

IZA DP No. 2050

**Ethnic Specialization and Earnings Inequality:
Why Being a Minority Hurts but Being a
Big Minority Hurts More**

Martin Kahanec

March 2006

Ethnic Specialization and Earnings Inequality: Why Being a Minority Hurts but Being a Big Minority Hurts More

Martin Kahanec

IZA Bonn

Discussion Paper No. 2050

March 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Ethnic Specialization and Earnings Inequality: Why Being a Minority Hurts but Being a Big Minority Hurts More^{*}

Social interaction is an important vehicle of human capital acquisition and its efficiency decreases in social distance. In this paper I establish that these two premises, given the socio-cultural differences between ethnic groups, explain the puzzling evidence that (i) minorities typically earn less than majorities and (ii) this earnings gap is increasing in the relative size of a minority in a given region. In particular, I argue that inter-ethnic social distance disadvantages smaller ethnic groups in human capital acquisition and that these efficiency differentials systematically expose minority and majority individuals to different incentives as concerns their choice of skills. As a result, minority and majority individuals tend to acquire different (combinations of) skills and the textbook substitution effect drives an efficiency unit of minority labor to sell at a relatively lower wage in a region with higher percentage of minority people. The conditions under which the efficiency disadvantage of the minority in social interaction and the substitution effect explain the abovementioned empirical findings are established. In addition, this study offers an answer why some minorities earn more than majorities, why minority individuals tend to spend more time socializing in families than in schools, and why integration may harm minorities.

JEL Classification: J15, J24, J70, O15

Keywords: human capital, earnings inequality, labor market, minority, network externalities, social interaction, ethnic specialization

Corresponding author:

Martin Kahanec
IZA
P.O. Box 7240
53072 Bonn
Germany
Email: kahanec@iza.org

^{*} I am grateful to Sjak Smulders, William J. Baumol, Patrick Francois, John Nye, Henri de Groot, Jean-Michel Glachant, Benito Arruñada, Mary Shirley, Piotr Stryszowski, Marta Kahancova, and audiences at various seminars at Tilburg University, IZA Bonn, University of Amsterdam, University of Lisbon, Ronald Coase Institute Workshop, 9th SPIE Annual Meeting, and the 2004 NAKE Day for their helpful comments on previous drafts of this paper. All remaining errors are mine. An earlier version of this paper appeared under a different name as CentER Discussion Paper No. 41, (2004). Financial support from Volkswagen Foundation for the IZA project on “The Economics and Persistence of Migrant Ethnicity” is gratefully acknowledged.

1. Introduction

Inequalities in socioeconomic conditions of Black and White Americans, Romany and White Europeans, and other minorities and majorities around the world are persistent and central features of the worldly history.¹ While there are many dimensions of socioeconomic inequality, labor income, as one of the major measures and determinants of socioeconomic inequality, is the principal focus of this study. Two robust empirical findings about the distribution of income between minority and majority peoples pose a challenge to economic theory. On the one hand, minorities typically earn less income per capita than majorities. On the other hand, minority-majority earnings disparity increases in the relative size of a minority in a region. The puzzling feature of these empirical regularities is that while being a member of the smaller social group in a region, the minority, is disadvantageous in earnings terms, minority people are relatively better off in regions where they are relatively less plentiful.

The scale puzzle that (i) minority individuals on average earn less than majority individuals and that (ii) this earnings differential is increasing in minority share in population in a given region has been corroborated in a sizeable empirical literature.² The early empirical studies on this topic include Blalock (1956, 1957), Heer (1959), Brown and Fuguitt (1972), and Frisbie and Neidert (1977). For example, Heer (1959) finds a correlation of -0.71 between the ratio of Black and White median per capita incomes and the percentage of Blacks. Frisbie and Neidert estimate the correlations between minority-majority income disparity and minority share in the population between 0.19 and 0.70. They go as far as to conclude that “one of the most consistent findings ... is that socioeconomic differentials vary directly with the relative numbers of a minority present in a given area ...”.³ More recently, in a micro-econometric study about the earnings of Black, Hispanic, Asian, and White men in the US, Tienda and Lii (1987) establish the existence of significant minority-majority income differentials and confirm that minority labor market percentages favor the majority while disadvantaging the minorities themselves. Finally, focusing

¹ Minority is understood to be a particular racial, ethnic, language, religious or national group of individuals who share socio-cultural characteristics such as culture, religion, language, history, beliefs, customs, values, and morals that make them distinct from the rest of the population – the majority. In a given region, the minority typically constitutes the smaller part of the population than the majority, but local concentrations may occur. The study does not deal with social groups formed on the basis of occupation, wealth, or other ordinal characteristics.

² Table A.1 in the Appendix summarizes the findings of the studies listed below in a greater detail.

³ Frisbie and Neidert (1977), p. 1007

on migrants, Borjas (1987) and Chiswick and Miller (2005) show that earnings of immigrants from a certain linguistic or ethnic group are decreasing in the concentration of similar people in the destination region.⁴

From the theoretical perspective, inter-ethnic earnings differentials attracted considerable attention. Becker (1957), Welch (1967), and Arrow (1972a, 1972b, 1973) argue that minority-majority economic inequality is a preference-driven phenomenon arising due to the so-called *taste for discrimination* of actors on the labor market.⁵ Williams (1947), Allport (1954), Blalock (1967), Reich (1971), and Bonacich (1972, 1976) argue that the hostility of a superordinate majority against minority people is increasing in the relative size of the minority. Another strand of literature, represented by Glenn (1964), Spilerman and Miller (1977), and Semyonov et al. (1984), advocates that discriminatory occupational structure creates an environment in which influx of minority workers crowds out majority workers into better jobs with higher pay. In an approach that understands discrimination as a consequence of a specific form of asymmetric information in the labor market, *statistical discrimination*, Lundberg and Startz (2002) and Coate and Loury (1993), building on the groundbreaking ideas of Phelps (1972), Arrow (1972a, 1972b, 1973), and Aigner and Cain (1977), argue that a priori actual or perceived asymmetries are maintained in the equilibrium through self-fulfilling expectations.

From what I denote the *local effects* perspective, Becker and Tomes (1979) and Loury (1981) argue that intergenerational transfers of ability to acquire human capital sustain human capital variation and thereby earnings inequality across families. Benabou (1993, 1996) and Durlauf (1994, 1996) explain persistent income stratification by the existence of local public goods or neighborhood externalities. In a similar vein, Steele (1992), Akerlof (1997), and Lundberg and Startz (1998) explicitly account for the role of social interaction in human capital distribution and suggest that it is the social or psychological (dynamic) externalities in segregated neighborhoods or workplaces that promote social and economic inequalities. In combination with the assumption of inferior initial conditions of minority people, as is often the case for immigrants or

⁴ Based on this empirical evidence, the proper interpretation of the scale puzzle involves comparing one minority across several regions of a given country, e.g. Blacks across U.S. counties, rather than different minorities in different countries, e.g. Chinese in Malaysia, Jews in the U.S., and Turks in Germany.

⁵ See also Darity (1982, pp. 72-75), Arrow (1998), and Loury (1998).

(past) discrimination, the local effects theories systematically explain persistent minority-majority earnings gap.

In this paper, applying the local effects approach to the specific conditions of minority-majority social interaction, I provide a theoretical explanation of the scale puzzle. First, I establish that positive external effects in skill acquisition and sociocultural differences between a minority and a majority that hinder their social interaction in social networks⁶ disadvantage the smaller social group, the minority, in terms of efficiency of human capital acquisition. As a consequence of this efficiency effect, a minority individual supplies less human capital and thus earns less than a majority individual, given the price of human capital. Second, the key insight of this paper is that in a world where heterogeneous skills are available in skill-specific social networks these efficiency differentials systematically expose minority and majority peoples to different incentives as concerns skill choice and, depending on the equilibrium organization of skill acquisition, make them acquire different (combinations of) skills. An important consequence of such differentiation, which has been corroborated by a number of studies,⁷ is that wages per efficiency unit of minority and majority labor typically differ, since these are no longer perfect substitutes. The imperfect substitutability of minority and majority labor in turn engenders the substitution effect, which in the present context implies that an efficiency unit of minority labor sells at relatively lower wage in regions where the minority is relatively larger.⁸ In effect, the efficiency and substitution effects work in opposite directions as concerns the relationship between minority share in the population and its relative earnings. The main result of this paper is that there are equilibrium regimes of skill acquisition under which the efficiency and substitution effect explain the scale puzzle. I classify these equilibrium regimes and establish the parametric conditions that support this result.

⁶ As in the previous paper, social network is understood to be a social structure between individual actors that facilitates social interaction among its members.

⁷ To wit, indicating a degree of differentiation on the labor market, Altonji and Blank (1998) report that minority workers are overrepresented in less skilled jobs and Blacks in the US are overrepresented in some kinds of jobs such as public administration. Occupational differentiation explored by e.g. Blalock (1957), Brown and Fuguitt (1972), and Hirschman and Wong (1984). From the empirical perspective, Grant and Hamermesh (1981), Grossman (1982), Borjas (1983, 1987), and Kahanec (2006) establish imperfect substitutability of minority and majority labor.

⁸ The substitution effect is a direct consequence of the textbook economic law of diminishing marginal product.

The argument proceeds as follows. In the following section I describe the social and economic environment of the model and elaborate on the main assumptions on which the argument is based. Next, I present a formal model and establish its main predictions. Finally, I discuss the robustness and relevance of the presented theory and conclude.

2. The Social and Economic Environment

2.1. The Main Assumptions

This paper draws on several insights about social embeddedness of human capital acquisition developed in the literature. That individuals learn from their peers, friends, and neighbors has been proposed by a number of scholars.⁹ As Lucas (1988) points out, “human capital accumulation is a *social* activity, involving *groups* of people in a way that has no counterpart in the accumulation of physical capital.”¹⁰ Allen (1982), Ellison and Fudenberg (1993, 1995), and Bala and Goyal (1998) investigate the role of social interaction in learning about optimal actions. Valente (1995), Feick and Price (1987), Gladwell (2000), and Foster and Rosenzweig (1995) substantiate such approach and observe that social networks are an important vehicle of information sharing. These authors document that colleagues, friends, or neighbors share information about their discoveries, experiment outcomes, or search results. Conley and Udry (2002), Foster and Rosenzweig (1995), and Munshi (2004) provide evidence that social interactions significantly affect farmers’ profitability upon adoption of new technologies, arguing that this finding implies that farmers learn about the best practices in social interaction with their peers and neighbors, rather than only mimicking their behavior.

A number of scholars, such as Glaeser et al. (2002), Foster and Rosenzweig (1995), and Lazear (1999), maintain that social interaction in social networks often involves positive externalities such that the aggregate resources of a network exceed the naïve sum of individual contributions. Foster and Rosenzweig (1995) develop a framework in which the efficiency of social learning improves in the number of involved individuals whenever social learning exhibits social

⁹ Early theories about human capital include Becker (1962), Mincer (1958), and Schultz (1961). The literature on social embeddedness of human capital acquisition includes Rees and Schultz (1972), Loury (1977), Bourdieu (1986), and Coleman (1988, 1990).

¹⁰ Italics are original, p.19.

memory.¹¹ Based on this literature, I adopt the premise that the benefits from social interaction are increasing in the number of people involved as the first essential assumption of this paper.¹² Namely, I assume that skill acquisition process exhibits external network effects¹³ that positively depend on the size of the social network in which the particular skill is acquired.

It is natural to argue that benefits from social interaction not only depend on the number of individuals one interacts with but also on who these individuals are. In the context of minority-majority social interaction, sociocultural differences between minorities and majorities are likely to determine the quality of social interaction in any network. To operationalize these sociocultural differences, in line with Poole (1927) and Lazear (1999), I define social distance to be the measure of subjective and objective dissimilarities between social groups that hinders social interaction between the members of these social groups.¹⁴ The natural corollary of the definition of social distance above is that agent's ability to benefit from social interaction in a given network negatively depends on her social distance to the other members of this network. Based on this, the second essential assumption of this paper is that individual benefits from network effects are decreasing in interpersonal social distance.¹⁵

To complete the description of the social organization of skill acquisition, given the omnipresent segregation of social institutions, it is assumed that institutionally exclusive and inclusive social networks exist in the economy.¹⁶ Specifically, while inclusive social networks permit any membership, a given exclusive network only permits memberships from one social group.¹⁷ The

¹¹ Goyal (2003) surveys the literature on social learning.

¹² Inefficiencies stemming from the size of social networks, such as inefficient herding, status, and congestion, are certainly possible. The focus of this paper is on the benefits from social interaction in social learning, however.

¹³ Network effects arise whenever benefits from a good or service, here the service of social network in skill acquisition process, increase in the number of individuals already owning that good or using that service. One consequence of a network effect is that the use of a network service by one individual indirectly benefits others who use it. This side effect in a transaction is known as network externality.

¹⁴ In contrast to Akerlof (1997), who studies endogenous social distance between homogeneous agents, I consider social distance between members of different social groups to be a predetermined variable that reflects the defining distinctiveness of social groups.

¹⁵ Note that social distance is fully symmetric on the individual level. Assuming a priori asymmetry of sociocultural differences, although trivially incorporable into the argument, would be largely ad hoc and racially prejudiced.

¹⁶ There is an enormous literature on social structure and ethnic segregation. Recent contributions include Massey and Denton (1993) and Farley and Frey (1994). Ethnic segregation has been documented by e.g. Farley and Frey (1994), Glaeser and Vigdor (2001), Reardon et al. (2000).

¹⁷ Thus, exclusive social networks are always segregated. Inclusive social networks may be integrated as well as segregated; the distinction made in this paper is that exclusiveness (inclusiveness) is understood as exogenous

prime examples of typically exclusive networks include families, kinships, social networks in ghettos, religious groups, expatriate communities, radical groups, and ethnically or religiously exclusive schools and clubs. Most schools, student societies, workplaces, academic communities, and cybernetworks¹⁸ are typically inclusive. These examples suggest that exclusive and inclusive social networks are typically different with respect to, inter alia, their complexity, objectives, functions, and the strength of ethnic or religious character. Arguably, these differences transpire into the different characteristics of skills acquired in exclusive and inclusive networks.¹⁹ On the one hand, inclusive social networks generally support more formal and cognitive skills such as those in e.g. mathematics, medicine, metal processing, machine operating, and banking. On the other hand, in exclusive social networks people typically acquire less formal and non-cognitive skills such as verbal and non-verbal communication skills including language skills, general social knowledge and socialization skills, and capability of self-motivation, but also particular arts and crafts skills whenever these are specific for a given social group. Therefore, I assume that the skills acquired in exclusive networks are generally different from those acquired in inclusive networks. I let “exclusive” and “inclusive” denote the respective skills and networks.

2.2. The Driving Mechanisms: The Efficiency and Substitution Effects

Having described the key assumptions, in this section I indicate the main mechanisms of the formal argument. First, through network effects, the efficiency of skill acquisition in a given social network is a function of its size. In any given network, due to the social distance between minority and majority, individuals benefit from a larger relative number of network members from their own social group. Moreover, if individuals from some social group choose to segregate, the size of their segregated social networks is limited by the size of their social group. Therefore, the efficiency effect favors relatively larger social groups and so offers an explanation of the first part of the scale puzzle: why minorities typically earn less than majorities.

Second, for a member of a given social group network effects and social distances generate efficiency benefits to joining and investing in that social network that is chosen by the other

institutional constraint on network membership while segregation (integration) as endogenous variable concerning equilibrium organization of social interaction.

¹⁸ Social networks in the cyber space, such as the users of the Internet.

¹⁹ Coleman et al. (1966) and Heckman (2000) discuss cognitive and non-cognitive skills.

members of his or her social group. These efficiency benefits induce minority and majority each to specialize in one, possibly different social network or, if exclusive and inclusive skills are both essential for an individual, to invest their time differently between exclusive and inclusive social networks. As a result, network effects and social distances direct minority and majority individuals to acquire different skills or they induce them to acquire different combinations of skills. Such differentiation engenders the substitution effect: individuals who supply skills that are scarcer earn higher wage for an efficiency unit of their labor than those that supply more abundant skills. It follows that the substitution effect rewards members of smaller social groups, as the aggregate supply of their skills is relatively smaller.²⁰ Consequently, the substitution effect offers an explanation of the second part of the scale puzzle: that smaller minorities earn relatively more than larger ones.

To summarize, as the relative size of a social group increases, it benefits from the efficiency effect while being hurt by the substitution effect in relative earnings terms and vice versa. In the analysis below I formally demonstrate that these two effects arise despite the fact that individuals only differ in their group membership, which determines to members of which groups a particular individual is socially close (distant), and the only asymmetry in the model concerns group sizes. In particular, I establish that the substitution and efficiency effects can produce a nonmonotonic pattern of earnings inequality between social groups that is consistent with the scale puzzle. The conditions under which this is the case are then identified and discussed.

3. The Model

3.1. Labor Demand

In this section, I study the demand side of the labor market where the society is divided into two social groups – the minority I and the majority J – and clarify the extent to which it accounts for the substitution and efficiency effects. Let i and j denote the respective members and I and J the respective measures of the continuums of minority and majority agents, where $I < J$ and I adopt a convenient normalization that $I + J = 1$. I assume that all individuals are identical with respect

²⁰ Besides the obvious reason that a smaller social group supplies lesser measure of skills on aggregate, due to the efficiency it also has a lower per capita supply of skills.

to their preferences and endowments, group membership excepting. Individual preferences are represented by a standard utility function $u(\cdot)$ that increases in individual consumption, C_k , where $k \in \{i, j\}$.

Let the consumption good be produced by combining labor input of minority individuals, H_i , and majority individuals, H_j , in a perfectly competitive industry according to the constant elasticity of substitution (CES) aggregate production function

$$C = \left(\left(\int_0^I H_i di \right)^{\rho-1/\rho} + \left(\int_0^J H_j dj \right)^{\rho-1/\rho} \right)^{\rho/(\rho-1)} \quad (1)$$

with the elasticity of substitution $\rho > 1$. According to this specification, labor of any given type has decreasing marginal returns, production exhibits constant returns to scale, and no type of labor is essential in production. Moreover, while members of the same social group are perfectly substitutable in production, the elasticity of substitution between minority and majority labor ρ is not a priori restricted to be finite. In particular, whenever minority and majority labor is perfectly substitutable, production does not distinguish between minority and majority labor.²¹

Similarly to the statistical discrimination literature, I adopt the asymmetric information hypothesis about the labor market. In particular, I assume that while employers observe the aforementioned observable characteristics of social group membership of individuals²² and the measures of labor they supply, H_i and H_j , they are not able to directly observe the marginal product of labor supplied by any individual.²³ From experience or statistical investigation, however, they understand that social group membership predicts the marginal product of individual labor and thus they know the expected marginal products of members of any social group. By corollary, employees from the same social group are not distinguishable with respect to their type of labor, as they do not perceptibly differ, and they always receive the same wage for a unit of their labor input.

²¹ The issue of substitutability of minority and majority labor is elaborated in the analysis of the supply side below. Proposition 3 implies that the production function (1) can be seen as a harmless simplification of a more general production technology with an arbitrary number of types of labor with a given elasticity of substitution.

²² E.g. skin color, group-specific name, or accent.

²³ This assumption also implies that employers cannot remunerate a worker separately for the skills and labor time he or she supplies.

Given the infinitesimal measure of any individual, all prices are taken as given at the individual level and the production function (1) gives rise to individual demands for labor

$$H_i = P_C^\rho W_i^{-\rho} C / I \quad (2a)$$

$$H_j = P_C^\rho W_j^{-\rho} C / J, \quad (2b)$$

where P_C is the price of consumption good C and W_i and W_j are the wages per unit of labor input of minority and majority individuals, respectively. As a result of the homogeneity of degree one of the CES production function, the sector does not generate any profits in the equilibrium and we can derive that $P_C = (W_i^{1-\rho} + W_j^{1-\rho})^{1/(1-\rho)}$. Combining the demands for H_i and H_j in (2a-b), one obtains the relative demand for labor

$$w = \left(\frac{I}{1-I} h \right)^{\frac{-1}{\rho}}, \quad (3)$$

where $w \equiv W_i/W_j$ and $h \equiv H_i/H_j$. Equation (3) is the main result from studying the demand side of the economy. It plainly reveals the substitution effect that, given a finite ρ , the relative wage w is decreasing in the relative size of the social group I and its relative supply of labor h . It also highlights the importance of ρ for the substitution effect: only a finite ρ makes the substitution effect operative.

Let us now turn to earnings as determined by the demand side. Premultiplying equation (3) by h and defining $\Omega_k \equiv H_k W_k$ to be the labor earnings (and the only income) of individual k , we derive the following expression for minority-majority ratio of per capita earnings:

$$\omega \equiv hw = h \left(\frac{I}{1-I} h \right)^{\frac{-1}{\rho}} = \left(\frac{I}{1-I} \right)^{\frac{-1}{\rho}} h^{\frac{\rho-1}{\rho}}. \quad (4)$$

Observing the properties of equation (4), at least two conditions about the relative supply of labor under which equation (4) generates patterns of income inequality consistent with the scale puzzle for some I , involving regularities that $\partial\omega(I)/\partial I < 0$ and $\omega(I) < 1$, can be identified. First, since $(I/(1-I))^{\frac{-1}{\rho}} > 1$, the efficiency effect must favor members of larger social groups so that

$h(I) < 1$ for some I and thus there exists ρ such that $\omega(I) < 1$ for some I .²⁴ Below, I will show how the model satisfies this condition when taking into account the organization of supply of labor and skill acquisition that is characterized by network effects and social distances, as described above.

The second condition constitutes the core of the argument. Equation (4) clearly reveals that whenever minority and majority labor is perfectly substitutable and thus $\rho \rightarrow \infty$, equation (4) boils down to $\omega = h$ and the substitution effect is non-operative, since wages per efficiency unit of labor are equal for all individuals. In such case, as established below, network effects and social distances still generate minority-majority earnings gap but the gap is diminishing in the relative size of minority, contrary to the scale puzzle. Therefore, an additional formal argument has to be made in order to substantiate the existence of the substitution effect. In particular, it is essential to establish that minority and majority labor is imperfectly substitutable; that is, that ρ is finite. In the following section I analyze the supply of labor and show how the supplies of minority and majority labor depend on the sizes of social groups through network effects and social distances. Most importantly, I establish that, under certain conditions, network effects and social distances determine the organization of supply of labor such that it justifies imperfect substitutability of minority and majority labor and thus the existence of the substitution effect.

3.2. Labor Supply

Individuals are each endowed with one unit of time that they divide between acquisition of human capital and time spent working. Human capital as well as time spent working increase the individual supply of efficient labor H_k , which I conceptualize to be the measure of labor in efficiency units that comprises labor time and human capital. In particular, it is assumed that efficient labor is a composite of time-empowered exclusive and inclusive skills. Denoting exclusive and inclusive skills and network types $m \in \{x, n\}$, respectively, I assume the constant elasticity of substitution technology of producing H_k efficiency units of labor

²⁴ That such ρ exists is clear from the fact that $\rho \rightarrow \infty$ implies $\omega \rightarrow h$. Thus, if $h < 1$ there always exists large enough ρ such that $\omega < 1$.

$$H_k = \left[(S_{k,x} (T_{k,x} - L_{k,x}))^{\varepsilon-1/\varepsilon} + (S_{k,n} (T_{k,n} - L_{k,n}))^{\varepsilon-1/\varepsilon} \right]^{\varepsilon/(\varepsilon-1)} \quad (5)$$

where $S_{k,m}$ is the measure of skills of type m of agent k , $T_{k,m}$ is the corresponding total time invested in acquiring skills and utilizing them in production, and $L_{k,m}$ is the corresponding time spent on acquiring skills.²⁵ Accordingly, $T_{k,m} - L_{k,m}$ is the time individual k allocates to utilizing skill m in production. The finite and positive parameter ε denotes the elasticity of substitution between time-empowered exclusive and inclusive skills in production of individual efficient labor and reflects their imperfect substitutability. Skills are acquired according to a decreasing-returns-to-scale technology in social networks

$$S_{k,m} = L_{k,m}^{\phi} (1 + N_{k,m}), \quad (6)$$

where $N_{k,m}$ is the external network effect benefit in network m enjoyed by its member k and $\phi \in (0,1]$ is the measure of decreasing returns to time spent in skill acquisition. Throughout the paper I assume that agents take network effects as given, given the infinitesimal measure of any individual.

As discussed above, exclusive and inclusive social networks differ in terms of membership they permit and skills they support. Given the difference of exclusive and inclusive skills, from the production technology (5) we see that the qualitative properties of individual labor are determined by the combination of skills that constitute efficient labor of a worker. I operationalize this qualitative variation of efficient labor such that efficiency units of labor that consist of different (combinations of) skills are imperfect substitutes on the labor market. Thus, for example, if the skills of one individual are predominantly exclusive and the skills of the other agent are predominantly inclusive, the elasticity of substitution of labor of these two individuals is finite. Formally, defining $s_k \equiv S_{k,x}/S_{k,n}$, whenever $s_k \neq s_{k'}$ ($s_k = s_{k'}$) for individuals k and k' , the elasticity of substitution between H_k and $H_{k'}$ is finite (infinite). Because s_k is determined by the organization of human capital acquisition, which is endogenous in the model, the elasticity of substitution between minority and majority labor ρ is in this sense endogenous as well.

²⁵ This technology of producing efficient labor H_k can be, without any bearing on the argument of this paper, reinterpreted as the production function of the intermediate good H_k , which is an input in the production of the consumption good C .

Turning to the individual problem of time allocation, individuals maximize their utility, taking their resource constraints, available technologies, network effects, wages per unit of their efficient labor²⁶, and the price level as given. From the properties of the utility function it follows that the agents' problem boils down to

$$H_k^* \equiv \underset{L_{k,m}, T_{k,m}}{\text{Max}} |H_k| \quad (7)$$

subject to (5), (6), and the resource constraints $T_{k,m} \geq 0$, $L_{k,m} \geq 0$, and $T_{k,x} + T_{k,n} \leq 1$. Solving the maximization problem, it is straightforward to see that individuals divide their time between acquisition and utilization of skills according to the rule²⁷

$$L_{k,m} = \frac{\phi}{1+\phi} T_{k,m}. \quad (8)$$

Thus, agent k spends a fixed share $\phi/(1+\phi)$ of the time that he allocates to skill m , $T_{k,m}$, on acquiring this skill. The rest of this time, $T_{k,m}/(1+\phi)$, or $T_{k,m} - L_{k,m}$, is spent on utilizing it. The following proposition characterizes the solution of the maximization problem (7). To save on notation in what follows I define $\tilde{N}_{k,m} \equiv \left(\left(\frac{\phi}{1+\phi} \right)^\phi (1 - \frac{\phi}{1+\phi}) (1 + N_{k,m}) \right)^{\varepsilon-1/\varepsilon}$.

Proposition 1

If $\varepsilon \geq (\phi+1)/\phi$, the optimal solution to agents' problem (7) arises as a corner solution where all the time available to an individual is spent on acquisition and utilization of the one skill whose acquisition is most efficient. In particular,

$$\left(\varepsilon \geq (\phi+1)/\phi \wedge \tilde{N}_{k,n} \geq \tilde{N}_{k,x} \right) \Rightarrow H_k^* = \tilde{N}_{k,n}^{\varepsilon/(\varepsilon-1)} \quad (9a)$$

and

$$\left(\varepsilon \geq (\phi+1)/\phi \wedge \tilde{N}_{k,n} \leq \tilde{N}_{k,x} \right) \Rightarrow H_k^* = \tilde{N}_{k,x}^{\varepsilon/(\varepsilon-1)}. \quad (9b)$$

$$\text{If } \varepsilon < (\phi+1)/\phi, \text{ the interior solution } H_k^* = \left(\tilde{N}_{k,x} T_{k,x}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} + \tilde{N}_{k,n} T_{k,n}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} \right)^{\varepsilon/(\varepsilon-1)} \quad (10)$$

Solves the agents' problem (7), where the optimal time allocation is governed by the arbitrage condition

²⁶ As a consequence of the assumption of asymmetric information in the labor market discussed above.

²⁷ Proof in the Appendix.

$$t_k \equiv \frac{T_{k,x}}{T_{k,n}} = \left(\frac{1 + N_{k,x}}{1 + N_{k,n}} \right)^{\frac{\varepsilon-1}{1+\phi-\varepsilon\phi}}, \quad (11)$$

giving rise to equilibrium time allocations $T_{k,x} = t_k / (1 + t_k)$ and $T_{k,n} = 1 / (1 + t_k)$.

Proof in the Appendix.

Corresponding to interior and corner solutions in Proposition 1, I define two classes of regimes, specialization and diversification, respectively. In particular, specialization regimes prevail if and only if $\varepsilon \geq (\phi + 1) / \phi$ and diversification regimes prevail if and only if $\varepsilon < (\phi + 1) / \phi$.²⁸ From Proposition 1 and equations (5), (6), and (8), under specialization the relative supply of labor is

$$h = \frac{1 + N_{i,m}}{1 + N_{j,m}} \quad (12)$$

and, under diversification,

$$h = \frac{\left(\frac{1}{1+t_i} \right)^{\phi+1} \left(\left(t_i^{\phi+1} (1+N_{i,x}) \right)^{\varepsilon-1/\varepsilon} + (1+N_{i,n})^{\varepsilon-1/\varepsilon} \right)^{\varepsilon/(\varepsilon-1)}}{\left(\frac{1}{1+t_j} \right)^{\phi+1} \left(\left(t_j^{\phi+1} (1+N_{j,x}) \right)^{\varepsilon-1/\varepsilon} + (1+N_{j,n})^{\varepsilon-1/\varepsilon} \right)^{\varepsilon/(\varepsilon-1)}}. \quad (13)$$

As one can see, the relative supply of labor is fully determined by network effects and time allocation in skill acquisition.²⁹ In the following section I specify network effects and investigate the allocation of individual involvements across networks k under the various equilibrium regimes of skill acquisition.

3.3. Network Effects

In the skill acquisition technology (6), external network effects play a pivotal role in determining the efficiency of acquiring skills in a social network and thus the equilibrium allocation of individual involvements across social networks. In line with the arguments above, network effects that any given agent k enjoys in social network m depend on the extent of social

²⁸ In general, there may be dual regimes in which agents of one social group diversify while agents of the other social group specialize. Because the choice to diversify or specialize entirely depends on parameters ε and ϕ of the model and these are assumed to be the same for every agent in the economy, I disregard these cases here. The rationale for this approach is that the emphasis in this paper is put on the question how network effects engender heterogeneity of human capitals of minority and majority and how this heterogeneity translates into income inequality when individual characteristics are the same for both social groups.

²⁹ In particular, it does not depend on wages. The reason is that individuals take wages as given, time has no other value but in skill acquisition, and skill acquisition does not involve any pecuniary exchange.

interaction therein, which is measured by the total time agents spend interacting in this network. Consistently with the assumption about social distance, agents benefit more from interaction with ethnically similar agents as compared to ethnically distant agents. This effect is captured by the social distance parameter $\delta > 0$. For the sake of clarity of exposition, I posit that the one-dimensional social distance parameter completely represents the multidimensional dissimilarities between the minority and the majority. Based on these premises, I assume network effect specifications

$$N_{i,m}(I_m, L_{i,m}, J_m, L_{j,m}, \delta) = \left(I_m L_{i,m} + \frac{1}{1+\delta} J_m L_{j,m} \right)^\gamma \quad (14)$$

$$N_{j,m}(I_m, L_{i,m}, J_m, L_{j,m}, \delta) = \left(\frac{1}{1+\delta} I_m L_{i,m} + J_m L_{j,m} \right)^\gamma, \quad (15)$$

where the parameter $\gamma \in (0,1]$ captures decreasing returns to involvement of individuals in a given social network and I_m and J_m are the numbers of, respectively, minority and majority members in network m . These numbers depend on the equilibrium organization of skill acquisition as discussed below.

3.4. Equilibria

I adopt the Nash concept of equilibrium where agents choose social networks (skills) freely and the equilibrium arises as the state where no agent has incentives to deviate, that is, to change his or her allocation of time across networks. Given this equilibrium concept, we can state the following general propositions about stable equilibrium regimes of skill acquisition:

Proposition 2

In any stable equilibrium, no agent is involved in more than one network of any given type, exclusive or inclusive.

Proof in the Appendix.

Proposition 3

In any stable equilibrium, all members of a given social group choose the same combination of skills to acquire.

Proof in the Appendix.

Proposition 2 is a consequence of network effects: in a stable equilibrium, there cannot be two (or more) equally efficient social networks of the same type for any given individual because any perturbation of agents' involvements makes one of them less efficient and causes this individual to abandon it. Proposition 3 is mainly due to asymmetric information in the labor market that implies that individuals take their wages as given with respect to their choice of skills, so this choice is purely driven by efficiency concerns. Similarly as in Proposition 2, no stable equilibrium involves two distinct equally efficient combinations of social networks. Consequently, given asymmetric information in the labor market, network effects and social distances in skill acquisition coordinate individuals such that at most two different types of labor are supplied – minority- and majority-specific. In this sense, as mentioned above, the production function (1) can be seen as a harmless simplification of a more general production technology with an arbitrary number of types of labor H_k .

3.4.1. Specialization

This section studies equilibrium regimes of skill acquisition under specialization. Given Proposition 2 and 3, under specialization all members of a given social group choose exactly one and the same social network to join and thus skill to acquire. In effect, five different non-empty networks may arise in the economy under specialization; these are highlighted in Table 1.³⁰

Table 1: Social networks under specialization.

<i>Social network type</i>	<i>Permitted membership</i>	<i>Possible membership</i>
Minority exclusive	Minority	<i>I</i>
		Empty
Majority exclusive	Majority	<i>J</i>
		Empty
Inclusive	Any	<i>I</i>
		<i>J</i>
		<i>I and J</i>
		Empty

³⁰ It is worthwhile to recall that whether a network is exclusive or inclusive is an institutional constraint and is not determined by who its members are. For example, it may happen that a school as a social network permits minority (and majority) people to participate, but these choose not to. The school then only has majority pupils, but it remains inclusive, nonetheless.

It turns out that there are three stable and two conditionally stable equilibrium allocations of minority and majority individuals across social networks and thus skills of different types under specialization. I list these equilibrium allocations and investigate their stability in Proposition 4. Table 2 depicts the properties of these equilibria.

Proposition 4

Under specialization the following equilibria are always stable in the Nash sense:

1. Each social group specializes in exclusive skills joining its exclusive network (EE)
2. Social groups specialize in inclusive skills joining the same inclusive network (II)
3. The minority specializes in inclusive and the majority in exclusive skills (IE)

The following allocations

4. The minority specializes in exclusive and the majority in inclusive skills (EI)
5. Social groups specialize in inclusive skills, acquiring them in two non-connected inclusive networks, each composed of members of only one social group (IS)

are stable if and only if

$$I \geq 1/(2 + \delta). \tag{16}$$

Proof in the Appendix.

Table 2: Specialization equilibria.

<i>Equilibrium</i>	<i>Segregation vs. Integration</i>	<i>Elasticity of substitution</i>	<i>Network membership</i>			
			<i>Exclusive</i>		<i>Inclusive</i>	
			<i>Minority</i> I_x	<i>Majority</i> J_x	<i>Minority</i> I_n	<i>Majority</i> J_n
EE	Segregation	Infinite	I	J	0	0
II	Integration	Infinite	0	0	I	J
IE	Segregation	Finite	0	J	I	0
EI	Segregation	Finite	I	0	0	J
IS	Segregation	Infinite	0	0	I	J

The most important insight of Proposition 4 is that there are stable equilibria, EI and IE, in which minority and majority people choose different types of skills. These equilibria substantiate

imperfect substitutability of minority and majority labor, that is, a finite ρ , and thus the existence of the substitution effect.³¹

To understand stability of specialization equilibria, one needs to recall that, due to asymmetric information, employers differentiate wages per unit of efficient labor across social groups but not within groups and that individuals cannot change their group membership. Therefore, individuals pick those networks to acquire skills that they are allowed to join and that offer the largest benefits from network effect (and thus are the most efficient for skill acquisition). In particular, this implies that deviation to an empty social network³² is never beneficial and for any individual it is preferable to be a member of the social network consisting of the other members of her social group.³³ This implies that the only possibly gainful and allowed deviation for an individual is that of switching to a non-empty inclusive social network of people from the other social group. In the IE equilibrium the only such possibility is a deviation of a majority individual to the inclusive network consisting of minority individuals. Such deviation would involve comparing the network benefits in the original social network, $N_{j,x}^{IE} = J^\gamma (\phi/(1+\phi))^\gamma$, to the network benefits obtained upon deviation, $N_{j,n}^{IE} = (1/(1+\delta)I)^\gamma (\phi/(1+\phi))^\gamma$. For this marginal deviation not to occur, the stability condition is $N_{j,x}^{IE} \geq N_{j,n}^{IE}$, which yields $J \geq I/(1+\delta)$. This condition is always satisfied. Consequently, the IE equilibrium is always stable. Similar arguments hold for the other specialization equilibria. The intuition behind the condition (16) is that in EI and IS equilibria minority individuals prefer segregated social interaction if and only if benefits from integration are low due to a relatively large size of the minority I or a large social distance δ .

³¹ Namely, $s_i^{EI} \rightarrow \infty$, $s_j^{EI} = 0$, $s_i^{IE} = 0$, and $s_j^{IE} \rightarrow \infty$. Therefore $s_i^{EI} \neq s_j^{EI}$ and $s_i^{IE} \neq s_j^{IE}$, which implies that ρ^{EI} and ρ^{IE} are finite. Clearly, $s_i^{EE} = s_j^{EE}$, $s_i^{II} = s_j^{II}$, and $s_i^{IS} = s_j^{IS}$ and thus $\rho^{EE} = \rho^{II} = \rho^{IS} \rightarrow \infty$. An example of minority specialization in a particular skill is the specialization of different Gypsy tribes in Romania in e.g. spoon-making (*Lingurari*), bear-leading, tinkering, and blacksmithing (*Ursari*), mining (*Rudari*), and goldsmithing (*Aurari*) (Fraser (1992)).

³² Deviating to an empty network involves setting it up. The implicit assumption of zero set-up costs of any network by any individual is a harmless simplification.

³³ See also Proposition 3 and its proof.

Relative Earnings under Specialization

In this section I turn to the particular pattern of earnings inequality as predicted by the model under specialization equilibria. Plugging the results from equations (12), (8), and (4), specifying the network effects according to (14) and (15), and taking the network sizes from Table 2, the following result for relative earnings ensues:

$$\omega^r(I) = \left(\frac{I}{1-I} \right)^{\frac{-1}{\rho}} \left(\frac{1 + I^\gamma (\phi/(1+\phi))^\gamma}{1 + (1-I)^\gamma (\phi/(1+\phi))^\gamma} \right)^{\frac{\rho-1}{\rho}}, \quad (17)$$

where the superscript r denotes the particular equilibrium; in equation (17) $r \in \{IE, EI\}$. Given this result, let us state one of the key propositions of this paper.

Proposition 5

Under the IE equilibrium for any $\rho > \rho' \equiv \left((2(\phi+1)/\phi)^\gamma + \gamma + 1 \right)^{-1}$ there always exists a range of relative minority size I such that the pattern of earnings inequality is consistent with the scale puzzle. There always exists sufficiently large social distance δ such that the same is true for the EI equilibrium. In particular, in such range $\partial \omega^r(I)/\partial I < 0$ and $\omega^r(I) < 1$, where $r \in \{EI, IE\}$.

Proof in the Appendix.

Thus, under the EI and IE equilibria, where minority and majority individuals acquire different skills and thus ρ is finite, the model predicts earnings inequality that is consistent with the scale puzzle for some range of I and large enough ρ . Intuitively, because ρ is finite in the EI and IE equilibria and hence the substitution effect is operative, there exists sufficiently small I for which minority labor is scarce enough to make minority earn more than majority.³⁴ On the other hand, a large enough ρ ensures that the substitution effect does not completely override the efficiency effect and there exist some I for which minority earns less than majority. For such finite and sufficiently large ρ , given the continuity of $\omega(I)$, there must be a downward sloping segment of $\omega(I)$ that is below one for some range of I . Such segment is congruent with the scale puzzle.

³⁴ Social distance δ has to be large enough to make such case under the EI equilibrium stable.

One can easily verify that under all specialization equilibria the efficiency effect favors relatively larger social groups and in particular that the relative supply of labor of minority individuals is increasing in their share in population, that is, $\partial h(I)/\partial I > 0$. To wit, under the EI, IE, EE, and IS

equilibria we obtain that $h(I) = \frac{1 + I^\gamma (\phi/(1+\phi))^\gamma}{1 + (1-I)^\gamma (\phi/(1+\phi))^\gamma}$, which is increasing in I , as follows from

the fact that the nominator is increasing in I and the denominator is decreasing in I . Similar result

holds for the II equilibrium, where $h(I) = \frac{1 + (I^\gamma + (1-I)^\gamma/(1+\delta))(\phi/(1+\phi))^\gamma}{1 + (I^\gamma/(1+\delta) + (1-I)^\gamma)(\phi/(1+\phi))^\gamma}$. Given that

$\partial h(I)/\partial I > 0$, the specialization equilibria EE, II, and IS predict that relatively larger minorities

earn relatively more than smaller ones, formally, $\partial \omega^r(I)/\partial I > 0$ for $r \in \{EE, II, IS\}$. To validate

this claim, knowing that $\partial h(I)/\partial I > 0$, one only needs to realize that ρ is infinity under these

equilibria and thus $\omega^r(I) = h^r(I)$ for $r \in \{EE, II, IS\}$.

3.4.2. Diversification

In the following sections I investigate whether the scale puzzle can be theoretically explained if the elasticity of substitution between skills ε is relatively small such that diversification arises.

As argued above, diversification equilibria arise in the equilibrium if and only if the two types of skills are complements or poor substitutes or there is sufficient degree of decreasing returns in

skill acquisition such that $\varepsilon < (\phi+1)/\phi$. Recalling that in diversification equilibria all agents

acquire both exclusive and inclusive skills, besides the optimality condition in equation (8), the

arbitrage condition in equation (11) is binding as well. Because all agents of a given type choose

the same set of networks and thus skills to acquire, as we know from Proposition 3, two different

equilibria can arise. In the DI equilibrium social groups acquire exclusive skills in their group-

specific social network and inclusive skills in one integrated inclusive social network where both

social groups interact. In the DS equilibrium, on the other hand, the inclusive skills are acquired

in two segregated minority- and majority- only inclusive social networks. In this sense the DI

equilibrium is integrated and the DS equilibrium segregated.³⁵ The following proposition

³⁵ Note, however, that there is a degree of segregation in the DI equilibrium as well, as the exclusive networks are by definition segregated.

discusses the stability of these equilibria; the intuition behind the condition (16) is the same as in Proposition 4 and concerns segregation of inclusive networks.

Proposition 6

The DI equilibrium of diversification is always stable. The DS equilibrium of diversification is stable if and only if the condition (16) holds.

Proof in the Appendix.

In the DI equilibrium, from Proposition 2 we know that $T_{k,x}^{DI} = t_k^{DI} / (1 + t_k^{DI})$ and $T_{k,n}^{DI} = 1 / (1 + t_k^{DI})$. For expositional convenience, I adopt here the network effects specifications $N_{i,m}(I_m, J_m, \delta) = (I_m + J_m / (1 + \delta))^\gamma$ and $N_{j,m}(I_m, J_m, \delta) = (I_m / (1 + \delta) + J_m)^\gamma$, assuming that network effects depend on the number of network members only. This network effect specification and the fact that all agents join all permissible networks under diversification such that $N_{i,x}^{DI} = I^\gamma$, $N_{i,n}^{DI} = (I + (1 - I) / (1 + \delta))^\gamma$, $N_{j,x}^{DI} = (1 - I)^\gamma$, and $N_{j,n}^{DI} = (I / (1 + \delta) + (1 - I))^\gamma$ result in the following specifications of relative times spent in any social networks:

$$t_i^{DI} = \left(\frac{1 + I^\gamma}{1 + (I + (1 - I) / (1 + \delta))^\gamma} \right)^{\frac{\varepsilon - 1}{1 + \phi - \varepsilon \phi}} \quad (18a)$$

and

$$t_j^{DI} = \left(\frac{1 + (1 - I)^\gamma}{1 + (I / (1 + \delta) + (1 - I))^\gamma} \right)^{\frac{\varepsilon - 1}{1 + \phi - \varepsilon \phi}}. \quad (18b)$$

Recalling that $I < J$, $\delta > 0$, and $\varepsilon < (\phi + 1) / \phi$ under the DI equilibrium, it is straightforward to observe that the ratios in the parenthesis in equations (18a-b) are less than one. For this reason, the results in equations (18a-b) reveal that all people spend more time in exclusive networks than in inclusive ones whenever $\varepsilon < 1$. This result arises as the consequence of skill complementarity that forces individuals to compensate for their lower efficiency in exclusive networks by the longer times spent in exclusive networks. Similarly, if $\varepsilon > 1$ and the DI equilibrium arises, all agents spend more time in inclusive networks. Noting from equation (8) that $t_k = l_k$ where $l_k \equiv L_{k,x} / L_{k,n}$, these results also hold for times spent on skill acquisition. Finally, if the

technology of combining skills is Cobb-Douglas and $\varepsilon=1$, individuals spend equal shares of their time in exclusive and inclusive networks.

Having computed the equilibrium time allocations for each social group, in this section I investigate differences in time allocations between social groups as they are closely related to the key question about substitutability of minority and majority labor. Proposition 7 states the first result in this respect:

Proposition 7

Under the DI equilibrium, minority individuals spend relatively more time in exclusive networks than majority individuals whenever $\varepsilon < 1$ such that complementarity of exclusive and inclusive skills prevails. Formally, $\varepsilon \gtrless 1$ implies $t_i^{DI} / t_j^{DI} = l_i^{DI} / l_j^{DI} \lesseqgtr 1$.

Proof in the Appendix.

These results stem from the relatively smaller efficiency of the exclusive networks of the minority as compared to those of the majority. As a result, if there is complementarity of skills such that $\varepsilon < 1$, as compared to the majority, minority individuals spend more time in their exclusive networks in order to compensate for this handicap. This finding reveals that the often-observed lesser involvement of minorities in formal educational institutions, as compared to the majority population, may be explained by inferior efficiency of social interaction in minority exclusive networks. The opposite result holds whenever skills are substitutes such that $\varepsilon < 1$ and diversification prevails. The same intuition governs the next proposition that states a much more important insight; namely, that minority and majority individuals choose different combinations of skills in the DI equilibrium.

Proposition 8

Social groups of different sizes choose different skill compositions in the DI equilibrium, in particular, $s_i^{DI} < s_j^{DI}$.

Proof in the Appendix.

Consequently, even though minority agents under some circumstances spend more time in their exclusive networks in the DI equilibrium, they unambiguously acquire relatively *less* exclusive skills than majority individuals. The key result here is that skill composition is different across social groups under the DI equilibrium of diversification and, therefore, the elasticity of substitution between labor of minority and majority individuals is finite.

Let us now turn to the DS diversification equilibrium under which minority and majority individuals join two disconnected inclusive networks as well as their exclusive networks. Under this equilibrium inclusive and exclusive networks provide the same network effect benefits for their members, as they are of the same size and composition. It follows that people distribute their time evenly between exclusive and inclusive networks and, as a consequence, have the same shares of exclusive and inclusive skills in the labor they supply. Therefore, their labor is perfectly substitutable on the labor market and they earn the same wage per efficiency unit of their labor.

Relative Earnings under Diversification

In this section I investigate relative earnings under the two diversification equilibria. Plugging the relative supply of labor (13) into (4), the relative income under the DI equilibrium is

$$\omega^{DI}(I) = \left(\frac{I}{1-I} \right)^{\frac{-1}{\rho}} \left(\frac{\left(\frac{1}{1+t_i^{DI}} \right)^{\phi+1} \left(\left(t_i^{DI} \right)^{\phi+1} \left(1 + N_{i,x}^{DI} \right)^{(\varepsilon-1)/\varepsilon} + \left(1 + N_{i,n}^{DI} \right)^{(\varepsilon-1)/\varepsilon} \right)^{\varepsilon/(\varepsilon-1)}}{\left(\frac{1}{1+t_j^{DI}} \right)^{\phi+1} \left(\left(t_j^{DI} \right)^{\phi+1} \left(1 + N_{j,x}^{DI} \right)^{(\varepsilon-1)/\varepsilon} + \left(1 + N_{j,n}^{DI} \right)^{(\varepsilon-1)/\varepsilon} \right)^{\varepsilon/(\varepsilon-1)}} \right)^{\frac{\rho-1}{\rho}} \quad (19)$$

where the respective $N_{k,m}^{DI}$ and t_k^{DI} are specified above.

Proposition 9

Under the DI diversification equilibrium of human capital acquisition there always exists ρ'' such that for any $\rho > \rho''$ there always exists a range of relative minority size I such that the pattern of earnings inequality is consistent with the scale puzzle. In particular, in this range $\partial \omega^{DI}(I) / \partial I < 0$ and $\omega^{DI}(I) < 1$.

Proof in the Appendix.

Proposition 9 shows that under the DI diversification equilibrium, in which minority and majority individuals acquire different skills and thus ρ is finite, the model predicts earnings inequality that is consistent with the scale puzzle for some range of I and large enough ρ . The intuition is similar to that of Proposition 5 for the specialization equilibria EI and IE. Similarly as in the EI and IE equilibria, small enough minorities outperform majorities in terms of earnings.

Under the DS equilibrium, however, minority and majority individuals, facing identical efficiencies in completely segregated social networks, spend equal times in each type of networks and acquire the same composition of skills, $s_i^{DS} = s_j^{DS} = 1$. Therefore, the elasticity of substitution of their labor ρ is infinite. Using this result and the facts that individuals divide their time evenly between acquisition and utilization of skills and that network effects are I^γ for minority individuals and $(1-I)^\gamma$ for majority individuals in any network they join, we obtain that

$$\omega^{DS}(I) = \frac{1 + I^\gamma}{1 + (1 - I)^\gamma}. \quad (20)$$

Clearly, $\omega^{DS}(I) = \frac{1 + I^\gamma}{1 + (1 - I)^\gamma} < 1$, or, in plain words, under the DS equilibrium minority individuals are always poorer than majority individuals. As it is easy to see, in conflict with the scale puzzle, equation (20) predicts that minority-majority earnings gap is decreasing in minority relative size.

4. Discussion

4.1. The Roles of Integration and Exclusion

In the presented model integration has a distinct role in determining relative income of minorities that challenges the habitual belief that integration leads to greater equality between social groups. While it is true that both minority and majority individuals benefit from integration through the increased network effects that integration brings about, integration disfavors minority individuals whenever it obliterates the substitution effect, which favors smaller social groups. If the obliteration of the substitution effect offsets the efficiency benefits of integration, integration

decreases the relative income of minority individuals.³⁶ Proposition 10 below states that this is possible in the case of integration from the EI or IE equilibrium into the II equilibrium and that, in particular, integration may increase inequality.³⁷ It is worthwhile to note that, in contrast to the specialization equilibria, integration produces imperfect substitutability under diversification equilibria and thus benefits minorities in terms of both efficiency and substitution effects.

Proposition 10

There exists I such that integration from the EI or IE equilibrium into the II equilibrium hurts minority individuals in terms of relative earnings, ω , that is, $\omega^r(I) > \omega^II(I)$ for $r \in \{EI, IE\}$. Whenever $\rho > \rho'$, there exists I for which integration increases earnings inequality as well.

Proof in the Appendix.

Exclusion in exclusive networks has insofar been accepted in the model as an exogenous institutional constraint on agents' choice. Although it is fully symmetric across social groups, it is sensible to put this constraint under scrutiny, as it may prevent agents from individually benefiting from integration.³⁸ In particular, it is informative to answer the question whether the explanation of the scale puzzle developed in this paper remains valid, or, in other words, whether the EI, IE and DI equilibria remain stable, if members of exclusive networks permit inclusion of individuals from the other social group. This is the case, if the excluded agents individually choose not to join exclusive networks of the other social group even if the institutional barriers to integrate are removed. From the proof of Proposition 4 it is clear that this holds whenever minority share in population or social distance is large enough such that the inequality (16) is satisfied.³⁹ It follows that exclusive behavior is not necessary to establish the main results of this

³⁶ Minority individuals benefit from integration relatively more than majority individuals do, as they gain access to social interaction with the larger pool of majority individuals as compared to the access to the smaller group of minority individuals gained by majority individuals.

³⁷ Forced integration can reduce minority earnings in absolute terms. Intuitively, this is the case whenever social distance is so large as to cause the efficiency benefits of integration to be smaller than its costs in terms of the substitution effect, e.g. when δ is relatively large or ρ relatively small.

³⁸ Here I consider marginal deviation from the equilibrium, in which the deviating minority individual can affect her efficiency of skill acquisition (network effects) but not her wage per unit of efficient labor.

³⁹ Note that marginal deviation to the minority exclusive network would never be beneficial for a majority individual.

study. In particular, equilibria in which segregation and thus skill differentiation across social groups arise are possible without institutional exclusion, giving rise to patterns of income inequality consistent with the scale puzzle as explained above. This argument generalizes the argument of the paper to societies without institutional exclusion.

4.2. Welfare

While the analysis thus far has focused on relative welfare⁴⁰ of minority and majority individuals, the various equilibria that the model generates can be welfare ranked according to aggregate consumption C . Given that whether specialization or diversification occurs fully depends on parameters in the condition $\varepsilon \geq (\phi+1)/\phi$, the policy maker cannot choose between diversification and specialization regimes. Therefore, from the policy perspective it is sensible to compare the efficiency of equilibria within rather than between these regimes.

Proposition 11

Under the specialization regime, the EI and IE equilibria and the EE and IS equilibria are equally efficient. The EI, IE, and II equilibria are more efficient than the EE and IS equilibria. Formally, $C^{EI} = C^{IE}$, $C^{EE} = C^{IS}$, and $C^r > C^{r'}$ for $r \in \{EI, IE, II\}$ and $r' \in \{EE, IS\}$. Under the diversification regime, the DI equilibrium is more efficient than the DS equilibrium, that is, $C^{DI} > C^{DS}$.

These results are intuitive and formal proofs are omitted: $C^{EI} = C^{IE}$ and $C^{EE} = C^{IS}$ because all the parameters and inputs in (1) are the same under the respective equilibria. The II equilibrium is more efficient than the EE and IS equilibria, because (i) it generates larger inputs $H_k^{II} > H_k^r$ for any k and $r \in \{EE, IS\}$ as a consequence of larger network benefits under integration and (ii) production is linear in each of these equilibria. While inputs are the same under EI, IE, EE, and IS equilibria, specialization into different skills under the EI and IE equilibria generates extra surplus in the CES production function (1).⁴¹ The same effect favors the DI equilibrium over the

⁴⁰ Individuals have no income but earnings and consume the same consumption good. Therefore their welfare equals their earnings.

⁴¹ It is a well-known property of the CES function that it exhibits returns to diversity of inputs, for given input levels. These returns are decreasing in ρ for $\rho \in (1, \infty)$.

DS equilibrium; furthermore, under the DS equilibrium segregation reduces network benefits and thus diminishes labor inputs in production.

The remaining issue concerns the comparison between the EI and IE equilibria on the one hand and the II equilibrium on the other hand. Clearly, this comparison involves weighing benefits of higher efficiency of skill acquisition under integration against the benefits of diversity of inputs in production under segregation. The two key parameters are thus the elasticity of substitution ρ and social distance δ . Intuitively, whenever ρ is very large under the EI and IE equilibria, the benefits of diversity are negligible and thus the II equilibrium is more efficient than the EI and IE equilibria. On the other hand, if δ is very large, integration offers no efficiency benefits and the EI and IE equilibria are preferable to the II equilibrium.

4.3. The Multiplicity of Equilibria

The model presented in this paper classifies equilibrium regimes of human capital acquisition and elucidates the conditions under which they are stable. Moreover, it pinpoints those equilibria under which minority and majority people acquire different human capitals and identifies the conditions under which they give rise to patterns of earnings inequality consistent with the scale puzzle, which are explicated in Propositions 5 and 9. Table 3 below summarizes the equilibria of the model and highlights (bold-typed) those that are reconcilable with the scale puzzle.

Table 3: Equilibrium regimes of skill acquisition.

<i>Regime type</i>	<i>Stability of equilibria</i>	
	$I < 1/(2 + \delta)$	$I \geq 1/(2 + \delta)$
$\varepsilon \geq (\phi + 1)/\phi$	EE, II, IE	EE, II, IS, IE, EI
$\varepsilon < (\phi + 1)/\phi$	DI	DI, DS

Table 3 shows that the elasticity of substitution between skills in production of efficiency units of labor ε , the degree of decreasing returns to scale in skill acquisition ϕ , social distance between minority and majority δ , and minority share in the population I determine which equilibria may arise in the equilibrium. In particular, the parameters ε and ϕ determine whether

specialization or diversification prevails, as depicted in Proposition 2. The other two parameters, δ and I , determine whether segregated equilibria EI, IS, and DS are stable. Notably, Table 3 shows that the conditions $I < 1/(2 + \delta)$, $\varepsilon < (\phi + 1)/\phi$, and $\rho > \rho''$ are sufficient for the patterns of earnings inequality to be consistent with the scale puzzle in some range of I . Whenever either of the conditions $I < 1/(2 + \delta)$ or $\varepsilon < (\phi + 1)/\phi$ is not satisfied, equilibria EE, II, IS, and DS, which are not reconcilable with the scale puzzle, are possible. Besides these constraints, which equilibrium occurs in any particular case is indeterminate in the model.

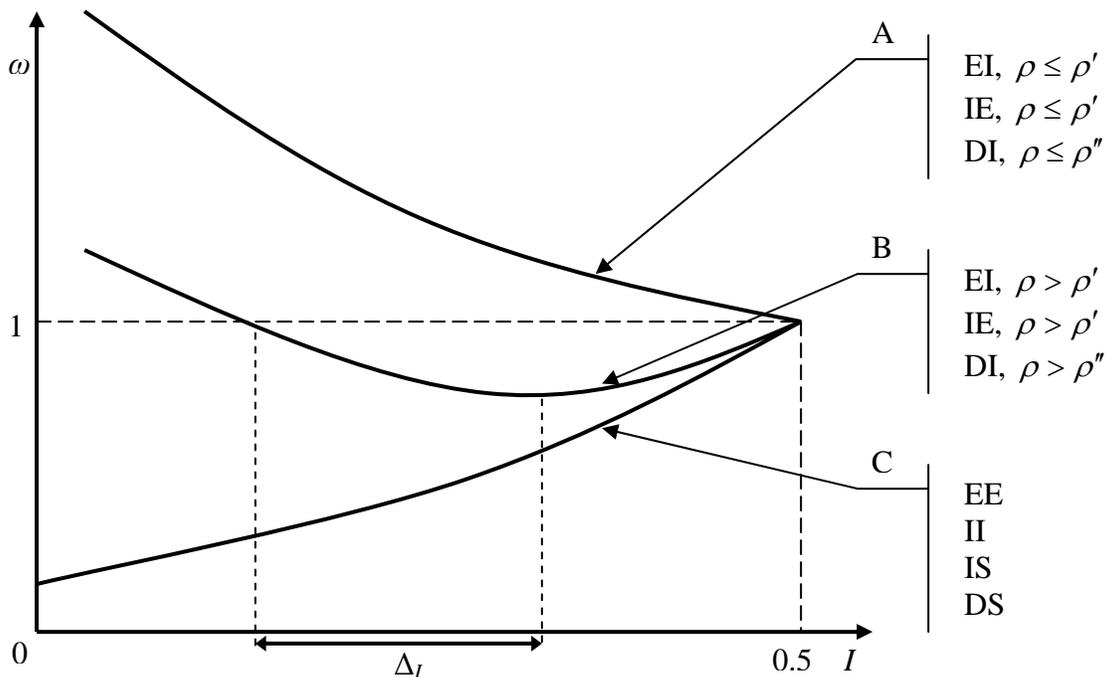
This indeterminacy is due to the absence of parametric restriction in the specification of the model, which was intentionally imposed in an attempt to foster its generality. There is, however, an intuitive parametric restriction that lends itself to elimination of some of the equilibria from consideration. In particular, recalling that exclusive skills typically involve those acquired in families, kinships, and other noninstitutional social networks and inclusive skills are predominantly acquired in schools, universities, and workplaces, one can reasonably argue that exclusive and inclusive skills exhibit complementarity such that they both are involved in individual human capital. Indeed, Coleman et al. (1966) argue that students' learning outcome is a function of family and school inputs. Heckman (2000) points out that skills acquired from informal, noninstitutional, sources such as families complement skills acquired in school and other formal institutions and thus determine success in life. If this is the case, one can impose the condition $\varepsilon < (\phi + 1)/\phi$, implying that the model only supports diversification equilibria DI and DS. There are two reasons to believe that the DI equilibrium is somewhat more representative of these two equilibria. First, the DS equilibrium is not robust with respect to coordination of minority individuals to join the inclusive network of majority agents.⁴² Second, in Western economies, where efforts are made to eradicate segregation in formal educational institutions, segregation in these institutions is less likely.

Figure 1 depicts the three stylized patterns of minority-majority earnings inequality as a function of minority percentage that the model generates. Pattern C represents the EE, II, IS, and DS equilibria under which only the efficiency effect is effective. Because the efficiency effect favors

⁴² This straightforwardly follows from the proof of Proposition 6. Note that for similar reasons the equilibria IS and EI are not robust with respect to such coordination; see the proof of Proposition 4.

larger social groups, pattern C is below one for any $I < 0.5$ and upward sloping. On the other hand, if any of the EI, IE, or, notably, DI equilibrium arises, the substitution effect that favors smaller social groups becomes operative. Pattern A depicts the case when ρ is relatively small such that the substitution effect overrides the efficiency effect to make the minority earn more than the majority for any minority size. The most interesting in the context of the scale puzzle, however, is pattern B. It depicts the intermediate case where the substitution effect works in favor of smaller minorities but does not completely outweigh minorities' efficiency disadvantage such that a range of I where the scale puzzle is replicated by the model, Δ_I , arises. This is so whenever ρ is large enough, as specified in Propositions 5 and 9.

Figure 1: Stylized patterns of minority-majority earnings inequality.



Given that the DI equilibrium is more representative than the DS equilibrium, the remaining question is whether the elasticity of substitution between minority and majority labor ρ is high enough such that a plausible segment Δ_I arises in the model under this equilibrium. To investigate this issue, I tentatively calibrate the model developed above. For this purpose, I let the key parameter ρ attain the value of 25, as estimated by Kahanec (2006). As concerns the remaining parameters, I assume that $\varepsilon = 0.9$, implying a degree of complementarity between exclusive and inclusive skills, $\delta = 0.25$, such that individual's benefit from interaction with a

socially distant individual is 80% of that with an akin individual, and $\phi = 1 = \gamma$, such that skill acquisition exhibits constant returns to time and social interaction.⁴³ With these parametric values the model predicts the segment Δ_I between 0.01% and 5.2%. This range of I covers three quarters of US counties with some Black population.⁴⁴ As concerns minority majority earnings gap, under the aforementioned parametric values the ratio of minority to majority earnings attains the minimal value of 0.8 in the labor market with 5.2% of minority people. In other words, in this labor market minority individuals earn about 20% less than majority individuals. These values correspond to the values of minority-majority earnings gaps found by Kahanec (2006). This exercise illustrates that plausible parametric values give rise to plausible predictions as concerns minority majority earnings differential in the model developed above.

5. Conclusion

This paper provides a novel explanation of the persistent patterns of income inequality between minorities and majorities. In particular, it extends the local effects literature by introducing heterogeneity of social networks and skills into a model with external network effects and minority-majority social distance in human capital acquisition. I establish that such extension provides a theoretical underpinning for the scale puzzle, reconciling the local effects approach with this empirical phenomenon. In particular, I first explicate how network effects and social distances in skill acquisition engender the efficiency effect, directly favoring the members of larger social groups. Next, I establish that introducing heterogeneity of available skills into this setup gives rise to equilibria in which minority and majority individuals, driven by network effects, choose different skills or combinations of skills to acquire, thus supplying imperfectly substitutable labor on the labor market. As a consequence, the substitution effect emerges and, through prices of minority and majority labor, favors smaller social groups, which on aggregate supply less labor measured in efficiency units, *ceteris paribus*. The conditions under which the efficiency and substitution effects explain the scale puzzle are then established and discussed.

⁴³ Constant returns are not necessary here, letting $\phi = 0.5$, $\gamma = 0.2$, and holding the other parameters as above, for instance, results in a plausible range of Δ_I between 1.1% and 9.5%.

⁴⁴ Author's computation, based on Census School District Tabulation Data, 2000, NCES.

The last section summarizes and classifies the outcomes of the model, offering a tool to address the various modes of minority-majority interaction observed in the reality.

Highlighting the role of segregation, the paper discusses the consequences of segregation on the efficiency of skill acquisition and relative wages. It is also shown that elimination of institutional exclusion in social networks does not necessarily lead to integration whenever the size of minority or the social distance between social groups is large enough. An interesting result of the model is that integration may increase as well as decrease the relative income of minority as compared to majority individuals. In this sense, although there are efficiency benefits of integration for both minority and majority individuals, integration is not necessarily a universal remedy against inequality between social groups. Similarly, integration may increase as well as decrease overall welfare. The model also predicts that under some conditions for a small minority the substitution effect outweighs minority's efficiency disadvantage such that minority earnings are higher than majority earnings. No discrimination on the labor market is necessary to obtain the results of this paper, which makes it an alternative to the existing discrimination-based theories of earnings inequality. Future research in this area should include empirical tests and case studies, as well as investigation of the link between organization of social interaction and, first, geographical location and, second, the dynamics of social distance.

References

- Allen, Beth**, "Some Stochastic Processes of Interdependent Demand and Technological Diffusion of an Innovation Exhibiting Externalities among Adopters," *International Economic Review*, Vol. 23, (1982), pp. 595-607.
- Aigner, Dennis and Glen Cain**, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, Vol. 30, (January, 1977), 175-87.
- Akerlof, George A.**, "Social Distance and Social Decisions," *Econometrica*, Vol. 65, No. 5. (Sep., 1997), pp. 1005-1027.
- Allport, G. W.**, *The nature of prejudice*. Reading, MA: Addison-Wesley, (1954).
- Altonji, J. and R. Blank**, "Race and Gender in the Labor Market", in eds. O. Ashenfelter and D. Card, *Handbook of Labor Economics*, Volume 3c, (1998).
- Arrow, Kenneth J.**, "Models of Job Discrimination," in A. H. Pascal, ed. *Racial Discrimination in Economic Life*, Lexington, Mass: D.C. Heath, (1972a), pp. 83-102
- Arrow, Kenneth J.**, "Some Mathematical Models of Race Discrimination in the Labor Market," in A. H. Pascal, ed. *Racial Discrimination in Economic Life*, Lexington, Mass: D.C. Heath, (1972b), pp. 187-204.
- Arrow, Kenneth J.**, "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds. *Discrimination in the Labor Markets*, Princeton, N.J.: Princeton University Press, (1973), pp. 3-33.
- Arrow, Kenneth J.**, "What Has Economics to Say about Racial Discrimination?" (Symposium: Discrimination in Product, Credit and Labor Markets), *The Journal of Economic Perspectives*, Vol. 12, No. 2. (Spring, 1998), pp. 91-100.
- Bala, Vanketash and Sanjeev Goyal**, "Learning from Neighbors," *Review of Economic Studies*, Vol. 65, (1998), pp. 595-621.
- Becker, Gary**, *The Economics of Discrimination*, Chicago: University of Chicago Press, (1957).
- Becker, Gary**, "Investment in Human Capital: A Theoretical Analysis," *The Journal of Political Economy*, Vol. 70, No. 5, Part 2: Investment in Human Beings, (Oct., 1962), pp. 9-49.
- Becker, Gary S.**, *The Economics of Discrimination*, 2nd edition, The University of Chicago Press, Chicago, IL, (1971).
- Becker, Gary S. and Nigel Tomes**, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, Vol. 87, (1979), pp. 1153-1189.
- Bell, W.**, "A Probability Model for the Measurement of Ecological Segregation," *Social Forces*, Vol. 32, (1954), pp. 357-64.
- Benabou, Roland**, "Workings of a City: Location, Education, and Production," *The Quarterly Journal of Economics*, Vol. 108, No. 3, (Aug. 1993), pp. 619-652.
- Benabou, Roland**, "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance" *The American Economic Review*, Vol. 86, No. 3. (Jun., 1996), pp. 584-609.
- Blalock, H.M., Jr.**, "Economic Discrimination and Negro Increase," *American Sociological Review*, Vol. 21, No. 5. (Oct., 1956), pp. 584-588.
- Blalock, H.M., Jr.**, "Per Cent Non-White and Discrimination in the South," *American Sociological Review*, Vol. 22, No. 6. (Dec., 1957), pp. 677-682.
- Blalock, H.M., Jr.**, "Status Inconsistency, Social Mobility, Status Integration and Structural Effects," *American Sociological Review*, Vol. 32, No. 5. (Oct., 1967), pp. 790-801.

Bonacich, Edna, "A Theory of Ethnic Antagonism: The Split Labor Market," *American Sociological Review*, Vol. 37, No. 5. (Oct., 1972), pp. 547-559.

Bonacich, Edna, "Advanced Capitalism and Black/White Race Relations in the United States: A Split Labor Market Interpretation," *American Sociological Review*, Vol. 41, No. 1. (Feb., 1976), pp. 34-51.

Borjas, George J., "The Substitutability of Black, Hispanic, and White Labor," *Economic Inquiry*, Vol. 21, (Jan. 1983), pp. 93-106.

Borjas, George J., "Immigrants, Minorities, and Labor Market Competition," *Industrial and Labor Relations Review*, Vol. 40, No. 3, (Apr., 1987), pp. 382-392.

Bourdieu, Pierre, "Forms of Capital," in *Handbook of Theory and Research for the Sociology of Education*, John G. Richardson, ed. Westport, CT: Greenwood Press, (1986), pp. 241-60.

Brown, David L. and Glenn V. Fuguitt, "Percent Nonwhite and Racial Disparity in Nonmetropolitan Cities of the South," *Social Science Quarterly*, Vol. 53 (December, 1972), pp. 573-82.

Chiswick, Barry R., and Paul W. Miller, "Do Enclaves Matter in Immigrant Adjustment?" *City and Community*, 4(1), (2005)

Coate, Stephen and Glenn Loury, "Will Affirmative Action Policies Eliminate Negative Stereotypes," *American Economic Review*, Vol. 83, (1993), pp. 1220-42.

Coleman J. S., E. Campbell, C. Hubson, J. McPartland, A. Mood, F. Weinfeld, and R. York, *Equality of Educational Opportunity*, Washington D.C.: U.S. Government Printing Office, (1966)

Coleman, J. S., "Social capital in the creation of human capital," *American Journal of Sociology*, Vol. 94, (1988), pp. S95-S120.

Coleman, J.S., *Foundations of Social Theory*. Cambridge, MA: Harvard University Press, (1990).

Conley, Timothy and Christopher Udry, "Learning about a New Technology: Pineapple in Ghana," mimeo, Yale University, (March, 2002).

Darity, William A., Jr., "The Human Capital Approach to Black-White Earnings Inequality: Some Unsettled Questions," *The Journal of Human Resources*, Vol. 17, No. 1, (Winter, 1982), pp. 72-93.

Durlauf, Steven, "Spillovers, Stratification, and Inequality," *European Economic Review*, Vol. 38, (1994), pp. 836-845.

Durlauf, Steven, "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, Vol. 1, (1996), pp. 75-93.

Ellison, Glenn and Drew Fudenberg, "Rules of Thumb for Social Learning," *Journal of Political Economy*, Vol. 101, (1993), pp. 93-125.

Ellison, Glenn and Drew Fudenberg, "Word-of-Mouth Communication and Social Learning," *Quarterly Journal of Economics*, Vol. 110, No. 1, (February, 1995), pp. 83-97.

Farley, Reynolds and William H. Frey, "Changes in the Segregation of Whites from Blacks During the 1980s: Small Steps toward a More Racially Integrated Society," *American Sociological Review*, Vol. 59, (1994), pp. 23-45.

Feick, Lawrence F. and Linda L. Price, "The Market Maven: A Diffuser of Marketplace Information," *Journal of Marketing*, Vol. 51, No. 1, (January, 1987), pp. 83-97.

Foster, Andrew D. and Mark R. Rosenzweig, "Learning by Doing and Learning from Others: Human Capital and Technical change in Agriculture," *Journal of Political Economy*, Vol. 103, No. 6, (1995), pp. 1176-1209.

- Fraser, Angus**, *The Gypsies*, Blackwell, Oxford, (1992).
- Frisbie, W. Parker and Lisa Neidert**, "Inequality and the Relative Size of Minority Populations: A Comparative Analysis," *The American Journal of Sociology*, Vol. 82, No. 5. (Mar., 1977), pp. 1007-1030.
- Gladwell, Malcolm**, *The tipping point: how little things can make a big difference*, Boston: Little, Brown, (2000).
- Glaeser, Edward L. and Jacob L. Vigdor**, "Racial Segregation in the 2002 Census: Promising News," The Brookings Institution, Center on Urban and Metropolitan Policy, Survey Series, (April, 2001).
- Glaeser, Edward L., David Laibson, and Bruce Sacerdote**, "An Economic Approach to Social Capital," *The Economic Journal*, Vol. 112, (November, 2002), pp. 437-458.
- Glenn, N.**, "The relative Size of the Negro Population and Negro Occupational Status," *Social Forces*, Vol. 43, (1964), pp. 42-49.
- Goyal, Sanjeev**, "Learning in Networks: A Survey," in *Group Formation in Economics: Networks, Clubs, and Coalitions*, Eds. G. Demange and M. Wooders, Cambridge University Press, (2003).
- Grant, James H. and Daniel Hamermesh**, "Labor Market Competition among Youths, White Women, and Others," *The Review of Economics and Statistics*, Vol. 63, No. 3, (Aug. 1981), pp. 354-360.
- Grossman, Jean Baldwin**, "The Substitutability of Natives and Immigrants in Production," *The Review of Economics and Statistics*, Vol. 64, No. 4, (Nov. 1982), pp. 596-603.
- Heckman, James J.**, "Policies to Foster human Capital," *Research in Economics*, Vol. 54, (2000), pp. 3-56
- Heer, David M.**, "The Sentiment of White Supremacy and Ecological Study," *The American Journal of Sociology*, Vol. 64, No. 6, (May, 1959), pp. 592-598.
- Hirschman, Charles and Morrison G. Wong**, "Socioeconomic Gains of Asian Americans, Blacks, and Hispanics: 1960-1976," *The American Journal of Sociology*, Vol. 90, No. 3, (Nov. 1984), pp. 584-607.
- Kahanec, Martin**, "The Substitutability of Labor of Selected Ethnic Groups in the US Labor Market" IZA Discussion Paper No. 1945, IZA, Bonn, (January, 2006).
- Lazear, Edward P.**, "Culture and Change," *The Journal of Political Economy*, Vol. 107, No. 6, (1999), pp. 95-126.
- Loury, Glenn C.**, "A dynamic theory of racial income differences." pp. 153-186 in P.A. Wallace and A. Le Mund, eds., *Women, minorities, and employment discrimination*. Lexington, MA: Lexington Books, (1977).
- Loury, Glenn C.**, "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, Vol. 49, No. 4. (Jul., 1981), pp. 843-867.
- Loury, Glenn C.**, "Discrimination in the Post-Civil Rights Era: Beyond Market Interactions," (Symposium: Discrimination in Product, Credit and Labor Markets), *The Journal of Economic Perspectives*, Vol. 12, No. 2. (Spring, 1998), pp. 117-126.
- Lucas, Robert E., Jr.**, "On the Mechanics of Economic Development," *Journal of Monetary Economics*, Vol. 22, (1988), pp. 3-42.
- Lundberg, Shelly and Richard Startz**, "On the Persistence of Racial Inequality," *Journal of Labor Economics*, Vol. 16, No. 2. (Apr., 1998), pp. 292-323.
- Lundberg, Shelly and Richard Startz**, "Race, Information, and Segregation," Working Paper, Department of Economics, University of Washington, (2002).

- Massey, Douglas S. and Nancy A. Denton**, *American Apartheid: Segregation and the Making of the Underclass*, Cambridge, MA.: Harvard University, (1993).
- Mincer, Jacob**, "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, Vol. 66, (August, 1958), pp. 281-302.
- Munshi, Kaivan**, "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics*, Vol. 73(1), (2004), pp. 185-215.
- Phelps, Edmund S.**, "The Statistical Theory of Racism and Sexism," *American Economic Review*, Vol. 62, (1972), pp. 659-61.
- Poole, Willard C., Jr.**, "Distance in Sociology," *The American Journal of Sociology*, Vol. 33, No. 1. (Jul., 1927), pp. 99-104.
- Reardon, Sean F., John T. Yun, and Tamela McNulty Eitle**, "The Changing Structure of School Segregation: Measurement and Evidence of Multiracial Metropolitan-Area School Segregation, 1989-1995," *Demography*, Vol. 37, No. 3, (August, 2000), pp. 351-364.
- Rees, Albert and George P. Schultz**, "Workers and Wages in an Urban Labor Market," *Urban Studies*, Vol. 9, No. 1 (February, 1972), pp. 130-132.
- Reich, Michael**, "The Economics of Racism," in David Gordon, ed., *Problems in Political Economy*. D.I.C Heath, (1971).
- Schultz, Theodore W.**, "Investment in Human Capital," *American Economic Review*, Vol. 51, (March, 1961), pp. 1-17.
- Semyonov, Moshe, Danny R. Hoyt and Richard Ira Scott**, "Place, Race and Differential Occupational Opportunities," *Demography* Vol. 21, 2, (1984), pp. 259-270.
- Spilerman, Seymour and Richard E. Miller**, "City Nondifferences Revisited," (Comment), *American Sociological Review*, Vol. 42, No. 6. (Dec., 1977), pp. 979-983.
- Steele, Claude M.**, "Race and the Schooling of Black Americans," *The Atlantic Monthly*, Vol. 269, No. 4, (April, 1992), pp. 67-78.
- Tienda, Marta and Ding-Tzann Lii**, "Minority Concentration and Earnings Inequality: Blacks, Hispanics, and Asians Compared," *The American Journal of Sociology*, Vol. 93, No. 1, (Jul. 1987), pp. 141-165.
- Valente, Thomas W.**, *Network Models of the Diffusion of Innovations*, Cresskill NJ: Hampton Press, Inc., (1995).
- Welch, Finis**, "Labor Market Discrimination: An Interpretation of Income Differences in the Rural South," *Journal of Political Economy*, Vol. 75, (1967), pp. 225-240.
- Williams, R. M., Jr.**, *The reduction of intergroup tensions*. New York: Social Science Research Council, (1947).

Appendix

Table A.1: The scale puzzle – empirical evidence.

Study	Data	Main Findings
Blalock (1956)	88 non-Southern and Southern Standard Metropolitan Areas (SMAs), (1950)	Finds a positive correlation .42 between percent Black population and Black-White income differentials. Controlling for subregion, white median income, size of SMA, and percent of employed males in manufacturing, the correlation was reduced to nonsignificant .19. However, for southern SMAs the correlation was .50 and increased to .70 when the same controls were included. When both Southern and non-Southern SMAs were included, the correlation was .67 irrespective of the controls. Suggests that the marginal Black relative losses due to percent increase are decreasing (non-linearity) in percent Black. Suspects a threshold at about 10% above which correlations significant.
Blalock (1957)	Sample of 150 Southern US counties having at least 250 non-white households, (1950)	Finds a positive correlation between percent Black population and income (.46) and educational (.68) differentials. The findings were robust with respect to the same controls as in Blalock (1956). Finds that in counties with low Black percentage (but not in those with high Black percentage where all results were non-significant) the income and educational gaps are disproportionately smaller, thereby supporting the non-linearity hypothesis for the low end of the density continuum. No relationship is observed for occupational differentials.
Heer (1959)	43 Southern Standard Metropolitan Areas, Census, (1950)	Finds negative correlation of -.71 between percent Black and the ratio of Black median income to White median income.
Brown and Fuguitt (1972)	878 non-metropolitan areas, Southern US, PH-5 census, (1960)	Report overrepresentation of the majority in all higher income groups (difference scores range between 31% and 45% when cumulative distributions are compared). Find that the association between percent Black and measures of racial income disparity is positive and ranges between .16 and .41. Moreover, they show that Black income decreases and White income increases with increases in percent Black. White component correlations range between .12 and .21. Black counterparts range between -.09 and -.31.
Frisbie and Neidert (1977)	40 Standard Metropolitan Areas in southwestern U.S. U.S. Bureau of the Census, (1971, 1972)	Report overrepresentation of the majority in all higher income groups (difference scores range between 7% and 20% (Mexican) and between 18% and 31% (Black) when cumulative distributions are compared). Finds correlations of .22 to .48 between percent Mexican and the Mexican-Anglo income differential for different income groups. The corresponding values for the Blacks range between 0.31 and 0.43. Controls included: % labor force in manufacturing, % labor force in services, Black median education, 1960-1970 % change in Black population, % Mexican. Uncontrolled correlations similar. Confirms that the majority income is positively correlated with percent minority (correlations between .06 and .21) and that minority income is negatively correlated with percent minority (correlations between -.48 and -.22).
Tienda and Lii (1987)	5% A File, men, Public Use Microdata Samples, Census (1980)	Confirms that minorities have lower income than majorities and that minorities lose from increases in their percentages, while the white majority gains. Blacks, Hispanics, and Asians lose 0.7%, 0.2%, and \$0.2% of their annual income, respectively, and the majority gains between 0.0-0.5% with every percentage increase of the respective minority density. These results are net of some observable individual characteristics and working time measures. Minority losses from their percentages most pronounced for educated minority people.
Borjas (1987)	5% A File, 18-64 years old non-military individuals working with pay, Public Use Microdata Samples, Census (1980)	Shows that while native populations' earnings are little decreased by inflow of immigrants of the same ethnicity (e.g. 10% increase in the supply of White immigrants decreases the pay of White natives by 2.5%), this inflow substantially reduces the earnings of other immigrants of the same ethnicity (10% increase in the number of White immigrants reduces the pay of white immigrants by 10.9%, the same increase of Black immigration reduces the pay of Black immigrants by 5.8%).
Chiswick and Miller, (2005)	US Census of Population (1990)	Establish that earnings of immigrants of a given linguistic group decrease in the share of people of the same linguistic group in the destination region.

Proof of Time Allocation Rule (8)

Take the allocation of time $\{T_{k,x}, T_{k,n}\}$ as given and rewrite the agents' problem (7) as follows:

$$\text{Max}_{L_{k,m}} |H_k|$$

s.t.

$$H_k = \left((S_{k,x} (T_{k,x} - L_{k,x}))^{(\varepsilon-1)/\varepsilon} + (S_{k,n} (T_{k,n} - L_{k,n}))^{(\varepsilon-1)/\varepsilon} \right)^{\varepsilon/(\varepsilon-1)}$$

$$S_{k,m} = L_{k,m}^\phi (1 + N_{k,m})$$

$$L_{k,m} \geq 0$$

Plugging the technological constraints into the objective function and deriving the first order conditions with respect to $L_{k,m}$, the optimality conditions on time distribution between acquisition

and utilization of skills are $L_{k,m} = \frac{\phi}{1+\phi} T_{k,m}$. The sufficiency conditions are also satisfied, as the

objective function is concave at the optimal allocation. ■

Proof of Proposition 1

Substitute for $L_{k,m}$ in the agent's problem (7) using (8). In addition, substitute for $S_{k,m}$ from the skill acquisition technology and use the definition of $\tilde{N}_{k,m}$. Consequently, the agent's problem is:

$$\text{Max}_{T_{k,x}, T_{k,n}} \left| \left(\tilde{N}_{k,x} T_{k,x}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} + \tilde{N}_{k,n} T_{k,n}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} \right)^{\varepsilon/(\varepsilon-1)} \right|$$

s.t.

$$T_{k,x} + T_{k,n} \leq 1$$

$$T_{k,m} \geq 0$$

Note that if $\varepsilon = (\phi + 1)/\phi > 1$ the problem boils down to

$$\text{Max}_{T_{k,x}, T_{k,n}} \left| \left(\tilde{N}_{k,x} T_{k,x} + \tilde{N}_{k,n} T_{k,n} \right)^{\varepsilon/(\varepsilon-1)} \right|$$

s.t.

$$T_{k,x} + T_{k,n} \leq 1$$

$$T_{k,m} \geq 0$$

and, obviously, the maximum is the corner solution with the higher $\tilde{N}_{k,m}$.

Now assume $\varepsilon \neq (\phi + 1)/\phi$. Form the Kuhn-Tucker Lagrangian

$$L(\bullet) = \left(\tilde{N}_{k,x} T_{k,x}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} + \tilde{N}_{k,n} T_{k,n}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} \right)^{\varepsilon/(\varepsilon-1)} - \lambda (T_{k,x} + T_{k,n} - 1)$$

Tucker conditions for a maximum

$$\begin{aligned}
& \tilde{N}_{k,x} (1 + \phi) T_{k,x}^{\frac{\phi\varepsilon - \phi - 1}{\varepsilon}} H_k - \lambda \leq 0 \\
& \tilde{N}_{k,n} (1 + \phi) T_{k,n}^{\frac{\phi\varepsilon - \phi - 1}{\varepsilon}} H_k - \lambda \leq 0 \\
& T_{k,x} \left(\tilde{N}_{k,x} (1 + \phi) T_{k,x}^{\frac{\phi\varepsilon - \phi - 1}{\varepsilon}} H_k - \lambda \right) = 0 \\
& T_{k,n} \left(\tilde{N}_{k,n} (1 + \phi) T_{k,n}^{\frac{\phi\varepsilon - \phi - 1}{\varepsilon}} H_k - \lambda \right) = 0 \\
& 1 - T_{k,x} - T_{k,n} \geq 0 \\
& \lambda (1 - T_{k,x} - T_{k,n}) = 0
\end{aligned}$$

First realize that $H_k \geq 0$, $H_k^* > 0$, and that both H_k and H_k^* are finite, for any admissible parametric values on the constrained domain. The finiteness follows from the limited time resources and the fact that, for any admissible parametric values, the production technology of efficient labor does not permit infinite output with limited resources. Moreover, it is always possible to allocate some resources to production of efficient labor such that it is positive and thus $H_k^* > 0$. To satisfy the first two Kuhn-Tucker conditions, λ must be positive, otherwise both $T_{k,x}$ and $T_{k,n}$ would have to be zero implying $H_k^* = 0$, which is inadmissible. Therefore, the time constraint is binding.

Now let us use the Kuhn-Tucker conditions to study the corner solution $T_{k,x} = 0$ and $T_{k,n} = 1$.

Realizing that $\left(\tilde{N}_{k,n} (1 + \phi) T_{k,n}^{\frac{\phi\varepsilon - \phi - 1}{\varepsilon}} H_k - \lambda = 0 \wedge \{T_{k,x}, T_{k,n}\} = \{0, 1\} \right) \Rightarrow \tilde{N}_{k,n} (1 + \phi) H_k = \lambda > 0$. Because

both $\tilde{N}_{k,n}$ and H_k are finite, λ is finite as well. Substituting for λ ,

$$\tilde{N}_{k,x} (1 + \phi) T_{k,x}^{\frac{\phi\varepsilon - \phi - 1}{\varepsilon}} H_k - \lambda \leq 0 \Leftrightarrow \tilde{N}_{k,x} T_{k,x}^{\frac{\phi\varepsilon - \phi - 1}{\varepsilon}} - \tilde{N}_{k,n} \leq 0. \quad \text{Because } T_{k,x} = 0 \quad \text{and}$$

$\tilde{N}_{k,x} T_{k,x}^{\frac{\phi\varepsilon - \phi - 1}{\varepsilon}} - \tilde{N}_{k,n} \leq 0$, it follows that $\tilde{N}_{k,x} T_{k,x}^{\frac{\phi\varepsilon - \phi - 1}{\varepsilon}}$ is finite (and well defined). This is the case whenever the exponent is larger than zero, that is, whenever $\varepsilon > (\phi + 1)/\phi$. This condition is thus the necessary condition for the studied corner solution to be the maximum. By symmetry, the same necessary condition must hold for the corner solution $T_{k,n} = 0$ then $T_{k,x} = 1$ to be the maximum.

If $T_{k,x} > 0$ and $T_{k,n} > 0$, in the interior solution, the first two Kuhn-Tucker conditions are equalities and the following result is obtained:

$$\tilde{N}_{k,x} (1+\phi) T_{k,x}^{\frac{\phi\varepsilon-\phi-1}{\varepsilon}} H_k - \lambda = 0 \wedge \tilde{N}_{k,n} (1+\phi) T_{k,n}^{\frac{\phi\varepsilon-\phi-1}{\varepsilon}} H_k - \lambda = 0 \Rightarrow$$

$$1 = \frac{\lambda}{\lambda} = \frac{\tilde{N}_{k,x} T_{k,x}^{\frac{\phi\varepsilon-\phi-1}{\varepsilon}}}{\tilde{N}_{k,n} T_{k,n}^{\frac{\phi\varepsilon-\phi-1}{\varepsilon}}} = \frac{\tilde{N}_{k,x} \left(\frac{T_{k,x}}{T_{k,n}} \right)^{\frac{\phi\varepsilon-\phi-1}{\varepsilon}}}{\tilde{N}_{k,n}} = \left(\frac{1+N_{k,x}}{1+N_{k,n}} \right)^{(\varepsilon-1)/\varepsilon} \left(\frac{T_{k,x}}{T_{k,n}} \right)^{\frac{\phi\varepsilon-\phi-1}{\varepsilon}}$$

It follows that $\frac{T_{k,x}}{T_{k,n}} = \left(\frac{1+N_{k,x}}{1+N_{k,n}} \right)^{\frac{\varepsilon-1}{1+\phi-\varepsilon\phi}}$. Using this result and the time constraint it follows that $T_{k,x} = t_k / (1+t_k)$ and $T_{k,n} = 1 / (1+t_k)$.

Consequently, there are three possible candidates for the maximum, two corner solutions and one interior solution. Given the results above, evaluating the objective function at each of these candidates, the values of H_k at the candidate time allocations are

$$H_k \Big|_{T_{k,x} = \frac{t_k}{1+t_k}, T_{k,n} = \frac{1}{1+t_k}} = \left(\tilde{N}_{k,x} \left(\frac{t_k}{1+t_k} \right)^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} + \tilde{N}_{k,n} \left(\frac{1}{1+t_k} \right)^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} \right)^{\varepsilon/(\varepsilon-1)}$$

$$H_k \Big|_{T_{k,x}=0, T_{k,n}=1} = \left(\tilde{N}_{k,x} T_{k,x}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} + \tilde{N}_{k,n} T_{k,n}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} \right)^{\varepsilon/(\varepsilon-1)} = \tilde{N}_{k,n}^{\varepsilon/(\varepsilon-1)}$$

$$H_k \Big|_{T_{k,x}=1, T_{k,n}=0} = \left(\tilde{N}_{k,x} T_{k,x}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} + \tilde{N}_{k,n} T_{k,n}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} \right)^{\varepsilon/(\varepsilon-1)} = \tilde{N}_{k,x}^{\varepsilon/(\varepsilon-1)}$$

Now I verify the sufficient conditions for each of these candidates to be the maximum. Note that from the Kuhn-Tucker first order conditions the time constraint is binding and therefore we can rewrite the agent's problem in the following way

$$\text{Max}_{T_{k,x}} \left(\tilde{N}_{k,x} T_{k,x}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} + \tilde{N}_{k,n} (1-T_{k,x})^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} \right)^{\varepsilon/(\varepsilon-1)}$$

s.t.

$$1 \geq T_{k,x} \geq 0$$

Consider the new objective function on the constrained domain. Differentiate the objective function to obtain

$$\frac{\partial^2 \left(\left(\tilde{N}_{k,x} T_{k,x}^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} + \tilde{N}_{k,n} (1-T_{k,x})^{\frac{(\phi+1)(\varepsilon-1)}{\varepsilon}} \right)^{\varepsilon/(\varepsilon-1)} \right)}{\partial T_{k,x}^2} =$$

$$\frac{(1+\phi)H_k \left(-\tilde{N}_{k,x} \tilde{N}_{k,n} (\phi\varepsilon - \phi - 1) (1-T_{k,x})^{\frac{1+\phi}{\varepsilon}} (T_{k,x} (1-T_{k,x}))^{\phi} T_{k,x}^{\frac{1+\phi}{\varepsilon}} - \varepsilon \phi (1-T_{k,x}) T_{k,x} \left(\tilde{N}_{k,x} T_{k,x}^{\phi} (1-T_{k,x})^{\frac{\phi+1}{\varepsilon}} - \tilde{N}_{k,n} (1-T_{k,x})^{\phi} T_{k,x}^{\frac{\phi+1}{\varepsilon}} \right)^2 \right)}{-\varepsilon (1-T_{k,x}) T_{k,x} \left(\tilde{N}_{k,x} T_{k,x}^{\phi+1} (1-T_{k,x})^{\frac{\phi+1}{\varepsilon}} + \tilde{N}_{k,n} (1-T_{k,x})^{\phi+1} T_{k,x}^{\frac{\phi+1}{\varepsilon}} \right)^2}$$

Now it is straightforward to see that

$$\text{Sign} \left\{ \frac{\partial^2 H_k}{\partial T_{k,x}^2} \right\} = \text{Sign} \left\{ \tilde{N}_{k,x} \tilde{N}_{k,n} (\phi\varepsilon - \phi - 1) (1-T_{k,x})^{\frac{1+\phi}{\varepsilon}} (T_{k,x} (1-T_{k,x}))^{\phi} T_{k,x}^{\frac{1+\phi}{\varepsilon}} + \varepsilon \phi (1-T_{k,x}) T_{k,x} \left(\tilde{N}_{k,x} T_{k,x}^{\phi} (1-T_{k,x})^{\frac{\phi+1}{\varepsilon}} - \tilde{N}_{k,n} (1-T_{k,x})^{\phi} T_{k,x}^{\frac{\phi+1}{\varepsilon}} \right)^2 \right\}$$

and that the term on the right hand side has a positive sign for any $1 > T_{k,x} > 0$ whenever $\phi\varepsilon - \phi - 1 > 0 \Leftrightarrow \varepsilon > (\phi+1)/\phi$. Thus, the interior solution cannot be the maximum if $\varepsilon > (\phi+1)/\phi$. As argued above, the objective function is certainly continuous and bounded on the constrained domain. Therefore, there must exist a maximum. Having excluded the interior candidate, the corner solution with the higher $\tilde{N}_{k,m}$ and thus $\tilde{N}_{k,m}^{\varepsilon/(\varepsilon-1)}$ is the maximum on the constrained domain.

From the Kuhn-Tucker conditions we know that if $\varepsilon < (\phi+1)/\phi$, none of the corner candidates can be the maximum. As above, because the objective function is continuous and bounded on the constrained domain, there must exist a maximum. Being the only remaining possibility, the interior solution is the maximum and $H_k^* = H_k(t_k/(1+t_k), 1/(1+t_k))$ whenever $\varepsilon < (\phi+1)/\phi$. For the implication in (9a-b), note that $\varepsilon/(\varepsilon-1) > 0$. This completes the proof.⁴⁵ ■

Proof of Proposition 2

Individuals invest their time in those social networks of a given type, exclusive or inclusive, that offer the largest network effects and thus are the most efficient for acquiring that type of skill. Assume there is an equilibrium with an individual violating the proposition thus involved in two networks of a given type, splitting his time between these networks. It must then be that the network effects in these two networks are the same for this individual; otherwise he would pick the one that is more efficient to spend his time. Such equilibrium is unstable, however. Given that network effects increase in agents' involvements, any marginal deviation in allocation of agents causes the network effects between the two networks to differ and, as a consequence, the agent to abandon the less efficient network. Obviously, the reaction of the other individuals to

⁴⁵ A simpler way to determine which of the three candidates is the maximum is possible, noting that $(\phi+1)\frac{(\varepsilon-1)}{\varepsilon}$ and $\frac{\varepsilon}{(\varepsilon-1)}$ fully determine the properties of the maximization problem.

such marginal deviation does not stabilize the equilibrium, as network effects are increasing in individual involvements. ■

Proof of Proposition 3

Given the asymmetric information on the labor market, individuals take the wage for the unit of their efficient labor as given with respect to their choice of skills. Therefore, individuals pick that combination of social networks and thus skills that is the most efficient in production of efficient labor. To prove Proposition 3 by contradiction, assume there is an equilibrium with two individuals from a given social group that are involved in two different combinations of social networks. Because the two individuals are free to choose between networks, in this equilibrium it must be that the efficiencies of these two combinations of social networks for the two individuals in production of efficient labor are the same. Such equilibrium is, however, unstable. Any marginal deviation from the equilibrium agent involvements across these two different combinations of networks causes their efficiencies to differ. As a result, the agent involved in the less efficient set of social networks switches to the more efficient one. As above, the reaction of the other individuals to the marginal deviations does not stabilize the equilibrium. ■

Proof of Proposition 4

Recall that, due to asymmetric information, employers differentiate wages per unit of efficient labor across social groups but not within groups and that individuals cannot change their group membership. Therefore, agents pick those networks to acquire skills that they are allowed to join and that offer the largest network effect benefits (and thus are the most efficient in skill acquisition). Applying (8) to network effect specification in (14) and (15), the network effects are:

$$N_{i,m}^r(I_m^r L_{i,m}, J_m^r L_{j,m}, \delta) = \left(I_m^r + \frac{1}{1+\delta} J_m^r \right)^\gamma \left(\frac{\phi}{1+\phi} \right)^\gamma$$

$$N_{j,m}^r(I_m^r L_{i,m}, J_m^r L_{j,m}, \delta) = \left(\frac{1}{1+\delta} I_m^r + J_m^r \right)^\gamma \left(\frac{\phi}{1+\phi} \right)^\gamma$$

where the superscript $r \in \{EE, II, EI, IE\}$ denotes the prevailing equilibrium.

As a first step, recalling Proposition 3 and its proof, note that deviation to an empty social network is never beneficial and for any individual it is preferable to be a member of the social network populated by the other members of her social group. Now let us consider deviations that are permitted and that involve switching to a non-empty social network, equilibrium by equilibrium.

As concerns the EE equilibrium, there is no permissible deviation to a non-empty social network. IT is thus always stable. The same holds for the II equilibrium.

Now investigate the IE equilibrium. Under this equilibrium a majority individual could switch to a minority inclusive network, which is non-empty. For such marginal deviation not to occur, the stability condition is $N_{j,x}^{IE} \geq N_{j,n}^{IE}$. Now note that $N_{j,x}^{IE} = J^\gamma (\phi/(1+\phi))^\gamma$ and $N_{j,n}^{IE} = (1/(1+\delta)I)^\gamma (\phi/(1+\phi))^\gamma$. Thus, the stability condition is equivalent to $J \geq I/(1+\delta)$, which always holds and the IE equilibrium is always stable.

As concerns the EI equilibrium, the argument is similar to the one just above. The possible deviation now involves minority individuals who compare $N_{i,x}^{EI} = I^\gamma (\phi/(1+\phi))^\gamma$ and $N_{i,n}^{EI} = (J/(1+\delta))^\gamma (\phi/(1+\phi))^\gamma$. This yields the condition $I \geq J/(1+\delta)$, which holds, if the minority (or social distance) is large enough.

Now consider the IS equilibrium. In any stable IS equilibrium it must be that agents prefer staying in the inclusive networks occupied by their own social group, enjoying network effects $N_{i,n}^{IS} = I^\gamma (\phi/(1+\phi))^\gamma$ and $N_{j,n}^{IS} = J^\gamma (\phi/(1+\phi))^\gamma$, to deviating to the inclusive social network composed of the members of the other social group and obtaining network benefits $\frac{1}{1+\delta} J^\gamma (\phi/(1+\phi))^\gamma$ and $\frac{1}{1+\delta} I^\gamma (\phi/(1+\phi))^\gamma$, respectively. These conditions hold if and only if $I \geq J/(1+\delta)$ and $J \geq I/(1+\delta)$, of which the second is always satisfied. Because $I + J = 1$, $I \geq J/(1+\delta) \Leftrightarrow I \geq 1/(2+\delta)$. ■

Proof of Proposition 5 (superscripts omitted)

Recall that $I + J = 1$. Realize that the $\omega(I)$ curve is continuous for any $I > 0$ and continuous from the right at $I = 0$. The proof consists of three steps. First, note that whenever ρ is finite, $\omega(I)$ approaches infinity as $I \rightarrow 0$, $\lim_{I \rightarrow 0} \omega(I) = \infty$. Second, realize that by symmetry $\omega(I) = 1$ at $I = 0.5$. In the last step tedious algebraic manipulations (partial differentiation of $\omega(I)$ with respect to I , computations omitted) yield that whenever the parametric condition $(\phi/(1+\phi))^\gamma (\rho\gamma - \gamma - 1) > 2^\gamma$ is satisfied the $\omega(I)$ curve is upward sloping at $I = 0.5$. In consequence, the continuous $\omega(I)$ curve lies above one at $I \rightarrow 0$ and has a positive slope at the point $\omega(0.5) = 1$ if the above parametric condition holds. As a result, under the IE equilibrium there always is a range of I where $\omega(I) < 1$ and $\partial\omega(I)/\partial I < 0$. The same is true under the EI equilibrium if social distance δ is large enough such that for some I from this range the EI equilibrium is stable. This is the case for any $\delta \geq 1/I - 2$ for I from such range, which is well defined, and for $\delta \rightarrow \infty$ in particular. Simple algebra shows that the parametric condition above is satisfied for all $\rho > \rho'$, where $\rho' \equiv ((2(\phi+1)/\phi)^\gamma + \gamma + 1)/\gamma > 0$. ■

Proof of Proposition 6

Recall that agents always pick those networks to acquire skills that they are allowed to join and that offer the largest network effects and thus are the most efficient. The only possibility for an individual to deviate in the equilibrium where both social groups acquire inclusive skills in one inclusive social network is to form his own inclusive social network. Because such network would offer zero network benefits, as compared to positive network benefits in the integrated inclusive network, such deviation is never profitable and therefore the equilibrium is stable.

If, on the other hand, inclusive skills are acquired in two segregated inclusive networks, for this equilibrium to be stable it must be that all individuals prefer staying in the inclusive networks occupied by their own social group, enjoying network effects $N_{i,n}^{DS} = I^\gamma$ and $N_{j,n}^{DS} = J^\gamma$, to deviating to the inclusive social network of the other social group and obtaining network benefits

$\frac{1}{1+\delta}J^\gamma$ and $\frac{1}{1+\delta}I^\gamma$, respectively. These conditions hold if and only if $I \geq \frac{1}{1+\delta}J$ and $J \geq \frac{1}{1+\delta}I$, of which the second is always satisfied. Note that because $I+J=1$, $I \geq \frac{1}{1+\delta}J \Leftrightarrow I \geq \frac{1}{2+\delta}$. ■

Proof of Proposition 7 and Proposition 8

First note that:

$$\begin{aligned} I < \frac{1}{2} &\Leftrightarrow \left(\frac{1}{1+\delta} \frac{I}{(1-I)} + 1 \right)^\gamma < \left(1 + \frac{1}{1+\delta} \frac{(1-I)}{I} \right)^\gamma \Leftrightarrow \frac{1 + \left(\frac{1}{1+\delta} I + (1-I) \right)^\gamma}{1 + \left(I + \frac{1}{1+\delta} (1-I) \right)^\gamma} = \\ &= \frac{1 + (1-I)^\gamma \left(\frac{1}{1+\delta} \frac{I}{(1-I)} + 1 \right)^\gamma}{1 + I^\gamma \left(1 + \frac{1}{1+\delta} \frac{(1-I)}{I} \right)^\gamma} < \frac{1 + (1-I)^\gamma}{1 + I^\gamma} \Leftrightarrow \frac{1 + I^\gamma}{1 + (1-I)^\gamma} \frac{1 + \left(\frac{1}{1+\delta} I + (1-I) \right)^\gamma}{1 + \left(I + \frac{1}{1+\delta} (1-I) \right)^\gamma} < 1 \end{aligned}$$

which yields:

$$1) \left(\frac{1 + I^\gamma}{1 + (1-I)^\gamma} \frac{1 + \left(\frac{1}{1+\delta} I + (1-I) \right)^\gamma}{1 + \left(I + \frac{1}{1+\delta} (1-I) \right)^\gamma} \right)^{\frac{\varepsilon-1}{1+\phi-\varepsilon\phi}} = \frac{t_i^{DI}}{t_j^{DI}} = \frac{l_i^{DI}}{l_j^{DI}} > 1 \Leftrightarrow \frac{\varepsilon-1}{1+\phi-\varepsilon\phi} < 1 \Leftrightarrow \varepsilon < 1 \wedge \varepsilon < (1+\phi)/\phi$$

$$2) \frac{t_i^{DI}}{t_j^{DI}} = \frac{l_i^{DI}}{l_j^{DI}} < 1 \Leftrightarrow \frac{\varepsilon-1}{1+\phi-\varepsilon\phi} > 1 \Leftrightarrow \varepsilon > 1 \wedge \varepsilon < (1+\phi)/\phi$$

$$3) \frac{t_i^{DI}}{t_j^{DI}} = \frac{l_i^{DI}}{l_j^{DI}} = 1 \Leftrightarrow \frac{\varepsilon-1}{1+\phi-\varepsilon\phi} = 0 \Leftrightarrow \varepsilon = 1 \wedge \varepsilon < (1+\phi)/\phi$$

$$4) \left(\frac{1 + I^\gamma}{1 + (1-I)^\gamma} \frac{1 + \left(\frac{1}{1+\delta} I + (1-I) \right)^\gamma}{1 + \left(I + \frac{1}{1+\delta} (1-I) \right)^\gamma} \right)^{\frac{1}{1+\phi-\varepsilon\phi}} = \frac{S_{i,x}^{DI}}{S_{i,n}^{DI}} \Big/ \frac{S_{j,x}^{DI}}{S_{j,n}^{DI}} < 1 \quad \forall \varepsilon < (1+\phi)/\phi$$

This completes the proof. ■

Proof of Proposition 9 (superscripts omitted)

To prove this proposition, I state and prove the following lemma:

Lemma 1:

There always exists $I > 0$ such that $h(I) < 1$.

Proof of Lemma 1

If $\delta \rightarrow \infty$, one obtains that $H_i(I)|_{I \rightarrow 0} < H_j(I)|_{I \rightarrow 0}$ directly from the inequality $H_i(I)|_{I \rightarrow 0, \delta \rightarrow \infty} = 2^{\frac{\varepsilon}{\varepsilon-1}-\phi-1} < 2^{\frac{\varepsilon}{\varepsilon-1}-\phi} = H_j(I)|_{I \rightarrow 0, \delta \rightarrow \infty}$. For the case of finite δ , in the first step I prove that $H_i(I)|_{I \rightarrow 0} < H_j(I)|_{I \rightarrow 0}$ and both are well defined if $\varepsilon \neq 1$. Using (5), (6), (18a-b), and the specifications of network effects under the DI equilibrium:

$$H_i(I)|_{I \rightarrow 0} = \left(1 + \left(\frac{1}{1 + (1/(1+\delta))^\gamma} \right)^{\frac{1-\varepsilon}{\phi\varepsilon-\phi-1}} \right)^{-1-\phi} \left(\left(\left(\left(\frac{1}{1 + (1/(1+\delta))^\gamma} \right)^{\frac{1-\varepsilon}{\phi\varepsilon-\phi-1}} \right)^{1+\phi} \right)^{\frac{\varepsilon-1}{\varepsilon}} + \left(1 + \left(\frac{1}{1+\delta} \right)^\gamma \right)^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

After some algebra one obtains $H_i(I)|_{I \rightarrow 0} = \left(1 + \left(\frac{1}{1+\delta} \right)^\gamma \right) \left(\left(1 + \left(\frac{1}{1+\delta} \right)^\gamma \right)^{\frac{\varepsilon-1}{\phi\varepsilon-\phi-1}} + 1 \right)^{\frac{1}{\varepsilon-1}-\phi}$. Now,

noting that $1 \leq \left(1 + \left(\frac{1}{1+\delta} \right)^\gamma \right) \leq 2$, one obtains that:

$$H_i(I)|_{I \rightarrow 0} = \left(1 + \left(\frac{1}{1+\delta} \right)^\gamma \right) \left(\left(1 + \left(\frac{1}{1+\delta} \right)^\gamma \right)^{\frac{\varepsilon-1}{\phi\varepsilon-\phi-1}} + 1 \right)^{\frac{1}{\varepsilon-1}-\phi} \leq 2 \left(\left(1 + \left(\frac{1}{1+\delta} \right)^\gamma \right)^{\frac{\varepsilon-1}{\phi\varepsilon-\phi-1}} + 1 \right)^{\frac{1}{\varepsilon-1}-\phi}.$$

It is easy to see that $H_j(I)|_{I \rightarrow 0} = 2^{\frac{\varepsilon}{\varepsilon-1}-\phi}$. Having these results, the condition

$$\left(\left(1 + (1/(1+\delta))^\gamma \right)^{\frac{\varepsilon-1}{\phi\varepsilon-\phi-1}} + 1 \right)^{\frac{1}{\varepsilon-1}-\phi} < 2^{\frac{1}{\varepsilon-1}-\phi} \quad (\text{A.1})$$

implies $H_i(I)|_{I \rightarrow 0} < H_j(I)|_{I \rightarrow 0}$. Realize now that the diversification condition $\varepsilon < (\phi+1)/\phi$ is

equivalent to: $\left((\varepsilon > 1) \wedge \left(\frac{1}{\varepsilon-1} - \phi > 0 \right) \right) \vee \left((\varepsilon < 1) \wedge \left(\frac{1}{\varepsilon-1} - \phi < 0 \right) \right) \vee \varepsilon = 1$, defining three

cases when the diversification condition is satisfied. Since the last one is excluded in the proposition, let us consider the first two, denoting them Case A and Case B.

Consider Case A when $(\varepsilon > 1) \wedge \left(\frac{1}{\varepsilon - 1} - \phi > 0\right)$. Due to the second inequality and noting that in

(A.1) it holds that $\left(1 + \left(\frac{1}{1 + \delta}\right)^\gamma\right)^{\frac{\varepsilon - 1}{\phi\varepsilon - \phi - 1}} + 1 \geq 1$, one can take away the powers without changing

the inequality in (A.1). Accordingly, rewrite (A.1) as follows: $\left(1 + \left(\frac{1}{1 + \delta}\right)^\gamma\right)^{\frac{\varepsilon - 1}{\phi\varepsilon - \phi - 1}} < 1$ and realize

that $\frac{\varepsilon - 1}{\phi\varepsilon - \phi - 1} < 0$. It follows that for finite δ it holds that $1 + \left(\frac{1}{1 + \delta}\right)^\gamma > 1$ and thus (A.1) holds,

which implies that $H_i(I)|_{I \rightarrow 0} < H_j(I)|_{I \rightarrow 0}$.

Consider now Case B where $(\varepsilon < 1) \wedge \left(\frac{1}{\varepsilon - 1} - \phi < 0\right)$. As in Case A, we can take away the powers but the direction of the inequality must be changed because of the powers are now

negative. Thus, we can rewrite the inequality in (A.1) as follows: $\left(1 + \left(\frac{1}{1 + \delta}\right)^\gamma\right)^{\frac{\varepsilon - 1}{\phi\varepsilon - \phi - 1}} > 1$. Now

realize that $\frac{\varepsilon - 1}{\phi\varepsilon - \phi - 1} > 0$. This inequality and that $1 + \left(\frac{1}{1 + \delta}\right)^\gamma > 1$ for any finite δ prove that

(A.1) holds and that $H_i(I)|_{I \rightarrow 0} < H_j(I)|_{I \rightarrow 0}$ in Case B as well.

Realizing that, by symmetry, $H_i(I)|_{I \rightarrow 0.5} = H_j(I)|_{I \rightarrow 0.5}$ and $h(I)$ is continuous for any $I \geq 0$, there exists $I > 0$ such that $H_i(I) < H_j(I)$ and thus $h(I) < 1$ for any admissible parametric values. Denote Γ_I the set of all such I .

If $\varepsilon = 1$ such that $H_k = \left(\left(\frac{\phi}{1 + \phi}\right)^\phi T_{k,x}^{\phi+1} (1 + N_{k,x}) \left(1 - \frac{\phi}{1 + \phi}\right)\right)^{1/2} \left(\left(\frac{\phi}{1 + \phi}\right)^\phi T_{k,n}^{\phi+1} (1 + N_{k,n}) \left(1 - \frac{\phi}{1 + \phi}\right)\right)^{1/2}$,

$h = \frac{(1 + N_{i,x})^{1/2} (1 + N_{i,n})^{1/2}}{(1 + N_{j,x})^{1/2} (1 + N_{j,n})^{1/2}}$. Evidently, network effects are larger for the majority and thus $h < 1$

whenever $\varepsilon = 1$. This completes the proof of Lemma 1.

Now recall that $\omega(I) = \left(\frac{I}{1 - I} h(I)\right)^{\frac{-1}{\rho}} h(I)$. Because $\lim_{\rho \rightarrow \infty} \left(\left(\frac{I}{1 - I} h(I)\right)^{\frac{-1}{\rho}}\right) = 1$ and $\left(\frac{I}{1 - I} h(I)\right)^{\frac{-1}{\rho}}$ is

a decreasing continuous function of ρ for any $I \in (0, 0.5)$, there always exists ρ large enough such that $\omega(I) < 1$ for some $I \in \Gamma_I$. In particular, it can be shown that for any

$\rho > \rho'' \equiv \text{Min}_{I \in \Gamma_I} \left| \text{Log} \left(\frac{I}{(1-I)} h(I) \right) / \text{Log}(h(I)) \right|$ it holds that $\omega(I) < 1$ for some $I \in \Gamma_I$.

Straightforwardly, ρ'' is well defined for any $I \in (0,1)$ and is positive and finite.

Similarly to the proof of Proposition 5, $\omega(I)$ is continuous for all $I \in (0,1)$, approaches infinity as I approaches zero, and $\omega(I) = 1$ for $I = 0.5$. Adding the fact that for $\rho > \rho''$, there is I such that $\omega(I) < 1$ suffices to ensure that there always is a range of I where $\omega(I) < 1$ and $\partial \omega(I) / \partial I < 0$. This completes the proof. ■

Proof of Proposition 10

Because $\omega''(I) < 1$ and there always exists I such that $\omega^{EI}(I) = \omega^{IE}(I) > 1$ (c.f. Proof of Proposition 5), it must be that $\omega^{EI}(I) = \omega^{IE}(I) > \omega''(I)$ for some I . This proves the first part. Moreover, if $\rho > \rho' \equiv \left((2(\phi+1)/\phi)^\gamma + \gamma + 1 \right)^{-1}$ there exists I such that $\omega^{EI}(I) = \omega^{IE}(I) < 1$. In addition, there also exists I such that $\omega^{EI}(I) = \omega^{IE}(I) = 1$. Since the functions $\omega^{EI}(I)$ and $\omega^{IE}(I)$ are continuous and because $\omega''(I) < 1$ for all I whenever $\delta > 0$, there exists I such that $1 > \omega^{EI}(I) = \omega^{IE}(I) > \omega''(I)$. ■