bar{\lambda})$), which means that $\lambda(\epsilon) \to 0$. But then for small enough ϵ, we see that for every household at i with equilibrium aspiration $a_i(\epsilon)$,

$$\Psi\left(\lambda(\epsilon), a_i(\epsilon)\right) \geq \Psi\left(\lambda(\epsilon), 0\right) > x,$$

because v is unbounded in w by assumption. That is, all households want to acquire skills, which contradicts the presumption that equilibrium $\lambda(\epsilon)$ is close to 0. This proves the claim.

Pick a selection $c(\epsilon)$ as given by the claim, with $c(\epsilon) \to c^* \in (\underline{\iota}, \bar{\iota})$ as $\epsilon \to 0$. It is easy to see that for all ϵ small enough, both $c(\epsilon) - \epsilon$ and $c(\epsilon) + \epsilon$ are in the interior of I. Define $\rho(\epsilon) \equiv \rho(c(\epsilon))$. We claim that $\rho(\epsilon) \to 1/2$. Recalling (6), we see that

$$\rho(\epsilon) = \frac{F([c(\epsilon), c(\epsilon) + \epsilon])}{F([c(\epsilon) - \epsilon, c(\epsilon) + \epsilon])} = \frac{\int_{c(\epsilon)}^{c(\epsilon)+\epsilon} f(i)di}{\int_{c(\epsilon)-\epsilon}^{c(\epsilon)+\epsilon} f(i)di}.$$

Because f is continuous and $f(c^*) > 0$, the claim is proved. As a trivial corollary of this claim, $\sigma(c(\epsilon)) \to 1/2$ as well.

To complete the proof, note that by indifference at the equilibrium cut $c(\epsilon)$, we have

$$\Psi\left(\lambda(\epsilon), a(\epsilon)\right) = x \text{ for all } \epsilon,$$

where $a(\epsilon)$ denotes aspirations at the cut $c(\epsilon)$, given by

$$a(\epsilon) = \sigma(c(\epsilon))\bar{w}\left(\lambda(\epsilon)\right) + [1 - \sigma(c(\epsilon))]\underline{w}\left(\lambda(\epsilon)\right).$$

Combine these last two equations with the fact that $\sigma(\epsilon) \to 1/2$ as $\epsilon \to 0$, and pass to the limit to obtain the desired result. $\square$

The proposition states that as the window size ϵ becomes vanishingly small, *all* purely segregated equilibria, no matter what their spatial structure, must generate the same aggregate quantity of skills. The intuition is very simple: as the window size becomes small, an indifferent individual placed at a cut sees approximately equal numbers of skilled and unskilled individuals, no matter what the density function at her location looks like.[19] The incentives of the indifferent individual(s) must pin down the equilibrium wage differential; hence we obtain (14) for *all* purely segregated equilibria as ϵ becomes vanishingly small.

How literally we take this result depends in large part on our intuition about cognitive windows. We are agnostic on this question. We believe, in line with a large literature on local interactions, that the case of small windows is of interest, but at the same time do not push the line that such windows must be *vanishingly* small relative to the economy as a whole. (If we did, we would not have reported the comparisons in the previous section: there would be nothing to compare.)

In the end, we view the case of vanishingly small windows as a convenient device to compare segregated versus unsegregated equilibria, to which we now turn. How compelling the following observations are to the reader will depend, in part, on how comfortable she is with extremely small window sizes.

Proposition 8 holds a simple clue to comparing unsegregated and segregated equilibria. To exploit this, we parameterize the skill bias of the technology by some parameter A. Write the production function as some $T(\lambda, 1-\lambda, A)$, where A is normalized to lie in $[0, 1]$. Think of higher values of A as indexing production functions with greater degrees of skill bias.

Write the skilled and unskilled wages as functions $\bar{w}(\lambda, A)$ and $\underline{w}(\lambda, A)$ and assume these are continuous in A. We will need the following minimal restrictions, placed on the 50-50 input configuration $\lambda = 1/2$.

[A.1] As $A \to 1$, $\underline{w}(1/2, A)$ converges to 0, and for small enough A, $\bar{w}(1/2, A) \leq \underline{w}(1/2, A)$.[20]

[A.2] As $A \to 1$, the skilled wage increases enough (relative to the unskilled wage, which is converging to 0), to make investment worthwhile even at zero aspirations. That is, $v(\bar{w}(1/2, 1), 0) - v(0, 0) > x$.

The following proposition shows that the skill bias of the technology informs the comparison across segregated and unsegregated equilibrium:

Proposition **9.** *Under conditions [A.1] and [A.2], there exist two threshold values of the skill bias, $\bar{A}$ and $\underline{A}$, such that*

1. If $A > \bar{A}$ then there is an unsegregated equilibrium with higher skills than any segregated equilibrium, provided that ϵ is small enough.

2. If $A < \underline{A}$ then for ϵ small enough, every purely segregated equilibrium generates higher skills than any unsegregated equilibrium.

[19] This observation is similar to a parallel step used in the theory of global games, in which the rank-order of a particular signal is uniformly distributed as the noise becomes small.

[20] A more symmetric way of writing this is to say that $\bar{w}(1/2, A)$ converges to 0 as $A \to 0$. But the weaker form that we adopt allows for the possibility that skilled labor can do unskilled jobs; see footnote 10.

Proof. We prove statement 1. The proof of statement 2 essentially reverses the argument below (we indicate the main steps at the end).

Include A explicitly in the various expressions below to indicate that the corresponding variables depend on this parameter. For instance, we define $\Psi(\lambda, a, A) \equiv v(\bar{w}(\lambda, A), a) - v\left(\underline{w}(\lambda, A), a\right)$.

LEMMA **1.** *Fix A. Suppose that for some λ_0 the following is true:*

$$\Psi(\lambda_0, a(\lambda_0, A), A) > x,$$

where for all λ, we define

$$a(\lambda, A) \equiv \lambda \bar{w}(\lambda, A) + (1-\lambda)\underline{w}(\lambda, A).$$

Then there exists an unsegregated equilibrium with aggregate skill $\lambda > \lambda_0$.

Proof. Certainly, it must be the case that $\lambda_0 < \bar{\lambda}(A)$, for skilled and unskilled wages are equalized at the latter value, and so $\Psi(\lambda', a(\lambda', A), A) < x$ for all $\lambda' \geq \bar{\lambda}(A)$. By the intermediate value theorem, there exists $\lambda > \lambda_0$ such that $\Psi(\lambda, a(\lambda, A), A) = x$, and this must be an unsegregated equilibrium. □

By [A.1] and [A.2], there exists a threshold value $\bar{A}$ such that

$$v(\bar{w}(1/2, A), 0) - v\left(\underline{w}(1/2, A), 0\right) - x > 0 \text{ for all } A > \bar{A}.$$

By complementarity, it must therefore be the case that

$$\Psi(1/2, a(1/2, A), A) - x = v(\bar{w}(1/2, A), a(1/2, A)) - v\left(\underline{w}(1/2, A), a(1/2, A)\right) - x > 0 \tag{15}$$

for all $A > \bar{A}$, where remember that

$$a(1/2, A) = \frac{1}{2}\left[\bar{w}(1/2, A) + \underline{w}(1/2, A)\right]. \tag{16}$$

The relationships (15) and (16) tell us that $\lambda_0 = 1/2$ satisfies the conditions of Lemma 1. We therefore conclude that for each $A > \bar{A}$, there exists an unsegregated equilibrium with $\lambda > 1/2$.

Fix any such A. Proposition 8 tells us that if λ^* is a limit (as $\epsilon \to 0$) of some sequence of purely segregated equilibria, then

$$\Psi\left(\lambda^*, \frac{\bar{w}(\lambda^*) + \underline{w}(\lambda^*)}{2}, A\right) = x. \tag{17}$$

If $\lambda^* \leq 1/2$ the proof is complete, because the unsegregated $\lambda > 1/2$. So $\lambda^* > 1/2$. But then, by complementarity,

$$\Psi\left(\lambda^*, \lambda^* \bar{w}(\lambda^*) + [1-\lambda^*]\underline{w}(\lambda^*), A\right) > \Psi\left(\lambda^*, \frac{\bar{w}(\lambda^*) + \underline{w}(\lambda^*)}{2}, A\right),$$

and combining this inequality with (17), we must conclude that

$$\Psi\left(\lambda^*, \lambda^* \bar{w}(\lambda^*) + [1-\lambda^*]\underline{w}(\lambda^*), A\right) > x.$$

This time $\lambda_0 = \lambda^*$ satisfies the conditions of Lemma 1. So there exists an unsegregated λ that strictly exceeds λ^*, and the proof of statement 1 is complete.

To prove statement 2, first note that Lemma 1 is also true when both inequalities are reversed. Next, using [A.1], we show the existence of a lower threshold $\underline{A}$ such that the reverse inequality holds in (15). Now follow the same lines as in the rest of the proof, reversing the inequalities where needed. □

The proposition suggests that in societies which depend heavily on skilled labor in production, there is a case for desegregation on the grounds of greater skill accumulation in the aggregate. Desegregation lowers the aspirational incentives of those in already-skilled neighborhoods, but raises incentives for the unskilled who come into contact with them. The net outcome, however, is positive provided that society is skill-intensive in the first place.

The opposite is true in poorer societies where either technological backwardness or paucity of physical capital (not explicitly modeled here) lowers the relative demand for skilled labor. Proposition 9 then suggests that segregation may be a better generator of skills. In line with the first part of the proposition, such a situation is likely to be associated with one in which a minority of the population is skilled in unsegregated equilibrium.

These results should be treated with some caution. More work needs to be done to see how they will hold up in more general specifications of the nature of externalities in investment incentives. For instance, one might allow for two components of aspirations: one "local" in the sense developed in this paper, and the other "global", related to per-capita income economy-wide. It is not difficult to see that in this case, the ordering of Proposition 9 would be preserved and even magnified in intensity. Because per-capita income is positively related to skill acquisition in the relevant range of equilibrium outcomes, a global component of aspirations would simply reinforce the advantages of segregation or desegregation (depending on which parameters of that proposition are in force).

6. Welfare Effects of Segregation

Two considerations are fundamental to any discussion of welfare in these models. First, we need to fix ideas on whether equilibrium skill aggregates are too small or large relative to some "first best" steady state.[21] Nothing in the logical structure of our model allows us to *necessarily* assert that we have over- or under-accumulation of skills, though the model is entirely compatible with the realistic scenario of under-accumulation. We therefore separately impose the assumption (already made implicitly in the previous section) that at any equilibrium outcome, more skills are socially desirable. It is easy enough to derive such an outcome if we assume (among other things) that the private cost of education outstrips the social cost, perhaps because of imperfections in the capital market.

Next, there is the question of what the v-function looks like. Our characterization of the spatial structure of equilibria was driven entirely by the complementarity of investment incentive. The precise nature of this complementarity did not matter — neither for the existence of segregated and unsegregated equilibria, nor for the levels of aggregate

[21]There is also the more subtle issue of dynamics. It is entirely possible that a steady state is Pareto-dominated by another steady state, while at the same time the former is Pareto-optimal. (After all, two different steady states display two different initial conditions.) Mookherjee and Ray (2003) contains a more detailed discussion of this issue, which we avoid here.

investment. However, for welfare it *does* matter whether we see v as increasing or decreasing in aspirations. Our default assumption would be that v decreases in a: having higher aspirations to start with renders any given achievement less attractive. But the opposite may well be true. For instance, the source of local interactions could be in provision of local public goods: wealthier neighbors may help contribute to create better libraries and schools which both raise educational incentives and raise utility levels.[22]

In what follows, we assume throughout that greater skill generation is a good thing, but we discuss both positive and negative effects of higher aspirations on individual welfare.

6.1. **Altruism as Pure Aspiration.** As a first benchmark, suppose that parental concern for the pecuniary success of their children is driven by parental aspirations alone. That is, having a child with a wage above the local average is considered a success (an achievement parents may brag about), while a wage below the neighbors' average is viewed as a failure, or at least reduces parental well-being. In particular, suppose that meeting exactly the given aspiration level neither enhances nor reduces a parent's satisfaction. A child's human capital thus has the character of a positional good for the parent, with no intrinsic value.[23]

A straightforward example of purely aspirational altruism is one in which v depends *only* on the difference between wage and aspiration. More generally, assume that $v(w, w)$ is constant (say 0). The assumption that v is increasing in its first argument then implies that $v(w, a)$ is decreasing in a: the greater the average achievement of an individual's peers, the less will that individual value any given achievement by her child. Assume, moreover, that v is strictly concave in its first argument. This implies that the aggregate utility associated with aspirational altruism is negative and bounded away from zero in an unsegregated equilibrium with skill ratio λ^u:

$$\lambda^u v(\bar{w}, a) + (1 - \lambda^u) v(\underline{w}, a) < v(a, a) = 0$$

where $a = \lambda^u \bar{w} + (1 - \lambda^u)\underline{w}$.

The aggregate aspirational utility in any segregated equilibrium λ^s, in contrast, approaches zero as $\epsilon \to 0$: the size of the transition zones between skilled and unskilled stretches becomes negligible and households will exactly match their aspirations inside any stretch.

Tack on to this the second set of conditions in Proposition 9, in which the economy has an unskilled-labor bias and segregation generates higher aggregate skills. Under these conditions, total welfare is unambiguously higher in a segregated equilibrium. Segregation entails greater human capital and net output, and it also creates a smaller welfare loss associated with the individual mismatch of wage and aspirations.

---

[22]This would come close to the idea that a has a beneficial effect on the individual *cost* of education due to local spillovers caused, for instance, by local school financing or parental involvement at school, role models, or peer effects (see Bénabou 1993; 1996; or Durlauf 1994; 1996, and the empirical references therein). But it is not the same thing. Such cost externalities directly affect only those agents who are, in fact, investing in their child's human capital. In contrast, we are discussing the effect of aspirations on both investors and noninvestors.

[23]A variation of our model might dispense with the OLG setup: the investment decision could refer to an indivisible good, say, a swanky car, whose value is only positional and decreasing in the local share of swanky-car owners, and whose price is determined in an economy-wide market. The spatial ownership patterns are likely to mimic our equilibria, but the welfare implications would probably be different, unless one associates the same kind of macroeconomic productivity effects with car ownership as we do with human capital.

If, on the other hand, the first set of conditions in Proposition 9 applies, the welfare ranking of equilibria is ambiguous: the greater net output under desegregation needs to be traded off against the global "simmering" of unskilled households with unmet aspirations.

To be sure, this sort of analysis presumes that segregated societies can't see beyond their own segregation in the sense that only local (rather than economy-wide) inequality affects individual welfare. That may be an incorrect point of view, though we need to be careful in taking a more nuanced position. One counter-claim is that there is no such thing as segregation: the entire society determines aspirations and therefore investment. We doubt that this is the case. A second, more subtle position is that local aspirations drive investment while global comparisons have no functional role but create resentments for those at the bottom. We don't know if either of these positions is correct but if they are, the welfare pronouncements of this section may need to be revisited.

6.2. **A Direct Concern for Descendant Incomes.** The somewhat disturbing argument in favor of segregation (in the case in which skills are in low demand) is driven by another important assumption. This is the presumption that all education is purely positional: v decreases in a, and any wage increase is entirely neutralized by a corresponding increase in aspirations. It is arguably more realistic to suppose that parents also care about the offspring's wage *per se*, not just relative to a. This scenario is easy enough to capture: assume that $\tilde{v}(w) \equiv v(w, w)$ is actually strictly increasing in w. One possibility would be that $v(w, a) = v_1(w - a) + v_2(w)$ where the function v_1 captures purely aspirational altruism as discussed above, and the function v_2 incorporates a direct concern for descendant incomes.

Even if we retain the assumption that v is concave in w, the aggregate utility associated with aspirational altruism in an unsegregated equilibrium may now very well be sizable. Segregation will also be associated with aspiration utility, but (presuming again that the cognitive window is small), the "aspirational utility difference" is now given by

$$\left\{\lambda^u v(\bar{w}, a) + (1 - \lambda^u) v(\underline{w}, a)\right\} - \left\{\lambda^s v(\bar{w}, \bar{w}) + (1 - \lambda^s) v(\underline{w}, \underline{w})\right\}$$

with $a = \lambda^u \bar{w} + (1 - \lambda^u)\underline{w}$, and this can no longer be signed unequivocally. So even in the low-skill case, the welfare comparison between segregated and unsegregated equilibria is ambiguous and depends on one's precise assumptions.

The difference between welfare without segregation and with it can be usefully decomposed by conducting two thought experiments. One is to hypothetically redistribute total incomes inside local neighborhoods such that every household has an equal income, and hence equal aspirations. Comparing the resulting aspirational welfare with that implied by the actual wage distribution inside local neighborhoods provides a measure of the welfare loss associated with *intra-neighborhood inequality*. We can thus decompose the total welfare in an unsegregated equilibrium as follows:

$$W^u(\lambda^u) \equiv \left\{F(\lambda^u, 1 - \lambda^u) - \lambda^u x + v(a, a)\right\} - \left\{v(a, a) - \left(\lambda^u v(\bar{w}, a) + (1 - \lambda^u) v(\underline{w}, a)\right)\right\}$$

with $a = \lambda^u \bar{w} + (1 - \lambda^u)\underline{w}$. The first term in brackets captures the "gross efficiency" of skill λ^u: net output plus the aspirational welfare that would result under an equal income distribution. The second term corrects this by accounting for the social cost of intra-neighborhood inequality associated with λ^u.

The second thought experiment is to hypothetically redistribute the aggregate wage of each local community across skilled and unskilled stretches such that there is no inequality between them. Comparing the resulting aspirational welfare to that implied by the actual wage distribution produces a measure of the welfare loss associated with *inter-neighborhood inequality*. Since there are no differences between distinct neighborhoods in an unsegregated equilibrium, the latter involves no such costs of inter-neighborhood inequality. In contrast, a segregated equilibrium involves no intra-neighborhood inequality (hence no associated welfare costs), but there are inter-neighborhood differences. Total welfare in a segregated equilibrium can therefore be understood as the sum of three components: gross efficiency, zero social costs from intra-neighborhood inequality, and positive social costs stemming from inter-neighborhood inequality (again we treat transition zones as negligible and consider a vanishing ϵ):

$$W^s(\lambda^s) \equiv \Big\{F(\lambda^s, 1-\lambda^s) - \lambda^s x + v(a,a)\Big\} - 0 - \Big\{v(a,a) - \Big(\lambda^s v(\bar{w},\bar{w}) + (1-\lambda^s)v(\underline{w},\underline{w})\Big)\Big\}.$$

So whilst a segregated equilibrium involves no intra-neighborhood inequality, an unsegregated equilibrium has no inter-neighborhood inequality. Our earlier assumption of purely aspiration-driven altruism associated no social costs with inter-neighborhood inequality: households focused completely on the comparison with local peers. So wage inequality at the macroeconomic level did not matter for their individual, and hence aggregate, welfare. The more general formulation considered now permits both types of inequality to matter. The costs which are linked to either, together with technology-driven efficiency considerations, determine whether a social planner views segregated or unsegregated equilibria more favorably.

6.3. **Utility-Enhancing Aspirations.** As indicated above, aspirations may have a positive effect on parental utility, while still generating the required complementarity in investment. We briefly consider this case. We therefore have the function $v(w,a)$, which is increasing in both arguments, concave in the first argument, and with w and a being complements.

If the complementarity is strong enough, this formulation will actually give rise to a strictly *convex* $\tilde{v}(w) \equiv v(w,w)$. For such a strong complementarity, one can again be fairly specific in comparing aspirational welfare in segregated and unsegregated equilibria. Once again, study the case of low skill bias, so that segregation is better in terms of net output. The convexity of $\tilde{v}$ implies that

$$\lambda^s v(\bar{w}^s, \bar{w}^s) + (1-\lambda^s)v(\underline{w}^s, \underline{w}^s) > v(a^s, a^s)$$

for $a^s = \lambda^s\bar{w}^s + (1-\lambda^s)\underline{w}^s$. Lower per-capita income under desegregation, the monotonicity of $\tilde{v}(w)$, and the concavity of v in its first argument jointly imply

$$v(a^s, a^s) > v(a^u, a^u) > \lambda^u v(\bar{w}^u, a^u) + (1-\lambda^u)v(\underline{w}^u, a^u)$$

with $a^u = \lambda^u\bar{w}^u + (1-\lambda^u)\underline{w}^u$. So the unsegregated equilibrium involves smaller aspirational utility than segregation, and, in view of also the latter's greater net output, a social planner would prefer segregation. This echoes the same finding for purely aspiration-driven altruism, even though the reasoning is different: here, skilled agents have a positive effect on their neighbors' utility. Because the effect is stronger, the higher is the neighbors' wage, the overall gain is maximized by placing the skilled agents next to each other.[24] For purely aspiration-driven altruism, skilled agents actually decrease their neighbors' utility; placing them together minimizes the "damage".

Finally, the case of a high skill bias or concave $\tilde{v}(w)$ involve ambiguous comparisons. We exclude a detailed analysis.

7. Conclusion

In summary, we have investigated steady state equilibria of a model where agents' locations are given on a one-dimensional interval, and investment decisions are made by parents in the education of their children. There are two key externalities affecting investment decisions. One is an economy-wide pecuniary externality, resulting from the dependence of returns to investment on economy-wide investment ratios. Higher investment ratios in the economy lower skill premia and reduce investment incentives. The other is a local externality: earnings of residents in local neighborhoods affect investment decisions by affecting parental aspirations. These aspirations constitute one possible source of neighborhood effects; others may include a preference for conformity, or access to better schools. All of these forms of local externality induce complementarity between investment decisions at the local level, in contrast to substitutability at the economy-wide level.

We showed that steady state equilibria generally exist in which segregation arises: the interval is partitioned into subintervals in which residents all invest or do not. Unsegregated equilibria also exist in general. The macroeconomic comparison between segregated and unsegregated equilibria depends on the extent of skill-bias in the production technology. If skill-bias is low and skilled agents form a minority, segregation is associated with a higher economy-wide investment ratio and lower skill premia. The converse is true if skilled agents form a majority.

While the macroeconomic comparisons are robust with respect to the precise source of neighborhood externalities, the welfare comparisons are not. Segregation is unambiguously welfare-enhancing if skilled agents form a minority and, first, parental utility depends on children's future earnings only insofar as they deviate from aspirations, or, second, parental utility is enhanced by higher earnings of their neighbors (e.g., owing to access to better local schools). If parents also care about the earnings of their children *per se*, however, the welfare comparison is ambiguous, as segregation is associated with greater inequality *across* neighborhoods.

These results indicate that identification of the precise source of neighborhood externalities is not important if we are interested in positive analysis, e.g., the spatial structure of steady state equilibria. These are driven entirely by local complementarity properties of investment incentives. However, the source of neighborhood effects does matter when evaluating the

---

[24] The same result would obtain if the local wage average reduced the costs x of investment, rather than raising its subjective benefits. Also in such a setting, an unsegregated equilibrium with $\lambda^u < 1/2$ would be dominated in terms of investment incentives by a segregated equilibrium, and hence the latter involves a higher skill ratio. In addition, the costs of those who invest are minimized by placing them together.

welfare effects of segregation. Then how the investments of neighbors affect the *level* of utilities matters, over and above their effect on marginal utilities.[25]

A number of important issues remain to be addressed. We focused entirely on steady states, and ignored issues of non-steady state dynamics. It is conceivable that the unsegregated steady states are locally unstable: shocks which lead to some local clustering of investment decisions may possibly cause the system to converge thereafter to some segregated steady state. Whereas segregation may be robust to small random perturbations. Such issues have been addressed in models of segregation based on agent mobility, following the seminal work of Schelling. It would be interesting to examine whether there is a natural tendency for non-steady-state dynamics to converge to segregated steady states in our setting, based entirely on local investment complementarities rather than agent mobility.

We've used a one-dimensional interval for the set of all possible locations. Extension to other spatial contexts would presumably broaden the applicability of the model. While many of the results — for instance, those regarding the macroeconomic comparison of distinct segregation patterns — would extend, e.g., to the circle or the plane, some others would not. For instance, there cannot be a segregated equilibrium with a single cut on a circle. At the same time, our example of the non-existence of multi-cut segregated equilibria on the line can be adapted to the circle. Hence, there are economies for which no purely segregated equilibrium exists on a circle; and the same is probably true for the plane as well.

Finally, we ignored equilibria which, while not purely segregated, exhibit patterns of segregation so fine-grained (relative to neighborhood structures) that more than two adjacent subintervals lie on any one "side" in some cognitive neighborhoods. These patterns fall between the unsegregated and purely segregated equilibria that we focused on in this paper. Yet other possibilities include equilibria which are unsegregated on some portions and segregated on others. The analysis of such geographic patterns is technically involved, but needs to be addressed in future work.

References

Akerlof, G. A. (1997). Social distance and social decisions. *Econometrica 65*(5), 1005–1027.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *Quarterly Journal of Economics 115*(3), 715–733.

Akerlof, G. A. and R. E. Kranton (2005). Identity and the economics of organizations. *Journal of Economic Perspectives 19*(1), 9–32.

Banerjee, A. V. and A. Newman (1993). Occupational choice and the process of development. *Journal of Political Economy 101*(2), 274–298.

Bardhan, P. (2002). Decentralizatoin of governance and development. *Journal of Economic Perspectives 16*(4), 185–205.

Bénabou, R. (1993). Workings of a city: location, education, and production. *Quarterly Journal of Economics 108*(3), 619–652.

[25]It is also worth mentioning that a similar analysis of the spatial structure of equilibria will obtain in contexts involving purchase of status consumption goods in a static setting, but the welfare properties of segregation will be quite distinct.

Bénabou, R. (1996). Heterogeneity, stratification, and growth: macroeconomic implications of community structure and school finance. *American Economic Review 86*(3), 584–609.

Bisin, A., U. Horst, and O. Özgür (2006). Rational expectations equilibria of economies with local interactions. *Journal of Economic Theory 127*(1), 74–116.

Carneiro, P. and J. J. Heckman (2002). The evidence on credit constraints on post-secondary schooling. *Economic Journal 112*(482), 705–734.

Durlauf, S. N. (1994). Spillovers, stratification, and inequality. *European Economic Review 38*(3-4), 836–845.

Durlauf, S. N. (1996). A theory of persistent inequality. *Journal of Economic Growth 1*(1), 75–93.

Fernández, R. (2003). Household formation, inequality, and the macroeconomy. *Journal of the European Economic Association 1*(2-3), 683–697.

Fernández, R. and R. Rogerson (2001). Sorting and long-run inequality. *Quarterly Journal of Economics 116*(4), 1305–1341.

Galor, O. and J. Zeira (1993). Income distribution and macroeconomics. *Review of Economic Studies 60*(1), 35–52.

Ghiglino, C. and S. Goyal (2008). Keeping up with the neighbours: social interaction in a market economy. mimeo, University of Essex and University of Cambridge.

Heckman, J. J. and A. B. Krueger (Eds.) (2003). *Inequality in America: What Role for Human Capital Policies?* Cambridge, MA: MIT Press.

Loury, G. C. (1981). Intergenerational transfers and the distribution of earnings. *Econometrica 49*(4), 843–867.

Mookherjee, D. and S. Napel (2007). Intergenerational mobility and macroeconomic history dependence. *Journal of Economic Theory 137*(1), 49–78.

Mookherjee, D. and D. Ray (2003). Persistent inequality. *Review of Economic Studies 70*(2), 369–393.

Napel, S. and A. Schneider (2008). Intergenerational talent transmission, inequality, and social mobility. *Economics Letters 99*(2), 405–409.

Pancs, R. and N. J. Vriend (2007). Schelling's spatial proximity model of segregation revisited. *Journal of Public Economics 91*(1-2), 1–24.

Ray, D. (1990). Income distribution and macroeconomic behavior. Mimeo, Boston University.

Ray, D. (1998). *Development Economics*. Princeton, NJ: Princeton University Press.

Ray, D. (2006). Aspirations, poverty and economic change. In A. V. Banerjee, R. Bénabou, and D. Mookherjee (Eds.), *Understanding poverty*. Oxford: Oxford University Press.

Schelling, T. C. (1978). *Micromotives and Macrobehavior*. New York, NY: Norton.