NONLINEAR NEIGHBORHOOD INTERACTIONS AND INTERGENERATIONAL TRANSMISSION OF HUMAN CAPITAL

by

Yannis M. Ioannides[1]

Department of Economics

Tufts University

Medford, MA 02155

+1.617.627.3294. yioannid @ tufts.edu

July 10, 2000

Abstract

The paper develops a model of the intergenerational transmission of human capital that reflects both individual choice and neighborhood interactions. It implies a law of motion for human capital that under certain conditions exhibits multiple equilibria that differ in terms of stability. As a result, the distribution of human capital across the population in the long run may not collapse into a single value. This result is entirely due to: first, presence of nonlinear neighborhood interactions, and, second, the elasticity of substitution between own human capital and neighborhood interactions being sufficiently larger than the the elasticity of substitution between one's child's human capital and own consumption. The model supports the Kuznets hypothesis and may be extended to allow for residential choice. Empirical analysis with geocoded data from the Panel Study of Income Dynamics, validates Kremer's results on the role of neighborhood effects, but also confirms by means of parametric and nonparametric methods the model's predictions of nonlinear effects of parents' education and of neighbors' education.

JEL Classification Codes: O40, O15, I21.

Keywords: Neighborhood effects, human capital, neighborhood interactions, economic growth, complementarity.

neinthu2.tex. Version: July 10, 2000

[1] I thank Costas Azariadis, Marcelo Coca, Steven Durlauf, Dennis Epple, Anna Hardman, Michael Kremer, Gib Metcalf, Scotte Page, Steve Ross, Danny T. Quah, Jeff Zabel and seminar participants in the University of Helsinki, the Athens University of Economics and Business, the London School of Economics and the University of Connecticut for useful comments, and Luis Ahn, Marcelo Coca, David Iaia, Pamela Jia and Anne K. Thompson for expert research assistance. Thanks are also due to the PSID staff at the University of Michigan for their help with the data, to Danny T. Quah for giving me access to tSrF and to Henry G. Overman for help with the program. This version is a major revision of my working paper with the same title, dated October 16, 1997.

Financial support by the John D. and Katherine T. MacArthur Foundation and the National Science Foundation is gratefully acknowledged.

1 Introduction

I develop a model of the evolution of human capital as an outcome of individual choice. A person chooses how much of her own human capital to devote to raising children and how much to allocate for own consumption. Production of offspring human capital is the only means by which individuals save. The model implies that under certain conditions the law of motion for the evolution of human capital exhibits multiple equilibria, two of which are stable and one is unstable. It is thus possible that the distribution of human capital across the population in the long run may not collapse into a single point.

This feature of the model is entirely dependent on two features: first, nonlinear neighborhood interactions must be present; second, the elasticity of substitution between own human capital and the neighborhood interactions must be sufficiently larger than the elasticity of substitution between one's child's human capital and own consumption. Neighborhood interactions in general represent the role of public education in the production of human capital. The model supports the Kuznets hypothesis, namely that the income distribution would worsen before it improves during the process of economic growth [Kuznets (1955)]. I obtain a complete characterization of the properties of the intertemporal evolution of human capital, for the family of the functional specifications that I assume, and when it is an outcome of optimization by parents in the presence of neighborhood interactions.

The theoretical assumptions I make in the paper are intended to justify empirical specifications made by several researchers, but especially by Kremer (1997). I also report here empirical findings, which are based on geocoded data from the Panel Study of Income Dynamics and are similar to Kremer's findings of significant linear effects on a person's human capital, measured by years of education, of the average education in the neighborhood where he or she grew up (suitably defined). However, my results show that a person's education is nonlinearly related to both father's and mother's education, when both are present in the regression, and to the mean, second and third moments of the distribution of education within the neighborhood where an individual was brought up. These findings are also supported by non-parametric estimates. They confirm a key prediction of the theory in this paper, namely that under certain conditions, the relationship between a child's education and that of his or her parent has a sigmoid shape, although the impact of the education of parents is different from that of education in the census tract where an individual grew up.

The model of the paper allows for the acquisition of human capital to involve both private and publicly provided inputs. Individuals and parents influence the former directly and can exercise some (possibly indirect) control over the latter. Examples of parents' direct choice over inputs into the educational process are sending one's children to private school, spending time helping their children with their schoolwork, or paying for private tutors. Even if individuals rely on public education, they still may exercise choice when they decide what schools to send their children to, if there is choice, or where to locate, if that might make a difference, as in societies where the public provision of education is locally controlled, as in the US and other countries. In societies where the provision of education is highly centralized, choice of location may still influence the quality of education via the characteristics of the other students who reside in the same community. The behavioral model of this paper emphasizes how parental choice of human capital investment for their children is affected by their own human capital and by a neighborhood effect.

Formal education, of course, is only one factor in the development of human capital. It is complemented by direct parental inputs and by health, recreation, and culture, all of which are influenced by the choice of community of residence. Similarly, access to information about educational requirements of jobs may be facilitated by social interactions in one's residential community.

The coexistence of individual and social factors in the intergenerational transmission of human capital has been addressed theoretically by Azariadis and Drazen (1990), Benabou (1996a; 1996b) and Durlauf (1996a; 1996b). Benabou and Durlauf, in particular, emphasize the role of of an economy's community structure on human capital accumulation in the presence of social spillovers. In the empirical literature, the paper is most closely related to Borjas (1992; 1995) and Kremer (1997).

In the remainder of this paper, I review in Section 2 key empirical findings, obtained by Borjas (1992) and Kremer (1997). In Section 3, I introduce a model that encompasses certain empirical models, especially those of Borjas (1992) and Kremer (1997), as special cases. I explore the properties of a family of models, in which both parental and social components contribute to the process of intergenerational transmission of human capital. These models are amenable to empirical testing. I investigate their dynamics, which turn out to be quite rich. In Section 4 I present empirical results which confirm the importance of *nonlinear* effects of parental and neighborhood education. In Section 5 I show that the dynamics of the model may predict intertemporal variations of the income distribution along the lines of the Kuznets hypothesis. I also show how to apply the model to the case of local communities, where individuals' being able to choose where to locate may serve to internalize neighborhood interactions. Section 6 concludes.

2 Neighborhood Effects in the Intergenerational Transmission of Human Capital

The empirical literature on neighborhood effects construes them literally, as effects on one's decisions of the characteristics and of the decisions of one's neighbors, as well as metaphorically, as interactions in non-geographical sense, such as social interactions. Empirical research in this area is hampered by paucity of suitable data. The two studies that I review below are significant also for their use of data.

2.1 Kremer's Findings on Neighborhood Effects

I use the term human capital to refer to the educational attainment in years of formal schooling. More generally, let $H_{i,t+1}$ denote human capital of a member of the ith dynasty in generation $t+1$, and let h_{it} denote that of a member of the same dynasty in generation t; agent i' is the spouse of agent i; $\nu(i)$ denotes the set of neighbors of agents i and i', and n_i is its size, $n_i = |\nu(i)|$.

In an important paper on neighborhood effects, Kremer (1997) *postulates* the following law of intergenerational transmission of educational attainment:

$$H_{it+1} = a_0 + \frac{\alpha}{2}(h_{it} + h_{i't}) + \beta h_{\nu(i)t} + \epsilon_i, \tag{1}$$

where a_0 denotes an exogenous intercept, $h_{\nu(i)t}$ is average education in the neighborhood of i's upbringing, $h_{\nu(i)t} = \frac{1}{n_i}\sum_{j\in\nu(i)} h_{jt}$, and ϵ_i is a stochastic shock, and H_{it+1} is capitalized because it is an endogenous variable. Kremer uses Equ. (1) to obtain estimates of coefficients α and β, and to study the intertemporal evolution of the variance of schooling (which may also be interpreted as a measure of the inequality of log earnings).

Kremer's estimates of neighborhood effects are large, when they are compared to the effect of parents' education, although there is no obvious metric. Kremer makes the point, however, that with the observed values for the parameters of the model *sorting* of individuals into neighborhoods contributes little to the magnitude of the variance of schooling in steady state, σ^2_∞. Specifically, with his findings of $\hat{\alpha} = .395$ (.051), $\hat{\beta} = .149$ (.072), $\hat{\sigma}_\epsilon = 1.79$ years, "living in an educated neighborhood increases the expected education for one's child by three-quarters as marrying an educated spouse, since the effect of each parent is half the total parental effect of .395." by only 1.7 %, if ρ_ν were to go up[2] from .2 to .4; it will go up by only .9%, if ρ_m were to go up from .6 to .8. Kremer's findings on the importance of neighborhood effects make it even more pressing to explore what sort of behavioral setting would give rise to equations like his key estimating equation (1).

2.2 Borjas' Findings on Ethnicity and Neighborhood Effects

Borjas (1992) assumes that individuals value their own consumption and the human capital of their descendant according to a constant elasticity of substitution utility function. He assumes that a child's human capital is expressed as a Cobb-Douglas function first of the fraction an individual allocates of her own human capital to her child's upbringing, and second, of a local interaction effect, represented by the mean human capital of her ethnic group. These assumptions imply an intergenerational human capital (or earnings) mobility equation that relates an individual's human capital, H_{igt+1}, to those of her parent, h_{igt}, and to the mean human capital in his ethnic group g, $\bar{h}_{gt}$ as follows:

$$H_{igt+1} = \gamma_1 h_{igt} + \gamma_2 \bar{h}_{gt} + \xi_{igt}, \qquad (2)$$

where all variables are measured as deviations from the mean, the shock ξ_{igt} may be decomposed as $\xi_{igt} = \epsilon_{igt} + \varepsilon_{gt}$, the random variables ϵ and ε are uncorrelated, and the random variable ξ_{igt} has the stochastic structure of the random effects model. The estimates with (2) with log wages data by Borjas range from .1829 to .2664, for γ_1, and from .1455 to .4589, for γ_2; the estimates with schooling data vary from .2566 to .3465, for γ_1, and from .0990 to .2983, for γ_2.[3]

[2] See *ibid.*, fn. 12: the estimates are .649, .633, and .620, for 1940, 1960, and 1980, respectively.

[3] Borjas (1995) explores further the nature of the ethnic externality by considering a potential link between parental and ethnic capital, on one hand, and residential segregation, on the other. Borjas estimates a version of (2) with data that provides for several observations from each residential "neighborhood" and thus allow him to account for neighborhood effects. All of his estimates of expected skills of individuals (measured by either educational attainment

3 Nonlinear Models of The Intergenerational Transmission of Human Capital

The differences between Kremer's and Borjas's results may have rather minor consequences for the dynamics of human capital formation in an economy, if one were to rely upon linear dynamic models. Results could differ dramatically if nonlinear processes were at work. Next I explore a class of models of neighborhood interactions and human capital, which imply the models of Borjas and Kremer as special cases. The general model is somewhat related to Azariadis and Drazen (1990) and exploits the behavioral model of Borjas (1992) to its full generality. It allows me to explore some of the same issues as those raised by those two and other researchers.

The economy consists of a large number of agents, each of whom lives for a single period. Each agent has an endowment of human capital, which she must allocate to the production of labor services and of human capital for her offspring. I adopt the assumption of Borjas (1992), Equ. (1), of a utility function for the parent, with a constant elasticity of substitution (CES) between the child's human capital, H_{it+1}, measured in efficiency units, and own consumption, C_{it} :

$$U = U(H_{it+1}, C_{it}) \equiv \left[\zeta (H_{it+1})^{1-\frac{1}{\sigma}} + (1-\zeta)(C_{it})^{1-\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \sigma > 0. \quad (3)$$

The parameter σ, which may exceed 1, denotes the elasticity of substitution between human capital stock of the child and own consumption. A parent and her child are always indexed by the same subscript. However, I do not consider marital sorting, and therefore I drop the subscript i from now on, when no confusion arises.

Let s_t denote the savings rate, that is, the fraction of own human capital that a parent devotes to the production of human capital of her child. The remainder is allocated to paying for her own consumption, $(1-s_t)R_t h_t = C_t$, where R_t denotes the real wage rate per efficiency unit of labor, with consumption as numeraire. A child's human capital is produced by the parent's own input, $s_t h_t$, and the local interaction effect, v_t, from the community where individual i lives. They combine

or log wages) against the average skills in the parent's generation, $\mathcal{E}[H_{igt+1}] = (\gamma_1 + \gamma_2)\bar{h}_{gt}$, show that accounting for fixed neighborhood effects reduces the estimates of mean convergence, denoted by $\gamma_1 + \gamma_2$, in the skills of ethnic groups. Roughly, the estimates of this effect go down from around .45 to .20, when neighborhood fixed effects are included and when a small number of neighborhood characteristics are included in the regression (%-age of population with at least high-school diploma, percentage with at least college diploma, labor force participation rate of men and women, the unemployment rate, %-age of workers in professional occupations, the %-age of families below the poverty level, and the %-age of families with at least 15,000$ of income).

through a constant elasticity of substitution production function to produce the child's human capital, H_{t+1}. This assumption of a CES production rather than of a Cobb-Douglas relationship, as in Borjas (1992), combines with (3) to enrich the model.[4] It is more convenient for the problem at hand to express this relationship in terms of C_t directly, instead of s_t, as the unknown decision variable:

$$H_{t+1} = \left[\eta(h_t - \frac{1}{R_t}C_t)^{1-\frac{1}{b}} + (1-\eta)(\upsilon_t)^{1-\frac{1}{b}}\right]^{\frac{b}{b-1}}, \tag{4}$$

where b, $b > 0$, denotes the elasticity of substitution between parental input and the social interaction effect in human capital production, and η, $0 < \eta < 1$, a parameter.

From the first order conditions for the maximization of (3) subject to (4), we have:

$$(H_{t+1})^{\frac{1}{\sigma}-\frac{1}{b}} = \frac{\eta\zeta}{(1-\zeta)R_t^{1-\frac{1}{\sigma}}} \frac{\left(h_t - \left[\frac{1}{\eta}(H_{t+1})^{\frac{b-1}{b}} - \frac{1-\eta}{\eta}\upsilon_t^{1-\frac{1}{b}}\right]^{\frac{b}{b-1}}\right)^{\frac{1}{\sigma}}}{\left[\frac{1}{\eta}(H_{t+1})^{\frac{b-1}{b}} - \frac{1-\eta}{\eta}\upsilon_t^{1-\frac{1}{b}}\right]^{\frac{1}{b-1}}}. \tag{5}$$

I develop the properties of the solution of (5) by working with its inverse, that is, with h_t as a function of H_{t+1} and υ_t. Simplifying by using the homogeneity of degree 0 of (5) yields:

$$h_t = H_{t+1}\eta^{-\frac{b}{b-1}}\left[1-(1-\eta)\left(\frac{\upsilon_t}{H_{t+1}}\right)^{\frac{b-1}{b}}\right]^{\frac{b}{b-1}} + \left(\frac{1-\zeta}{\eta\zeta}\right)^{\sigma}\eta^{-\frac{\sigma}{b}}R_t^{\sigma-1}H_{t+1}\left[1-(1-\eta)\left(\frac{\upsilon_t}{H_{t+1}}\right)^{\frac{b-1}{b}}\right]^{\frac{\sigma}{b-1}}. \tag{6}$$

3.1 Law of Motion of The Intertemporal Evolution of Human Capital

A child's human capital follows, from (6), as an implicit function of her parent's human capital and of the local interaction effect,

$$H_{t+1} = \mathcal{H}(h_t, \upsilon_t). \tag{7}$$

I shall refer to the above equation as the *law of motion*. My results are summarized in the propositions which follow. Proposition 1 deals with general properties of the law of motion. Proposition 2 deals with properties of specific cases, many of which have been utilized by the empirical literature

[4]Borjas (1992) works with a Cobb-Douglas version of (4):$H_{t+1} = (s_t h_t)^{\delta_1}\upsilon_t^{\delta_2}$, $\delta_1 + \delta_2 < 1$. Borjas' assumption of a Cobb-Douglas production function for child quality does exclude complementarity in the production of human capital [*ibid.* Equ. (3)]. Since the endogenous savings function may not be solved in closed form anyway, it is interesting to assume the most general framework that involves a CES functional form with complementarity in the production of the child's human capital and thus seek the most general results that might be possible within such a framework.

but had not been explored as following from the same general behavioral framework. Both proofs are elementary and therefore relegated to the Appendix.

Proposition 1. *The time map associated with the law of motion (7) of the evolution of human capital is a monotone increasing function of* (h_t, υ_t), *with* $H_{t+1} = (1-\eta)^{\frac{b}{b-1}} \upsilon_t$, *if* $h_t = 0$. *It is an increasing (decreasing) function of* R_t, *if* $\sigma \leq 1 (> 1)$. *In addition:*

1. *If* $\sigma \geq b$, *or* $\sigma < b$, *and* $1 > b - \sigma$, *then the time map is an increasing concave function of* h_t, *with* $\lim_{h_t \to \infty} \frac{\partial H_{t+1}}{\partial h_t} < \eta^{\frac{b}{b-1}} < 1$.

2. *If* $\sigma < b$, *and* $1 < b - \sigma$, *then in general there exists an interval, defined by a pair of threshold values for* R_t, $[R_{\min}(\upsilon_t), R_{\max}(\upsilon_t)]$, *within which the time map is an increasing S-shaped (sigmoid) function of* h_t, *for low values of* h_t.

3. *The time map is an increasing concave function of* υ_t, *if* $\sigma < b < 1$ *or* $\sigma > b > 1$. *If* $b > \sigma > 1$, *and* b *is sufficiently larger than* σ, *then the time map may become a convex function of* υ_t.

The essence of Proposition 1 may be summarized as follows. Given b, the elasticity of substitution in (home) production between own input and the neighborhood interaction effect, there exists a maximum value of σ, the elasticity of substitution in consumption between the offspring's human capital and own consumption, below which the law of motion is sigmoid. In other words, provided that $b > 1$, own input and neighborhood interaction input are not essential in the production of offspring human capital, the more complementary are own consumption and offspring human capital the more likely it is that the law of motion is sigmoid. This property is subject to a threshold value of the real wage rate and therefore is similar to the emergence of sigmoid time maps in overlapping-generations neoclassical growth models, that also involve thresholds [see Galor and Ryder (1991) and Azariadis (1993), p. 203]. However, there is an important difference from Galor and Ryder's finding, where emergence of a sigmoid map requires that the economy be sufficiently productive and that the elasticity of substitution between capital and labor in the aggregate production function be less than 1 (in which case capital and labor are essential). Proposition 1 generalizes in the form of maximum and minimum thresholds for productivity, and of a maximum threshold $b - 1$ for the elasticity of substitution between own consumption and offspring human capital. A notion of complementarity that rests on a comparison of preference and human capital production parameters heeds in part Matsuyama's point that complementarity should *not be assumed* but *derived* [Matsuyama (1995; 1996)]. The main model implies results similar to those of Ciccone and Matsuyama (1999). It is precisely because own input and neighborhood interaction input are not essential in the production of offspring human capital that we may have three positive steady states. In contrast, the fact that inputs are essential in the Galor-Ryder case gives rise to at most two equilibria. This result serves as a specific example of the general possibility of a sigmoid time map, recognized by Galor and Tsiddon (1997) and attributed by them to third derivative properties.[5]

As the above discussion and Figures 1 and 2 make clear, as R_t changes, the dynamical system (7) undergoes a *saddle-node bifurcation* [Azariadis (1993), p. 92]. The fact that the critical values depend upon the social interaction term υ_t makes clear that in general the dynamics of the model are quite complicated, especially when interactions are endogenous. An exhaustive analysis would be based on the properties of H_{t+1} as a function of υ_t, which may be obtained from (5). However, unlike (6), we cannot solve (5) in closed form for H_{t+1} as a function of υ_t, (nor its inverse), a fact that complicates such an analysis. Additional elements of complication are introduced if an individual's own human capital is related to the interaction effect υ_t, as is likely to be the case in the model of community selection that I briefly touch upon further below.

It is straightforward to work out the basic properties of the law of motion for special sets of values of the substitution elasticities. I summarize my results in

Proposition 2. *The law of motion (7) encompasses the following as special cases.*

1. *If* $\upsilon_t = 0$, *or* $\eta = 1$, *no interaction effects, the time map is linear in* h_t :
$$H_{t+1} = \eta^{\frac{b}{b-1}} \left(1 + \left(\frac{1}{\eta}\right)^{\frac{b-\sigma}{b-1}} \left(\frac{1-\zeta}{\eta\zeta}\right)^{\sigma} R_t^{\sigma-1} \right)^{-1} h_t. \tag{8}$$

2. *If the elasticity of substitution in consumption is greater than 1 and in human capital production equal to 1,* $\sigma > 1$ *and* $b = 1$, *then the time map, expressed by its inverse,*
$$h_t = H_{t+1}^{\frac{1}{\eta}} \upsilon_t^{1-\frac{1}{\eta}} + \left(\frac{1-\zeta}{\eta\zeta}\right)^{\sigma} R^{\sigma-1} (H_{t+1})^{\frac{\eta+\sigma(1-\eta)}{\eta}} \upsilon_t^{\sigma(1-\frac{1}{\eta})}. \tag{9}$$
is an increasing concave function of (h_t, υ_t). *It follows that* $H_{t+1} = 0$, *if* $h_t = 0$;

[5] I was unaware of Galor and Tsiddon (1997) at the time when first version of the paper was completed.

$\lim_{h_t \to \infty} \frac{\partial H_{t+1}}{\partial h_t} = 0.$

3. *If* $\sigma = 1$ *and* $b > 2$*, the elasticity of substitution in consumption is equal to 1 and in human capital production greater than 2, then the time map*

$$h_t = \left[\frac{1}{\eta}(H_{t+1})^{\frac{b-1}{b}} - \frac{1-\eta}{\eta} v_t^{\frac{b-1}{b}}\right]^{\frac{b}{b-1}} + \left(\frac{1-\zeta}{\eta\zeta}\right)(H_{t+1})^{1-\frac{1}{b}} \left[\frac{1}{\eta}(H_{t+1})^{\frac{b-1}{b}} - \frac{1-\eta}{\eta} v_t^{\frac{b-1}{b}}\right]^{\frac{1}{b-1}}. \tag{10}$$

may be a sigmoid curve and is independent of R_t*. It follows that* $H_{t+1} = (1-\eta)^{\frac{b}{b-1}} v_t$*, if* $h_t = 0$*;* $\lim_{h_t \to \infty} \frac{\partial H_{t+1}}{\partial h_t} < \eta$*.*

4. *If* $\sigma \to \infty$*, the utility function is linear, then the time map is given by*

$$H_{t+1} = (1-\eta)^{\frac{b}{b-1}} \left(\frac{\left(\frac{1-\zeta}{\zeta\eta} R_t\right)^{b-1}}{\left(\frac{1-\zeta}{\zeta\eta} R_t\right)^{b-1} - \eta} \right)^{\frac{b}{b-1}} v_t, \tag{11}$$

and is a function of v_t*, only, and thus independent of* h_t*.*

5. *If* $b \to \infty$*, the production function for human capital is linear, then the time map is given by*

$$H_{t+1} = \left(1 + \left(\frac{1-\zeta}{\zeta\eta}\right)^{\sigma} \eta^{1-\sigma} R_t^{\sigma-1}\right)^{-1} \left[\eta h_t + (1-\eta) v_t\right], \tag{12}$$

and is proportional to a convex combination of (h_t, v_t)*.*

6. *If* $\sigma = b$*, the elasticity of substitution in consumption is equal to that in human capital production then the time map is given by*

$$H_{t+1} = \left[\eta \left(\frac{1}{1 + \left(\frac{1-\zeta}{\eta\zeta}\right)^{\sigma} R_t^{\sigma-1}}\right)^{1-\frac{1}{\sigma}} h_t^{1-\frac{1}{\sigma}} + (1-\eta)(v_t)^{1-\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{13}$$

Borjas (1992) works with a case similar to (10) above, except that his Cobb-Douglas production function assumes decreasing returns to scale, whereas mine assumes constant returns to scale. Kremer (1997) does not assume a specific behavioral model, but the model he estimates, Equ. (1), is reminiscent of (12), which is a special case of our general model.

The law of motion for the intergenerational transmission of human capital implies a solution for a child's human capital, H_{t+1}, as a function of both the human capital of the parent, h_t, and of the interaction effect, v_t. An increase in the interaction effect shifts the entire map upwards. However, the impact of such an increase in the interaction effect upon the steady state depends whether or not we have multiple equilibria.

This basic model may be adapted to express different institutional environments via the specification of the relationship between one's human capital h_t and the interaction effect v_t. I return below to the role of the specification of v_t. The properties of $\mathcal{H}(\cdot)$ that I have developed allow broad predictions about the dynamic evolution of dynastic, i.e., a family's, human capital. The possibility that the time map has a sigmoid shape leads readily to a multiplicity of steady-state equilibria, where one of these equilibria will be unstable in certain circumstances. Nonlinear dynamics are caused by the dependence of offspring human on the interaction effect v_t. In the context of an ethnic capital interpretation of the model (in the style of Borjas), observing the occupational distribution among ethnic groups and applying the standard Roy model of occupational choice would make one conclude that that there are intrinsic skill differences across ethnic groups. In contrast, as Matsuyama (1995) notes, "a theory of pattern formation would suggest that there are some complementarities in the processes of skill acquisition, limited within an ethnic group, so that some small differences in skills or some random events happen to end up sorting different groups into different occupations" [*ibid.*, p. 68.] It is precisely such complementarities that my model seeks to articulate. The skill differences do not have to be very large. It suffices, according to Proposition 1, to have an unstable equilibrium that is symmetric (loosely speaking, although it could be rendered precisely so by means of suitable choice of parameter values) relative to two other stable but asymmetric equilibria. Such a variety of outcomes for the intergenerational evolution of human capital depends upon behavioral parameters and the parameters of the production function for human capital, on one hand, and initial conditions (h_0, v_0), on the other.

4 Empirical Results

Kremer (1997) works with data from the Panel Study of Income Dynamics (PSID), augmented by means of the geocodes. This additional information allows the researcher to identify the census tract in which a respondent lived when the interviews took place. US census tracts typically comprise approximately 5000 people. Education and other data by census tracts is publicly available and reported in terms of brackets [Adams (1991a)]. Such a linkage is made possible by means of tract identifiers, which are documented in Adams (1991b) and made available to investigators by means of special arrangements only. These identifiers allow one to link individuals to the "neighborhoods" of their residence.

I follow Kremer (1997) and define the sample within the PSID to include only individuals who have completed their education. I have selected individuals, male or female, who have been interviewed at least once and are at least 28 years of age by 1992 and both of whose parents have been interviewed at least once. That is, I work with the same age cutoff as Kremer (1997), except that he applies it to 1988 only while I extend it to 1992. Kremer defines the neighborhood of upbringing as the census tract where the child lived in 1968. For individuals who became 28 years of age after 1988 I use the 1980 Census Extract Data Sets. About 74% of the sample used in regressions are associated with neighborhood data from the 1970 Census, and the remainder are from the 1980 Census. In spite of this difference, the resulting samples are quite similar. In order to use self-reported data from parents and children I need to use split-offs from the original Panel Study of Income Dynamics (PSID) sample. This renders the sample no longer representative of the US population because it does not include observations added to the PSID to replace respondents lost to attrition. Moreover, the PSID oversamples the poor. I therefore weight observations, using the most recent set of PSID weights as suggested by the PSID staff [Hill (1992), just as Kremer (1997), p. 123].

I follow Kremer in defining the values of years of education to associate with each interval. These brackets and the respective means in 1970 and 1980 are as follows:

Age group	Mean 1970	Mean 1980	Value
Percent of age 25 or older with 0–8 years	24.96	16.70	6
Percent of age 25 or older with 9–11 years	20.32	15.99	10
Percent of age 25 or older with 12 years	31.30	33.65	12
Percent of age 25 or older with 13–15 years	11.59	16.63	14
Percent of age 25 or older with 16 + years	11.85	16.98	17

I am interested in exploring further the presence of nonlinearities in the relationship between individuals' education and those of parents and the distribution of education in the neighborhood where individuals were brought up.[6]

I follow Kremer in constructing neighbors' education. In a novel step, I use the frequency distribution for the educational attainment of males over twenty five years of age in the census tract in which the respondent grew up, to compute the first, second and third moments of the distribution of education in the census tract of respondents' residence. I believe this is the first such use of the frequency distributions for education within census tracts. It appears to be crucial for the appearance of significant nonlinear effects.

4.1 Parametric Estimations

My OLS regressions basically follow Equ. (1) and are reported in Table 1. It is worth recalling the fundamental identification issues which may in principle affect this model [Manski (1993)]. Here, as in Manski's Equ. (13), with "social forces act[ing] on the individual with a lag," the coefficient of neighbors' education (the social effect) may be identified if I assume that the process is observed out of equilibrium [*ibid.*, p. 540].

Column 1 reports Kremer's main regression for the purpose of comparison. Column 2, 4 and 6 report our results for Equ. (1) with a similarly defined sample, based on the random subsample of the PSID. Column 3 and 5 report our results with the entire PSID sample, which includes oversampling of the poor. All regressions are weighted with the appropriate weight. Columns 1 and 2 are quite similar, broadly speaking, although in my results the total effect of parents' education is numerically less important than that in Kremer, and the opposite is true for the average education in the neighborhood. I do not have a ready explanation for these differences.

A key prediction of Proposition 1, part 2, is that parents' education would have a sigmoid effect on children's education in the presence of an interaction effect. Kremer reports that inclusion of quadratic terms for parents' education and neighbors' education, and an interaction term for parents' and neighbors' education is not significant and an F-test cannot reject linearity.[7] I test this

[6] I am not aware of any previous work on nonlinear interactions between education of parents and their offspring. Eide and Showalter (1999) follow a quantile regression approach to studying the intergenerational evolution of earnings. In one of the regressions of sons' earnings against fathers' characteristics that they report, *ibid.,* Table 3, the quantile effects are roughly similar to the predicted relationships by Proposition 1, Part 2, above.

[7] Kremer also estimates a Markovian model, where he codes educational attainment in terms of six ordered categories and estimates probabilities that a child is in each category conditional on each parent's category. He repeats such an estimation for four educational categories of neighbors. In all cases, the computation of the steady-state distribution of education allows one to examine the impact of parental or neighborhood sorting. The standard deviation

prediction by including linear, quadratic and cubic terms for father's education and for mother's education. The results for the polynomial structure, which are reported in Table 1, Columns 3 and 5, respectively for the random subsample of the PSID include both father's and mother's education. A polynomial structure for father's education on its own is significant, but two of the terms lose their significance when the terms for mother's education are added. The results imply a sigmoid shape for the relationship between parents' education and children's education, with two of the terms for mother's education being statistically significant and implying an inflection point at 11.75. Mothers generally spend more time with their children, which enables them to instill more of their own values upon their offspring. Moreover, mothers' own educational attainments can also be seen as better proxies for "social class", especially at the time when the data were collected. Marital sorting suggests that education of spouses are related to one another, and such dependence may cloud the interpretation of the two different coefficients. The data reject the hypothesis that the linear terms for fathers' and mothers' education are equal. Nonetheless, imposing equality and including quadratic and cubic terms yields dynamics very similar to those implied by the terms for mother's education.

I test the nonlinearity of education within the neighborhood where an individual was brought up by including the second and third moments of the distribution of education within the appropriate census tract. This is statistically significant overall. The estimated coefficients are not statistically significant for the random sample (Column 4) but are more significant and do imply a nonlinear effect for neighbors' education for the entire PSID sample (Column 5).

Comparison of the signs of the coefficients of the polynomial terms for mother's education with those for the moments of neighbors' education suggests a puzzling asymmetry. The marginal effect of mother's education, which is quadratic, attains a maximum within the range of values, is positive for most of them but is ultimately decreasing. The marginal effect of neighbors' education, which is also quadratic, attains a minimum within the range of values, and is ultimately increasing. Similar differences are present in all econometric experiments that I performed with both samples. Overall, my results with the random subsample of the PSID are not as statistically significant as those with the entire sample, yet both sets of results are very similar.

---

of education increases only slightly if the proportion of male population whose spouses with the same education were to increase from .6 to .8, or if the correlation coefficient between neighbors' education were to increase from .2 to .4.

In a further attempt to explain the structure of interactions among neighbors I explore another, generally little known (even among researchers who have used the PSID geocoded data), feature of the PSID. That is, PSID employed cluster sampling techniques which result in several observations from each census tract, and within tracts additional groupings, to be referred to as sampling clusters. The number of observations per tract are fairly evenly spread between one and seven. However, there are about 2% of the sample that come from tracts that contribute more than ten observations each, with a maximum of thirty three. I reestimated the basic regressions reported in Columns 2 and 4 by allowing for a random effect associated with observations belonging to the same tract, but excluding the dummy indicating whether the neighborhood data come from the 1970 census. There are 227 clusters with an average of 3.9 observations per cluster. The estimated coefficients differ little from those reported, but the random effects structure is significant. The fraction of the variance that is explained by the random effect varies from 11.02%, for Column 2, to 10.83%, for Column 4.[8]

In a further attempt to probe the structure of nonlinearities in the intergenerational transmission of human capital, I estimate the model in logs and report the results on Columns 7 and 8, Table 1. The model in logs also addresses an important concern, namely the numerical significance of the coefficient for the dummy of whether 1970 neighborhood data are used. In contrast to the results in levels, the numerical effect on the mean is an order of magnitude smaller. Also particularly interesting are the results for the effect of parental education, which are both positive and have positive marginal effects and have sigmoid shapes. The terms for father's education are more significant than those in Column 4, although still insignificant at conventional levels of significance, and imply different dynamics. The marginal effect of the log of father's education is quadratic and ultimately increasing at an increasing rate. The terms for mother's education are all significant at conventional levels of significance, and imply similar dynamics to those of Column 4. The marginal effect of the log of mother's education is quadratic, attains a maximum and ultimately decreases. Again, I interpret this asymmetry as evidence of different mechanisms at work. The effect of paternal education may be more likely to reflect income, whereas the maternal one to reflect

[8] Additional evidence of the co-dependence of education of individuals who grew up in the same neighborhood is obtained by regressing an individual's education against neighborhood education and the mean education among all other individuals who grew up in the same neighborhood. The respective coefficients are .178 (6.92) and .564 (21.81), and $R^2 = .225$.

mothers' own values, in which case it would make sense to be more closely related to her own values. The inclusion of all nonlinear terms is significant. Allowing for a random effect associated with observations from the same sampling cluster, is significant and yields that the fraction of the variance that is explained by the random effect varies from 20.48%, for Column 7, to 18.84%, for Column 8.

I address the question of whether the impact of parental education and of the distribution of educational attainment within a relevant neighborhood is more general than that of either polynomial terms of parents' education or through the first three moments of the distribution of neighbors' education. I work with a CES specification for interaction effects that is similar to the one proposed by Benabou (1996b). Equ. (1) may be specified more generally as

$$\ell n H = a_0 + a_p \ell n \left(D h^{1-\frac{1}{\phi}} + (1-D) h_{\prime}^{1-\frac{1}{\phi}} \right) + \ell n v + \epsilon, \qquad (14)$$

and

$$v = \varpi(\mu_{\nu(i)}(\cdot)) \equiv \left(\sum_{j \in \nu(i)} \mu_{\nu(i)j} h_j^{1-\frac{1}{\iota}} \right)^{\frac{\iota}{\iota-1}}, \qquad (15)$$

where μ^j denotes the frequency of the value h^j within the distribution of educational attainment of population in neighborhood $\nu(i)$ at time t, and $\mu(\cdot)$ the entire distribution.[9]

Estimation of the above model with nonlinear interaction structures as in (14) and (15) by means of nonlinear least squares did not work well. The best of the results I obtained with both parental and neighborhood data are reported on Table 1, column 6, where parental educations enter in logs. Inclusion of a CES structure for neighborhood education according to (15) is marginally significant relative to the OLS case with parental education in logs, and their estimated coefficients differ little, and the actual estimate $\hat{\iota} = .626$ is not statistically significant. I obtained clearer results when I tried

[9] See Benabou (1996b). If $\phi < 0$, then parental educations are "complements," and the corresponding "isoquants" associated with the RHS of (14) (in levels) are concave. If $\phi > 0$, then parental educations are "substitutes," and the corresponding "isoquants" associated with the RHS of (14) (in levels) are convex. Similarly, if $\iota < 0$, then individual levels of educational accomplishments are "complements," ϖ is convex with respect to the educational attainment of the population, and diversity is a source of loss: $\varpi_{\nu(i)t} < \bar{h}_{\nu(i)t}$. If $\iota > 0$, then individual levels of educational accomplishments are "substitutes," ϖ is concave with respect to the educational attainment of the population, and diversity is a source of gain: $\varpi_{\nu(i)t} > \bar{h}_{\nu(i)t}$. Commonly made assumptions about interaction effects are implied as special cases. That is, if $\iota \to +\infty$, ϖ equals the arithmetic mean, $\varpi_{\nu(i)t} = \sum_{j\in_{\nu(i)t}} \mu_j h_{jt}$; if $\iota \to 1$, ϖ coincides with the Cobb-Douglas function and thus equals the geometric mean, $\varpi_{\nu(i)t} = \prod_{j\in_{\nu(i)t}} h_j^{\mu_j}$; As $\frac{1}{\iota}$ decreases from $+\infty$ to $-\infty$, ϖ spans the whole range of interaction technologies from Leontieff, or "weakest link – one apple spoils the bunch", $\varpi = \min_{j\in\mathcal{I}_{\nu(i)t}} \left\{ \frac{h_j}{\mu_{\nu(i)j}} \right\}$, which occurs when $\frac{1}{\iota} \to +\infty$, to "best shot," that is to role models where the best individual sets the standard, $\varpi = \max_{j\in\mathcal{I}_{\nu(i)t}} \{h_j\}$, which occurs when $\frac{1}{\iota} \to -\infty$.

to estimate nonlinear structures with parental education and neighborhood education in separate regressions. A significant CES structure for parental education was estimated with $\hat{\phi} = -.415$ (t-statistic of 2.45), and $\hat{D} = .634$ (t-statistic of 8.07), and $R^2_{\text{adj}} = .224$. A significant CES structure for neighborhood education was estimated with $\hat{\iota} = -.206$ (t-statistic of 1.91, significant at 5.6%), and $R^2_{\text{adj}} = .115$. Both these results imply that parental educations and neighbors' educations are "complements," when considered separately. These results are consistent with the notion that parents' education and neighbors' education serve as role models. All of the results taken together suggest that nonlinearities for both parents' education and neighbors' education are important.

4.2 Nonparametric Results

While the parametric results do provide support for the nonlinear effects of parental and neighborhood education, they do depend on restricted notions of nonlinearity. Next I turn to discussion of nonparametric estimations of stochastic kernels for various versions of Equ. (1). I use Danny Quah's tsrf program.[10]

To understand the construction of the stochastic kernel, consider the kernel showing the child education conditional on father's education, $H_i = G(h_i, \varepsilon_i)$, reported in Figures 3.a, b. To estimate that stochastic kernel, the program first derives a non–parametric estimate of the joint distribution $f(h_i, H_i)$. Then numerically integrate under this joint distribution with respect to H_i to get the marginal distribution of father's education $f(h_i)$. Next the conditional distribution $f(H_i|h_i)$ is estimated by $\hat{f}(H_i|h_i) = \frac{\hat{f}(h_i,H_i)}{\hat{f}(h_i)}$. Under regularity conditions, this gives us a consistent estimator for the conditional distribution for any value of father's education h_i.

The stochastic kernels plot this conditional distribution of H_i (`maxed`) for all values of h_i (`dad`), Figure 3.a, and the corresponding contours are given in Figure 3.b. Figures 4.a and 4.b give a glimpse at marital sorting in terms of education by reporting kernel estimates of mother's education (`mom`) conditional on the father's $f(h_i'|h_i)$. In view of the complexity of marital sorting, I examine the dependence, due to selection, between average neighborhood education (`nschup`) and average parents' education (`parent`), with the stochastic kernels for $f(h_{\nu(i)}|\frac{1}{2}(h_i + h_i'))$ being reported in Figures 5.a, b.

[10]The program is available at http://econ.lse.ac.uk/~dquah/tsrf.html.

Figures 6.a, b report the stochastic kernels for $\hat{f}\left(H_i|\frac{1}{2}(h_i + h_i')\right)$, and Figures 7.a, b report the stochastic kernels for $\hat{f}\left(H_i|h_{\nu(i)}\right)$. This last set of estimations provide persuasive evidence of the presence of nonlinear effects of parents' and neighborhood education. To appreciate that, consider drawing a curve connecting the modes in Figures 6.b and 7.b. The resulting curves once transposed look very similar to Figures 1 and 2. This is particularly interesting finding, given that the theoretical predictions admit many possibilities. [11]

5 Applications

I offer next some applications of the basic model. One explores the properties of distribution of human capital. A second addresses the dynamics of the income distribution and the Kuznets hypothesis. A third allows individuals to choose the community of their residence and therefore neighborhood interactions as well.

5.1 Dynamics of the Distribution of Human Capital

I assume that the entire population benefits from the same interaction effect. I modify the original model slightly. Individuals live for two periods. In the first period of their lives, they are supported by their parents. In the second period, they give birth to offspring (one each) and use some of their labor to bring up their offspring. They supply the remainder of their labor to the labor market, receive earnings and spend them on consumption. Such a simple overlapping-generations model is sufficient for the essential elements of the model [Durlauf (1996c)]. I assume an interaction effect as in (15).

Let aggregate output be produced by labor only by means of a constant returns to scale aggregate production function: $C_t = A\sum_{i\in\mathcal{I}}(1 - s_{it})h_{it}$. Consequently, the equilibrium wage rate is given by $R_t = A$.

If one were to assume the conditions of Proposition 2, part 5, which yield a child's human capital

[11] I also explored the fact that the data on neighborhood education is reported in terms of frequencies. While I have not found a way to build this directly into the nonlinear estimations reported above, I explored the dependence of within-neighborhood dispersion on average neighborhood education. A bit to my surprise, there is relatively little variation of within neighborhood dispersion across neighborhoods: its mean standard deviation is 3.12 years, the standard deviation of the standard deviations across all tracts in the data is .303, and the respective minimum and maximum values are 2.426 and 4.362.

as a linear convex combination of that of the parent's and of the interaction effect, then one could show that however strong the interaction effect might be in the direction of implying diversity as a source of gain, the distribution of human capital will collapse to a single point in the long run. Thus, I concentrate on the case of child's human capital being a nonlinear function of that of the parent's and of the interaction effect.

Recall that under the conditions of Proposition 1, part 2, the time map would be a sigmoid function for low values of h_t. Under those same conditions, it isan increasing concave function of v_t. It follows that in every t, there exist three intersections with the time map with 45^o-degree line, of which the middle would correspond to an unstable equilibrium for the dynamic evolution of human capital. With $R_t = A$, the time map given by (7) is time invariant. Therefore, if the middle of the three intersection points is contained in the interior of the support of the initial distribution, then only the stable equilibria will survive in the steady state [Arthur *et al.* (1994)]. Theorem 5.1, *ibid.,* p. 193, ensures that the stable equilibria will be attained in the steady state with positive probability. Theorem 5.2, *ibid.,* p. 195, ensures that the unstable equilibrium will be attained in the steady state with zero probability.

A definite statement requires that one examines the role of the interaction term. Since $\mathcal{H}$ is increasing concave in v, it has an equalizing effect on human capital, unless diversity is a source of gain. One would expect that for certain parameter values the equalizing effect of the interaction term, due to concavity, would be offset by the disequalizing effect due to social interactions. Even when diversity is a source of loss, the equalizing effect may be offset by the strength of the instability.

Let $\mathcal{H}_{(1)}(v)$ and $\mathcal{H}_{(2)}(v)$ be the two stable fixed points of $\mathcal{H}(h, v)$, as functions of v, and let μ_1, μ_2 be the corresponding population proportions with $\mu_1 + \mu_2 = 1$. At equilibrium,

$$v^{1-\frac{1}{\iota}} = \mu_1 \mathcal{H}(v)_{(1)}^{1-\frac{1}{\iota}} + \mu_2 \mathcal{H}_{(2)}(v)^{1-\frac{1}{\iota}},$$

must be satisfied. Therefore, if the preference for diversity is not too strong, that is, $|\iota|$ is not very large, then the r.h.s. of the above equation is concave in v and the l.h.s. is convex. Therefore, for given (μ_1, μ_2), a unique equilibrium value of the interaction effect exists and is unique, and the equilibrium values of human capital follow.

Because of the global nature of the interaction effect here, no mechanism exists in this model, through which it may be evaluated, and the population proportions are indeterminate. Therefore, it is interesting to invoke a planner who chooses the population proportions so as to maximize average utility. In the homogeneous case, when all individuals have the same amount of human capital, the interaction effect is also equal to that. If diversity is a source of gain, by creating heterogeneity, the planner could cause the interaction effect to improve utility above what would be its value in the homogeneous case. With this intuition, the heterogeneity associated with the solution to the planner's problem could be interpreted as resulting from competition. across the population at the steady state. Such a planner's problem may be stated as follows. Let $V(\mathcal{H}_{(j)}(v),\ j = 1,2$ denote the indirect utility for each type. The planner's problem is to choose $(\mu_1, \mu_2),\ \mu_1 + \mu_2 = 1,\ \mu_1, \mu_2 > 0$, so as to maximize $\mu_1 V\left(\mathcal{H}_{(1)}(v)\right) + \mu_2 V\left(\mathcal{H}_{(2)}(v)\right)$, subject to the definition of the interaction effect above. The corresponding first-order condition is:

$$\mu_1 V'\left(\mathcal{H}_{(1)}(v)\right) \mathcal{H}'_{(1)} \frac{\partial v}{\partial \mu_1} = \mu_2 V'\left(\mathcal{H}_{(2)}(v)\right) \mathcal{H}'_{(2)} \frac{\partial v}{\partial \mu_2}.$$

This discussion highlights two important issues pertaining to the persistence of a non-trivial distribution of human capital. One, residential choice is critical for such a persistence; and two, the distribution reflects the nature of social interactions.

5.2 The Kuznets Hypothesis

I consider next the case with an exogenously growing A_t. That is, labor productivity in output production increases, while the productivity of human capital as an input to its own production remains stagnant. As $R_t = A_t$ increases, and depending upon the value of v_{it}, it will at some finite time reach the interval of threshold values $[R_{\min}, R_{\max}]$. Therefore, as the economy evolves, it will inevitably get into a temporary ("local") tendency for the distribution of human capital to become polarized – which confirms the Kuznets hypothesis: the distribution of income during economic growth will worsen before it improves [Kuznets (1955)].[12]

Depending upon whether $\sigma < 1$ or $\sigma > 1$, the economy will ultimately emerge below or above the area where the time map is sigmoid. Depending upon parameter values, the economy could cycle locally. Finally, I note that even in the case of the previous paragraph when $R_t = A$ and thus constant, and if $A \ni [R_{\min_0}, R_{\max_0}]$, the dependence of the threshold values upon the interaction

[12] Glomm (1997), the latest review of the evidence on Kuznets hypothesis, emphasizes that while Kuznets' "inverted U curve" is consistent with the facts for some countries, it does not describe the evolution of income distribution everywhere.

term v_t implies that this region "moves" and could encompass A in finite time. The economy would again satisfy the Kuznets hypothesis. Which of the two cases of Proposition 1 applies depends upon upon the magnitude of σ relative to 1 and initial conditions.

I note here the similarity of my finding with that by Galor and Tsiddon (1997), which I was unaware of at the time of first version of the paper was completed. My finding of a Kuznets curve is similar to that of Galor and Tsiddon (1997) but different from that by Glomm and Ravikumar (1998), which rests on short-run increasing returns to scale in learning technology. My finding, like that of Galor and Tsiddon (1997), rests on the complementarity that renders sigmoid the time map of the intertemporal evolution of human capital, and which is endogenous, in that it is defined in terms of certain conditions on preference and human capital production function parameters.

5.3 Neighborhood Choice

Neighborhood effects function as an externality in this model, which could be internalized through choice of community by individuals. Here I consider the same model as in the previous subsections. Individuals' human capital depend on parents' human capital and on the neighborhood effect in the community of their residence, which was chosen by their parents.[13]

Let $\ell_{i,t+1}$ denote the community where individual i, born at time t, chooses to reside when her child is born in period $t+1$, $\mathcal{L} \equiv \{1, \ldots, L\}$, the set of communities in the economy, and $\{\bar{q}_1, \ldots, \bar{q}_L\}$, the land areas they occupy, which are assumed to be given. Individuals choose where to locate so as to maximize utility. I close the model by assuming that each individual receives in the beginning of the second period of her life a lump sum transfer, $\bar{p}_{t+1}$, equal to the aggregate housing rents (across all communities) per capita.

The modification of the behavioral model to allow for land is as follows. Conditional on the choice of a community of residence, individual i's problem is the same as before, except that consumption consists of housing and non-housing consumption, (Q_{t+1}, F_{t+1}), which is evaluated

[13] This model combines essential features of the models in Benabou (1996a, 1996b). As a model of growth resembles the overlapping generations model in the appendix of Benabou (1996a), although it is somewhat more standard and include assumptions about the acquisition of human capital which differ from his. E.g., Benabou assumes that local spillovers exhibit a threshold effect. The behavioral model in Benabou (1996a) is more general, but is explored fully in only a two-period setting. Benabou (1996b), on the other hand explores a richer pattern of interactions, which include both global and local linkages, which are limited to the Cobb-Douglas functional forms only. The model here is conceptually very similar to Durlauf (1996a, 1996b) but involves only deterministic tools.

by a subutility function for consumption, $C_{t+1} \equiv \chi^{-\chi}(1-\chi)^{-(1-\chi)} Q_{t+1}^{\chi} F_{t+1}^{1-\chi}$, where parameter χ satisfies $0 < \chi < 1$. With non-housing consumption as the numeraire, and $p_{\ell,t+1}$ the rental rate of housing in community ℓ, labor earnings are denoted by E_{t+1} and consumption expenditure is written as $E_{t+1} + \bar{p}_{t+1} = p_{\ell,t+1} Q_{t+1} + F_{t+1}$. The corresponding indirect consumption subutility function is given by $C_{t+1} \equiv p_{\ell,t+1}^{-\chi}(E_{t+1} + \bar{p}_{t+1})$, in the absence of tax. The individual pays a tax on his housing expenditure at a community-specific rate $\tau_{\ell,t+1}$. As a result, it is the after tax expenditure, $(1 - \chi\tau_{\ell,t+1})(E_{t+1} + \bar{p}_{t+1})$, that enters the indirect consumption subutility function.

By using the consumption subutility C_{t+1} in the utility function (3), utility is expressed in terms of offspring human capital and consumption expenditure,

$$U \equiv \left[\zeta (H_{t+1})^{1-\frac{1}{\sigma}} + (1 - \zeta_{\ell,t+1}) \left(E_{t+1} + \bar{p}_{t+1}\right)^{1-\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{16}$$

where $1 - \zeta_{\ell,t+1} \equiv (1-\zeta) \left(p_{\ell,t+1}^{-\chi}(1 - \chi\tau_{\ell,t+1})\right)^{1-\frac{1}{\sigma}}$. Conditional on the choice of community ℓ, an individual is assumed to maximize her utility function (16), subject to her nonlinear budget constraint (4), except that E_{t+1} takes the place of C_t.

Summarizing, an individual i's choice of community ℓ at time $t+1$ is reflected upon the analytics through: first, the community-specific "taste" parameter $1 - \zeta_{\ell,t+1}$, defined in above, and second, the neighborhood interaction effect in the production of offspring human capital $v_{\ell,t+1}$ which also depends on the community.

6 Conclusions

I develop here a model of the evolution of human capital as a result of individual choice. The model implies a law of motion for the evolution of human capital with multiple equilibria that differ in terms of stability properties.

This feature of the model is entirely dependent on two features: first, nonlinear neighborhood interactions must be present; second, the elasticity of substitution between own human capital and the neighborhood interactions must be sufficiently larger than the elasticity of substitution between one's child's human capital and own consumption. Neighborhood interactions in general represent the role of public education in the production of human capital. I show that under certain conditions the economy segregates into two types of human capital. The model also supports the

Kuznets hypothesis, namely that the income distribution would worsen before it improves during the process of economic growth. An extension of the model allows one to bring together two strands of the literature, those emphasizing the role of neighborhood effects in growing economies with those examining equilibria in economies with local public goods where individuals have a choice where to locate, and therefore they may segregate according to human capital.

I report empirical findings, which are based on geocoded data from the Panel Study of Income Dynamics and are similar to Kremer's findings of substantial (linear) effects on a person's human capital, measured by years of education, of the average education in the neighborhood where he or she grew up. My results also show that a person's education is nonlinearly related to both father's and mother's education, when both are present in the regression, and to the mean, and the second and third moments of the distribution of education within the neighborhood where an individual was brought up. These findings, and nonparametric estimates that I also report, key predictions of the theory regarding r certain conditions, the relationship between a child's education and that of his or her parent has a sigmoid shape.

7 References

Adams, Terry K. (1991a), "Excerpts from Documentation for 1970 and 1980 Census Extract Datasets," Economics Program, Institute for Social Research, University of Michigan, July.

Adams, Terry K. (1991b), "Excerpts from Documentation for 1968-1985 PSID-Geocode Match Files," Economics Program, Institute for Social Research, University of Michigan, August.

Arthur, W. Brian, Yuri M. Ermoliev, and Yuri M. Kaniovski (1994), "Path Dependence Processes and the Emergence of Macrostructure," Ch. 3, 33–48, and "Strong Laws for a Class of Path-Dependent Stochastic Processes," Ch. 10, 185–201, in Arthur W. Brian, *Increasing Returns and Path Dependence in the Economy,* University of Michigan Press, Ann Arbor.

Azariadis, Costas (1993) *Intertemporal Macroeconomics,* Blackwell Publishers, Oxford.

Azariadis, Costas, and Allan Drazen (1990), "Threshold Externalities in Economic Development," *Quarterly Journal of Economics,* CV, May, 501–526.

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

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

Borjas, George J. (1992), "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics,* February, 123–150.

Borjas, George J. (1995), "Ethnicity, Neighborhoods, and Human Capital Externalities," *American Economic Review,* 85, 3, 365–390.

Ciccone, Antonio, and Kiminori Matsuyama (1999), "Efficiency and Equilibrium with Dynamic Increasing Returns due to Demand Complementaries," *Econometrica,* 67, 3, 499–525.

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

Durlauf, Steven N. (1996b), "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," in W. Barnett, G. Gandolfo and C. Hillinger, *Dynamic Disequilibrium Modelling,* Cambridge University Press.

Eide, Eric, R., and Mark H. Showalter (1999), "Factors Affecting the Transmission of Earnings across Generations: A Quantile Regression Approach," *The Journal of Human Resources,* 24, 2.

Galor, Oded, and Harl E. Ryder (1991), "Dynamic Efficiency of Steady-State Equilibria in an Overlapping-Generations Model with Productive Capital," *Economics Letters,* 35, 385–390.

Galor, Oded, and Daniel Tsiddon (1997), "The Distribution of Human Capital and Economic Growth," *Journal of Economic Growth,* 2, 93–124.

Glomm, Gerhard (1997), "Whatever Happened to the Kuznets Curve? Is It Really Upside Down?" *Journal of Income Distribution,* 7, 1, 63–87.

Glomm, Gerhard, and B. Ravikumar (1998), "Increasing Returns, Human Capital, and the Kuznets Curve," *Journal of Development Economics,* 55, 353–367.

Kremer, Michael (1997), "How Much Does Sorting Increase Inequality," *Quarterly Journal of Economics,* February, 115–139.

Kuznets, Simon (1955), "Economic Growth and Income Inequality," *American Economic Review,* 45, 1–28.

Manski, Charles F. (1993), "Identification of Endogenous Social Effects," The Reflection Problem," *Review of Economic Studies,* 60, 531-542.

Matsuyama, Kiminori (1995), "Comment on Paul Krugman's 'Complexity and Emergent Structure in the International Economy,' " in Levinson, Jim, Alan V. Deardorff and Robert M. Stern (1995), *New Directions in Trade Theory,* University of Michigan Press, Ann Arbor, 53–69.

Matsuyama, Kiminori (1996), "Why Are There Rich and Poor Countries? Symmetry-Breaking in the World Economy," *Journal of The Japanese and International Economies,* 9, 419–439.

FIGURES 1 and 2

Table 1: Intergenerational Transmission of Human Capital

Dependent Variable: Educational attainment at 28 years of age (years). Sample restricted to individuals whose parents report their own educational attainment. $t-$ statistics in parentheses. All regressions are weighted by latest PSID weights. + denotes the estimate of the coefficient of the interaction term.

Column 1 reports Kremer's main regression for the purpose of comparison. Column 2 and 4 reports our results with a similarly organized sample, which is based on the original random sample of the PSID. Column 3, 5 and 6 reports our results with the entire PSID sample. Columns 7 and 8 report results with the same sample os that for Columns 2 and 4 but the dependent variables and all regressors (except for the dummy for 1970 neighbors data) are in logs.

Model	1	2	3	4	5	6	7	8
Observations	880	881	1764	885	1764	881	881	881
$R^2_{\text{adj.}}$	.231	.2385	.2212	.2624	.2033		.2219	.2411
F		70.28	126.27	32.48	62.62		63.80	35.98
LLF		-3463.12				-3462.36		
Mean dep. var.	13.18	13.96	13.75	13.96	13.75	2.62	2.62	2.62
Intercept	6.96 (7.48)	6.38 (10.83)	7.24 (18.31)	7.19 (.79)	5.05 (0.80)	2.81 (.98)	1.32 (12.58)	1.56 (3.96)
Father's educ (12.11, 3.85)	.288 (7.38)	.192 (7.29)	.186 (10.33)	-.353 (1.39)	-.263 (1.50)	.152 (6.96)	.164 (6.89)	.671 (1.08)
Father's education squared				.035 (1.26)	.027 (1.44)			-.436 (1.36)
Father's education cubed				-.0006 (.67)	-.0004 (.66)			.092 (1.74)
Mother's educ (12.9, 2.85)	.154 (2.85)	.166 (4.79)	.139 (5.88)	-.643 (1.11)	-1.184 (5.10)	.110 (3.95)	.127 (4.21)	-.573 (1.99)
Mother's education squared				.109 (2.06)	.141 (5.84)			.483 (2.38)
Mother's education cubed				-.004 (2.51)	-.004 (5.72)			-.091 (2.29)
Neighbors' educ (11.29, 1.53)	.150 (2.08)	.232 (4.46)	.191 (5.27)	1.916 (.71)	3.24 (1.68)	$.159^{+}$ (.43)	.224 (4.91)	.193 (4.20)
Neighbors' educ second moment				-.229 (.91)	-.344 (1.89)			
Neighbors' educ third moment				.008 (1.13)	.011 (2.15)			
1970 neigh'rs data (.739, 1.532)		.344 (2.58)	.365 (3.81)	.350 (2.63)	.348 (3.68)		.027 (2.66)	.027 (2.61)
Interaction parameter (ι)						.626 (1.17)		

Proof of Proposition 1.

The first derivative with respect to H_{t+1} of the r.h.s. of (6) is given by:

$$\frac{\partial h_t}{\partial H_{t+1}} = \eta^{-\frac{b}{b-1}} \left[1-(1-\eta)\left(\frac{v_t}{H_{t+1}}\right)^{\frac{b-1}{b}}\right]^{\frac{1}{b-1}}$$

$$+\,\eta^{-\frac{\sigma}{b-1}}\left(\frac{1-\zeta}{\eta\zeta}\right)^{\sigma} R_t^{\sigma-1}\left[1-(1-\eta)\left(\frac{v_t}{H_{t+1}}\right)^{\frac{b-1}{b}}\right]^{\frac{\sigma}{b-1}}\left(1+\frac{\sigma}{b}\frac{(1-\eta)\left(\frac{v_t}{H_{t+1}}\right)^{\frac{b-1}{b}}}{1-(1-\eta)\left(\frac{v_t}{H_{t+1}}\right)^{\frac{b-1}{b}}}\right), \qquad (17)$$

and is positive. Therefore, h_t is an increasing function of H_{t+1}. The first term in the r.h.s. of (6) is an increasing convex function of H_{t+1}, defined for $H_{t+1} \geq (1-\eta)^{\frac{b}{b-1}} v_t$. The curvature properties of the second term in the r.h.s. of (6) depend upon parameter values and the magnitude of H_{t+1}.[14]

Specifically, by differentiating (17) with respect to H_{t+1} we have:

$$\frac{\partial^2 h_t}{\partial H_{t+1}^2} = \frac{\frac{1-\eta}{\eta} v_t^{\frac{b-1}{b}}}{\frac{1}{\eta}(H_{t+1})^{\frac{b-1}{b}} - \frac{1-\eta}{\eta} v_t^{\frac{b-1}{b}}}$$

$$\times \left\{ \frac{1}{\eta b} H_{t+1}^{-1-\frac{1}{b}} \left[\frac{1}{\eta}(H_{t+1})^{\frac{b-1}{b}} - \frac{1-\eta}{\eta} v_t^{\frac{b-1}{b}}\right]^{\frac{1}{b-1}} + \frac{\sigma}{b}\left(\frac{1-\zeta}{\eta\zeta}\right)^{\sigma} R_t^{\sigma-1} H_{t+1}^{-\frac{\sigma}{b}} \left[\frac{1}{\eta}(H_{t+1})^{\frac{b-1}{b}} - \frac{1-\eta}{\eta} v_t^{\frac{b-1}{b}}\right]^{\frac{\sigma}{b-1}} \right.$$

$$\left. \times \left(H_{t+1}^{-1} - \frac{b-1}{b} \frac{\frac{1}{\eta} H_{t+1}^{\frac{b-1}{b}}}{\frac{1}{\eta}(H_{t+1})^{\frac{b-1}{b}} - \frac{1-\eta}{\eta} v_t^{\frac{b-1}{b}}} \right) \right\}. \qquad (18)$$

Therefore, if $\sigma \geq b$, h_t is an increasing convex function of H_{t+1}, for $H_{t+1} \geq ((1-\eta))^{\frac{b}{b-1}} v_t$, which implies that H_{t+1} is an increasing concave function of h_t. This is depicted in Figure 1.

If, on the other hand, $\sigma < b$, the second term within the braces in the r.h.s. of (18) will be negative for sufficiently low values of H_{t+1}. It is thus possible for that second term to dominate the first in absolute value, conferring a sigmoid shape to the time map.

The two possibilities when the time map is sigmoid have been drawn on: Figure 1, for the case of $\sigma > 1$; and, on Figure 2, for the case of $\sigma < 1$.

The threshold value of $R_{\max}(v_t)$, for the case of $\sigma > 1$, is obtained by eliminating $h = h_t = H_{t+1}$ between (6) and (17), where I set $\frac{\partial h_t}{\partial H_{t+1}} = 1$, and ensure that the r.h.s. of (18) is positive ($\mathcal{H}(h_t, \cdot)$ is concave). The threshold value of $R_{\min}(v_t)$ is obtained by eliminating $h = h_t = H_{t+1}$ between (6) and (17), where I set $\frac{\partial h_t}{\partial H_{t+1}} = 1$, and ensure that the r.h.s. of (18) is negative ($\mathcal{H}(h_t, \cdot)$ is convex). It can be shown that both threshold values $R_{\max}$ and $R_{\min}$ are decreasing functions of v_t. See Figure 1.

For the case of $\sigma < 1$, the conditions are reversed. The threshold value of $R_{\max}(v_t)$, for the case of $\sigma < 1$, is obtained by eliminating $h = h_t = H_{t+1}$ between (6) and (17), where I set $\frac{\partial h_t}{\partial H_{t+1}} = 1$, and ensure that the r.h.s. of (18) is negative ($\mathcal{H}(h_t, \cdot)$ is convex). The threshold value of $R_{\min}(v_t)$

[14] Note that I am interested in drawing conclusions for the curvature properties of the time map but I will be working with its inverse. Therefore, concavity of the time map requires convexity of its inverse and so on.

is obtained by eliminating $h = h_t = H_{t+1}$ between (6) and (17), where I set $\frac{\partial h_t}{\partial H_{t+1}} = 1$, and ensure that the r.h.s. of (18) is positive ($\mathcal{H}(h_t, \cdot)$ is concave). It can be shown that both threshold values $R_{\max}$ and $R_{\min}$ are increasing functions of v_t. See Figure 2.

Regarding Part 3, by totally differentiating (6) we have $\frac{\partial \mathcal{H}}{\partial v_t} = -\frac{\partial h_t}{\partial v_t} / \frac{\partial h_t}{\partial H_{t+1}}$. This yields:

$$\frac{\partial \mathcal{H}}{\partial v_t} = (1-\eta)\frac{\left(\frac{v_t}{H_{t+1}}\right)^{-\frac{1}{b}}}{1 + (1 - \frac{\sigma}{b})\frac{1}{R_t^{1-\frac{\sigma}{b}}\frac{1}{\eta}\left[\frac{1}{\eta} - \frac{1-\eta}{\eta}\left(\frac{v_t}{H_{t+1}}\right)^{\frac{b-1}{b}}\right]^{\frac{1-\sigma}{b-1}} + \frac{\sigma}{b}\frac{1}{1-(1-\eta)\left(\frac{v_t}{H_{t+1}}\right)^{\frac{b-1}{b}}}}} \tag{19}$$

By inspection, it follows that if $\sigma > b > 1$, or $\sigma < b < 1$, $\mathcal{H}$ is a concave function of the interaction effect. In all other cases, the sign of the second derivative of $\mathcal{H}$ is ambiguous. In those other cases, if $\sigma << b$, and R_t is sufficiently large, the denominator of (19) dominates the magnitude of the derivative.

Q.E.D.