1994.

"Neighborhood effects on belief formation and the distribution of education and income"

by

John E. Roemer *
and
Roger J-B Wets*

-mail: JEROEMER@UCDAVIS.EDU fax: 916-752-9382

---

* Departments of Economics and Mathematics, respectively, University of California at Davis, CA 95616. We are grateful to Glenn Loury for a useful discussion, and for his contribution to Appendix 1, and to Robert Triest for suggestions.

Abstract: "Neighborhood effects on belief formation and the distribution of education and income"

Authors: John E. Roemer and Roger J-B Wets
Departments of Economics and Mathematics
University of California
Davis CA 95616

We study a society with a continuum of families, segregated in neighborhoods perfectly by income. There is a deterministic, non-linear relationship between years of education attained in youth and earnings in adult life. Youths choose years of education to acquire in order to maximize a utility function whose arguments are income as an adult and years of education attained, where education is experienced as costly. Talent is randomly distributed at each generation, and high talent reduces the welfare cost of attaining education: otherwise, all utility functions are identical. Youths do not know the true returns-to-education function, but form rational beliefs about that function by observing the education-income pairs of adults in their neighborhood, and making linear extrapolations. Our task is to compute the stationary distribution of income and education in this society. In that distribution, we show that, at many levels of income, there is a broad distribution of talent. Thus, income earned (or education attained) are poor signals for talent (or potential) when neighborhoods are segregated according to income.

JEL Categories: I2, D30, D60

1. Introduction

Following the sociological work of Wilson[1987], Jencks[1993] and others, a small literature has emerged in economics studying the generation of persistent inequality among a population due to neighborhood effects of various kinds. These effects all have the consequence of inducing sub-optimal levels of education for a group of the population, who then earn low incomes. The neighborhood effects consist of three types, which we shall call investment, role-model, and belief. *Investment* effects occur when the poor are segregated in a community; because the tax base is small, funding of local education is low, and hence children receive less education than in richer communities. Durlauf [1992,1993] has studied this problem, and Benabou [1991] also has an investment model in which there are positive externalities from being a student among other students who have enjoyed high inputs of family-specific human capital. *Role-model* effects occur when children do not observe high-flying adults of their type (let us say, African-American) and therefore conclude that high-flying lives are infeasible for them. This problem is discussed more by sociologists than economists. The third category, *belief* effects, is not clearly analytically distinct from the second; children form beliefs about the efficacy of education from observations which generally constitute a biased sample, due to neighborhood segregation by income. Streufert [1991] proposes a model in which students form inaccurate estimates of the marginal product of effort in school, due to the biased sample of adults they observe.

Our study falls in the third category. We shall assume that families are perfectly segregated by income, and that children form rational conjectures concerning the relationship of future income to education based upon observing the income-education pairs of adults in their neighborhood. Children all have the same preferences over education and income, but they differ with respect to talent, a trait that affects the disutility of education to a child. (Talent, however, has no direct effect on income as an adult, which is a deterministic function of years of education achieved only.) Talent of children is randomly distributed at each level of parental income; we shall assume that this distribution is identical at all income levels. The corresponding interpretation would be that talent is some inborn trait, which is independent of the parent's talent[1].

In our model, there is a deterministic, monotonic relationship between years of education and income in later life. Our task shall be to calculate the long-run stationary bivariate distribution of income and talent. Since all children have the same utility function $u(y,x;\ s)$ where y is income, x is years of education, and s is the talent parameter, there would be a perfect monotonic relationship between income and talent if everyone knew the true relationship of income to education. But with the incorrect but rational beliefs that

---

[1]Alternatively, one might assume that the distribution of talent among children varies with parental talent or parental income. In the former case, the interpretation would be that talent is an inherited or culturally acquired characteristic; in the latter case, it would be that talent is an acquired characteristic. We have chosen to assume that the distribution of talent is independent of parental talent and income for several reasons: first, we wish to examine the consequences for long-run income inequality in the extreme case that the cohorts of children at all parental levels of income are statistically identical (i.e., have the same preferences and distribution of talent); second, the pure relationship between parental and child IQ (if that is a proxy for talent) is not well-established; third, to link an acquired talent to parental income would muddy the waters, in the sense that we are trying to calculate the effect of parental income on educational choices of children as articulated through the observations children make about the returns to education.

children shall form in our model about the returns to education, there will, in the stationary distribution, be a non-degenerate distribution of talent levels at each level of income. If, in this stationary distribution, the variance of talent is quite large at a particular income level, then we conclude that one cannot accurately infer the talent (or, let us say, potential) of a person by observing her level of income.

Indeed, this is the conclusion to which our analysis shall lead us. We show that under our assumptions there is a quite fat spread of talent, in the stationary income-talent distribution, at each level of income (or education). This means considerable inefficiency exists due to the beliefs induced in children by neighborhoods which are segregated by income. Alternatively put, education and income (which in our model are perfectly correlated) are poor signals for talent.

In the next section, we present our model. In section 3, we calculate the stationary income-talent distribution gotten by parameterizing the model using U.S. data for white and black income-education relationships. Section 4 concludes.

2. The model

We study a multi-generation model consisting of children who form their beliefs about the monetary returns to education by observing adults in their community. Communities are perfectly segregated according to income of adults. Each person lives for two periods, and gives birth to one child in the second period of his life. The child chooses her educational level by maximizing a utility function where years of education are a cost and income as an adult is the benefit. All children have the same utility function $u(y,x;s)$ where y is income as an adult, x is years of education, and s is the 'talent' of the child.

Utility is increasing in y, decreasing in x, and increasing in s. (Think of a larger s as making a given amount of education less costly to acquire.) Talent has no effect on earnings, other than indirectly, through its effect on years of education attained.

The true relation of income to education is given by a (non-stochastic) function $\Phi(x)=y$. Children and parents do not know the function Φ. A child, however, forms rational beliefs, in this sense: she estimates the returns to education as the best linear regression of the income against education of the parents in her neighborhood. As neighborhoods are perfectly segregated by income, and as there is a continuum of families, we can mathematically represent a child's linear regression as the <u>tangent line to the point characterizing the education-income pair</u> of her parent, (x_0,y_0), as depicted in Figure 1.

We note that, unlike the Streufert model referred to earlier, children do calculate the correct <u>marginal</u> returns to education, calculated at their neighborhood's prevalent level of education. Children in our model are only incorrect because the true returns to education are non-linear. In fact, average income in the U.S. is a convex function of years of education until there are many such years; see Figures 2ab below, which graph the (average) income of white and black male adults in the U.S. in 1990 against years of education attained. Under the postulated Φ of Figure 1, it is clear that poor students who behave as we have postulated would believe that there are low marginal returns to education, while children from rich families would believe there are high marginal returns. This mistake children make is in assuming that returns to education are linear[2].

---

[2]The aggregate Bureau of the Census data yield and convex function Φ, while micro econometric studies generally show that income is a concave function of

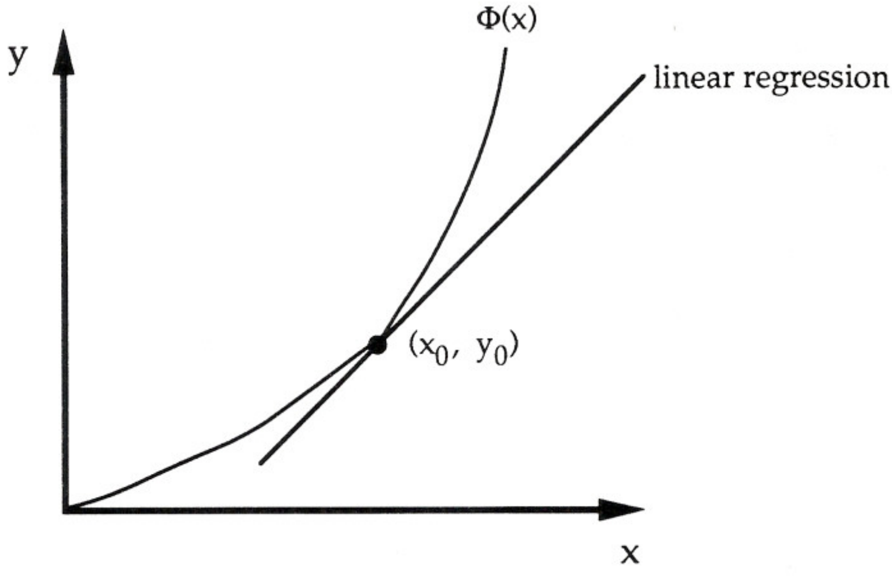


Figure 1

Our problem is thus the following: Given u, Φ, and the distribution of s in the generation of children at each level of parental income, what is the stationary (i.e., long-run) joint distribution of income and talent in this society?

Let the density of talents of children in each generation (and, by our assumption, at each income level of parents) be $\sigma(s)$, and let $S(s)$ be the C.D.F. of σ. Let $F_t(x)$ be the the C.D.F. of the distribution of education levels at generation t from some starting initial point which shall be irrelevant for us. Let $\Phi^e(x,x_0)$ be the beliefs of a child in a family whose parent has an income of

---

years of education (see, e.g., Willis [1986, Table 10.5]). But the micro studies enter years of experience as an independent variable as well. If the macro Census data are consistent with the micro data, it is because years of education are positively correlated with years of job tenure on average. Since our model simplifies reality by ignoring the relationship of income to years of experience, we use only the aggregative data.

x_0 about the returns to education. (I.e., $\Phi^e(x,x_0)=y$ is the tangent line illustrated in Figure 1.) The child, whom we may identify by his relevant traits (x_0,s), thus solves the following maximization problem :

choose x to maximize $u(\Phi^e(x,x_0), x; s)$,

which gives a solution that we may write $x^e(x_0,s)$. We now wish to calculate the relationship between $F_t(x)$ and $F_{t+1}(x)$. Condsider the equation $x^e(x_0,s)=x$. For a fixed value of x, and for a certain domain of x_0 , this typically will produce a unique solution in s, call it $s(x,x_0)$. The domain of x_0 for which the equation $x^e(x_0,s(x,x_0))=x$ is valid will typically be an interval, which we denote $[\alpha(x), \beta(x)]$. In the problems we shall study, the function $s(x,x_0)$ is monotone decreasing in x_0: that is, if we look at children who attain a given level of education x, their talent is inversely related to their parent's income. (This is due to the convexity of Φ: a child of a high income parent estimates high marginal returns to education, which makes her more 'optimistic.') For values of parental education x_0 below $\alpha(x)$, it is the case that all children, regardless of their talent, end up in the next generation acquiring education less than or equal to x. From this discussion, we can write the equation:

$$F_{t+1}(x) = F_t(\alpha(x)) + \int_{\alpha(x)}^{\beta(x)} S(s(x,x_0))dF_t(x_0) \quad .$$

The interpretation is as follows. Recall that $S(s(x,x_0))$ is the fraction of children with skill level less than or equal to $s(x,x_0)$, which is that skill level that leads a child of a parent with education x_0 to acquire education x. Therefore $S(s(x,x_0))$ is the fraction of children of parents with education x_0 who will acquire education less than or equal to x. Thus, the integral in the above equation is the fraction of children who will acquire skill level less than or equal to x, averaged over all families from which these children might come.

The first term in the above equation is the fraction of children who come from parents who have such low education levels that, regardless of the skill of the child, she will not seek education greater than x. The sum of these two terms is the fraction of children of parents at generation t who have acquired, by generation t+1, education less than or equal to x.

Hence, the stationary cumulative distribution function, F(x), satisfies the equation :

$$F(x) = F(\alpha(x)) + \int_{\alpha(x)}^{\beta(x)} S(s(x,x_0))dF(x_0) \quad . \qquad (2.1)$$

By differentiating equation (2.1), we obtain an equation involving the density function f(x) of F. Our task is to solve this equation for the unknown density function f. In Appendix 1, we present an example of a model in which (2.1) can be solved analytically. This example should also help clarify how the functions $\alpha(x)$ and $\beta(x)$ arise.

In general, (2.1) is a functional integral equation which cannot be solved analytically. Our procedure was to solve this equation by approximation methods, which are described in Appendix 2. Before proceding to the results, we introduce the functions u and σ with which we worked. We chose $\sigma(s) = \frac{1}{B(p,q)}(\frac{s}{200})^{p-1}(1-\frac{s}{200})^{q-1}$, where p=7.64 and q=7.65, and B(p,q) is the beta function. Thus σ is the density of a beta distribution with parameters (p,q), on the compact domain [0,200]. This is an approximation to a normal distribution with mean and median at 100. The choice of σ suggests that we think of talent s as being measured by 'IQ.' We take no position on the nature/nurture debate concerning the explanation of IQ: for us, all that is relevant is that a child with higher 'talent' will find education less psychologically costly to attain.

We chose the utility function for a child whose parent has education level x_0 to be:

$$u(x;\, x_0,s) = \frac{-k\,x^b}{(s/100)^a} + \int_0^x y_L e^{-\delta t}dt + \int_x^{52} \Phi^e(x,x_0)e^{-\delta t}dt \ . \qquad (2.2)$$

Its parameters are (a,b,k,δ,y_L). The first term is the disutility of education, which is larger the smaller is s. (k is a positive constant, to be estimated below.) We measure x in years of education acquired past the seventh grade: that is, x is years of education minus 7. It is assumed that all children complete the seventh grade, and, at that point, each calculates the number of years to continue in school. y_L is the annual income while in school, and δ is the time rate of discount. The first integral on the r.h.s. of (2.2) is the present value of income during the period of voluntary schooling; the second integral is the present value of income for the rest of one's life after schooling is finished, where the annual income is constant[3] during those years and assumed to be given by $\Phi^e(x,x_0)$. The top limit of that integral, 52, corresponds to one's age 52 years after finishing the seventh grade, or about 65 years.

We have estimated utility functions for white male adults and black male adults in the U.S. For each population, we estimated the 'true' function $\Phi(x)$ from 1990 data of the Bureau of the Census, which report mean income as a function of number of years of education, for the two population groups. These data exist for six values of x, from x=0 to x=12 (that is, from 7 to 19 years of education). Then, for each group, we calculated the function $\Phi^e(x,x_0)$. We chose, rather arbitrarily, a=3, b=2 for both groups. For whites, we chose δ=.06 and y_L= \$10,000; for blacks, we chose δ=.07 and y_L = \$8,000.

---

[3]Of course empirically, income increases with years of experience, a phenomenon we have not modeled.

For the white group, we fitted the points by pasting together two quadratic functions: Figure 2a plots the data points and our fitted functions. For the black group, one quadratic sufficed: Figure 2b shows the fit.

[Place figures 2a and 2b about here]

Thus, we arrived at an estimated function for whites:

$$\Phi_{white}(x) = \begin{cases} .264\ (x+7)^2 - 3.198\ (x+7)+25.41\ , & \text{for } x \leq 9 \\ -.644\ (x+7)^2+25.858\ (x+7)\ -207.04, & \text{for } x>9. \end{cases} \tag{2.3a}$$

(This function is differentiable at $x=9$, and has an inflection point there.) For blacks, the function is:

$$\Phi_{black}(x) = .177\,(x+7)^2 - 2.415(x+7) +22.9 \quad . \tag{2.3b}$$

After computing the functions $\Phi^e_{white}(x,x_0)$ and $\Phi^e_{black}(x,x_0)$ from (2.3ab), according to the recipe illustrated in Figure 1. We calibrated the constant k in the two utility functions as follows. For whites, we assumed that a child with $s=100$ from a family whose parent acquired just a high school diploma ($x=5$) would herself just acquire a high school diploma. This produced a value of $k=2.415$. For blacks, we assumed that a child with $s=100$ from a family whose parent acquired a tenth grade education ($x=3$) would acquire just a tenth grade education. This produced $k=.959$. We checked our utility function by looking at its predictions for other children (x_0,s) and got optimal education levels that looked reasonable.

3. <u>Estimating the relationship of talent to years of education attained</u>

Given the functions Φ^e and σ, we then solved the integral equation (2.1) by discretizing it, for an approximation to the density function of the stationary distribution of education, $f(x)$. (See the Appendix 2.) We then interpolated, fitting an eighth degree polynomial to the fifty points in our discrete approximation, giving us an analytic function for the density function. We shall report the story for whites first, and then for blacks. The discrete points and our fitted polynomials for whites are reproduced in Figure 3a.

[Place figure 3a about here]

The next task is to compute, for each value of education x achieved, the distribution of talent at that level of education. Recall that, were each child to know the function Φ, then in the stationary distribution (which would be achieved in one generation), there would be a perfect correlation between years of education achieved and talent. This turns out to be far from the case in our environment.

Let $\theta^x(s)$ be the density function of the distribution of talent (s) at $x+7$ years of education in the stationary distribution. Recalling the function $s(x,x_0)=s$, define its inverse as $x_0(x,s)=x_0$. (Thus, $x_0(x,s)$ is the educational level of a parent which will produce in a child of talent s an achieved education of x.) Then we have the formula:

$$\theta^x(s) = f(x_0(x,s))\ \sigma(s)/c, \qquad (3.1),$$

where c is a constant calculated so that the integral of $\theta^x(s)$ w.r.t. s is one. Equation (3.1) says that the frequency of s-talented people at education level x is proportional to the product of the frequency of children of s-talent (in the previous generation) with the frequency of parents of educational attainment $x_0(x,s)$ (in the previous generation). From (3.1), we calculated the densities $\theta^x(s)$ for a number of values of x. Figures 4a-4e reproduce those density functions for $x=1,2.5,5,7.5,10$ (recall that years of education attained are $x+7$).

Note the that the last four densities are bimodal. This is a consequence of the fact that our function $\Phi(x)$ has a convex and a concave part. Consider, for example, figure 4d. The second hump in the density function is associated with children who have high talent but came from parents who had low levels of education ($x_0<9$); the first hump is associated with children with low talent who came from parents with high levels of education ($x_0>9$). The cusp at s around 105 corresponds to children from parents at $x_0=9$.

[Place figures 4a-4e about here]

The means and medians of these five distributions are all fairly close: see Table 1.

x	median	mean
1	92	94
2.5	101	101.8
5	108	107
7.5	108	108
10	99.5	104

Table 1: Means and medians of talent by years of education for whites

Thus, the average level of talent changes very little over a very wide range of achieved educational levels. In other words, achieved level of education is a poor statistic for talent in a world where neighborhood effects are important in influencing children's beliefs about the returns to education. That the highest mean talent is not associated with the highest level of education is due to the bimodality of the density functions of talent. At x=7.5 (i.e., 14.5 years of education), there are a lot of talented children from poor families (the second hump of the density function has substantial mass). There are, however, very few talented children from poor families who achieve x=10 (17 years of education), and this pulls down the mean talent level at x=10.

For blacks, our discretized density function, and our polynomial fit to it are reproduced in Figure 3b.

[Place figure 3b about here]

The density of talent, associated with the five levels of education at the stationary distribution x=1, 2.5, 5, 7.5, and 10 are reproduced in Figures 5a-5e. Note that these functions, unlike the white densities, are unimodal. This is a consequence of the fact that there is no inflection point in the black stationary distribution of education levels (see Figure 3b).

[Place figures 5a-5e about here.]

Finally, we computed the means and medians of these density functions: see Table 2.

x	median	mean
1	88.5	89.4
2.5	88.5	90.7
5	100.2	101.9
7.5	113.3	114.6
10	127.8	129.2

Table 2: Means and medians of talent by years of education for blacks

Evidently, the mean and median talent vary much more with education levels at the stationary distribution for blacks than they do for whites.

4. Conclusion

Our purpose has been to show that, when young people form rational beliefs about the returns to education -- rational in the sense that they are based on linear extrapolations from the evidence they have -- then, if the sample upon which their extrapolations are based is biased, the income and education levels that they achieve will not accurately reflect their true potential. Here, we are thinking of the potential income and educational achievement of a person as that she would achieve were she to solve her optimization problem with correct information. We have made a minimal effort to calibrate the parameters of the problem we have studied by using real data, but we do not pretend this is an empirical paper. Were it to be seriously empirical, we would have to account for the fact, for example, that income is not a one-to-one function of education. We offer our exercise as an illustration of a phenomenon that we believe occurs in the real world.

Our general result is that, due to incorrect beliefs, the stationary distribution of 'talent' at any level of income/education has a non-trivial variance. An observor who cannot see talent cannot, therefore, infer a person's talent with any degree of precision by observing her educational/income achievement. Viewed from another angle, the result means that there is a great deal of inefficiency due to incorrect beliefs, in the sense that, were young people to optimize given correct beliefs, then educational levels would be a monotone increasing function of talent.

In our model, it turns out that education is a somewhat better predictor of talent for blacks than for whites; that is, the mean (and median) level of talent as a function of achieved level of education is a more sharply increasing function for blacks than for whites. We do not have enough confidence in the empirical accuracy of our model to assert that this is true in real life. For instance, it may be the case that blacks are, in reality, more sharply segregated in neighborhoods by income level than whites, and so white children may have a less biased sample from which to form their beliefs than black children. This would tend, of course, to reverse this result.

We tried to do some dynamics with the model. Namely, we calculated a discretized transition matrix, whose ij^{th} element is the probability of moving from i years to j years of education in one generation. Using this matrix, we tried to compute the expected number of generations to move from, say, finishing grammar school to finishing high school. But the series did not converge sufficiently fast to give us reasonable estimates using reasonable amounts of computer time.

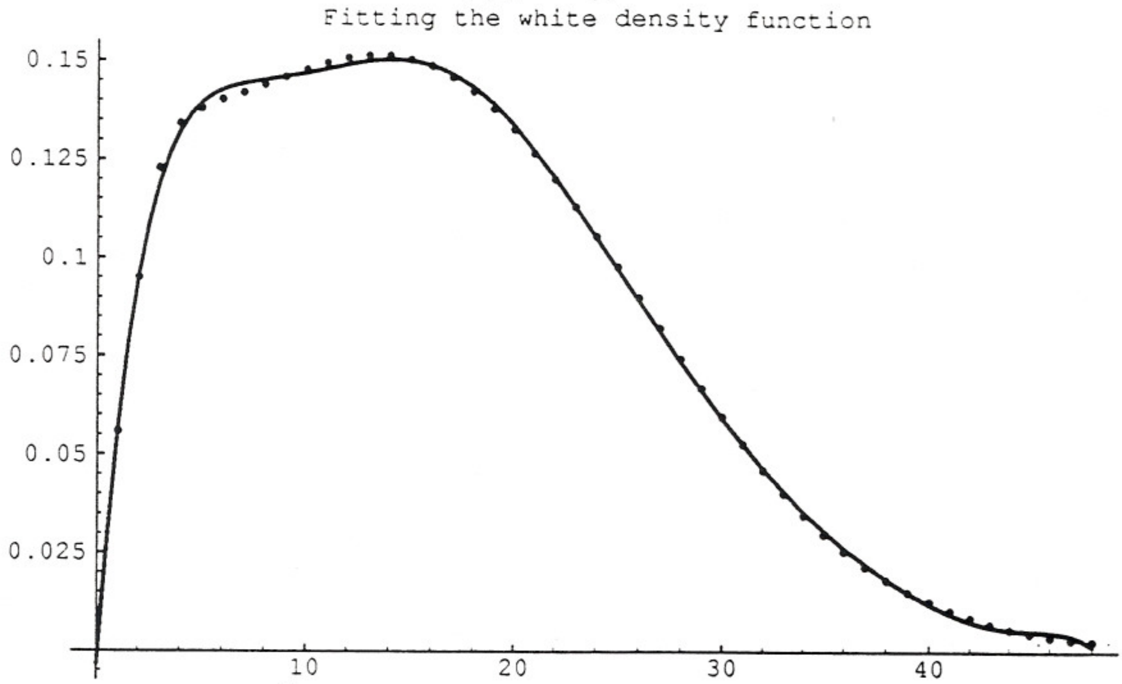


Figure 3a

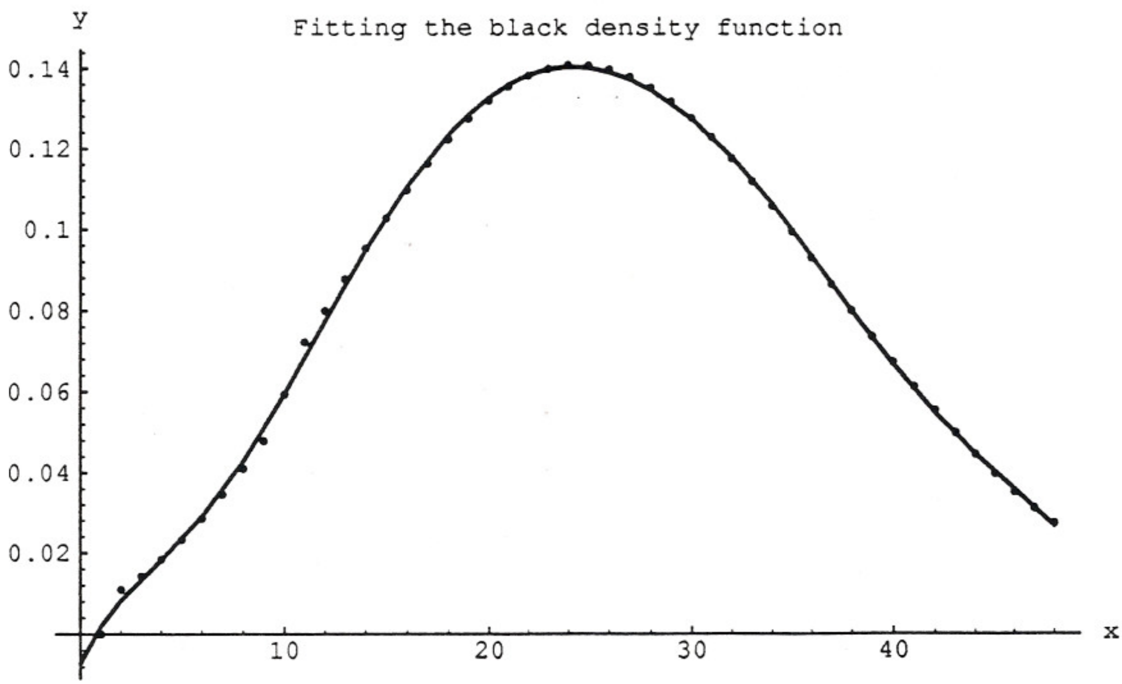


Figure 3b

Figure 4: Densities of talent by years of

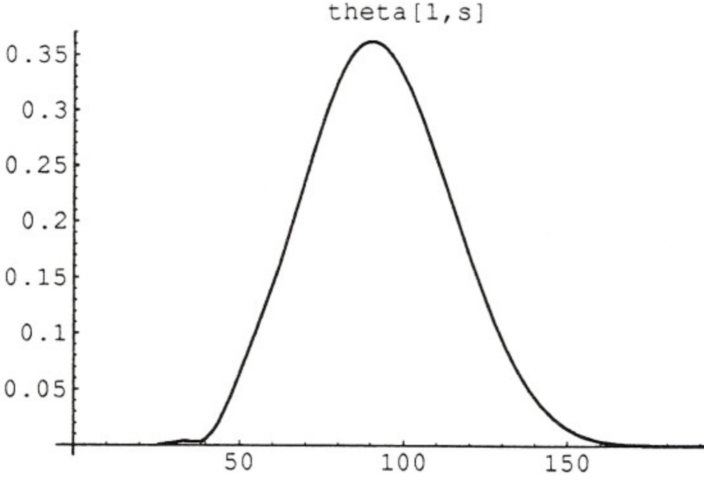


Figure 4a

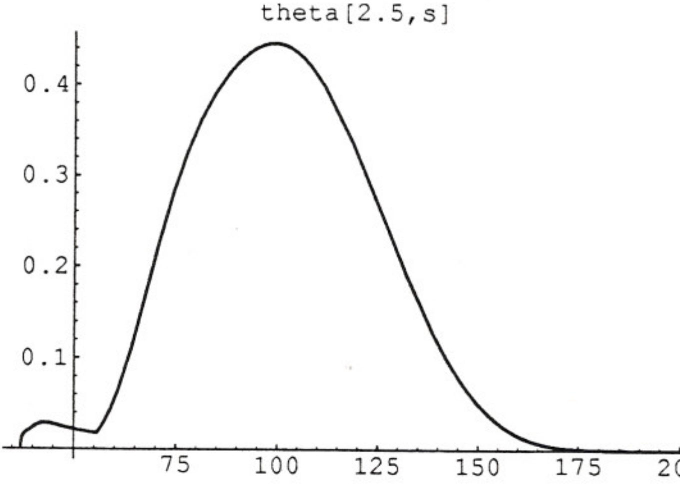


Figure 4b

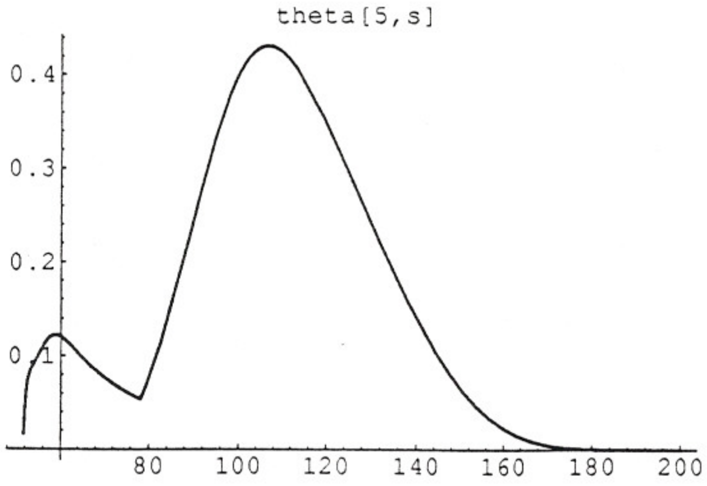
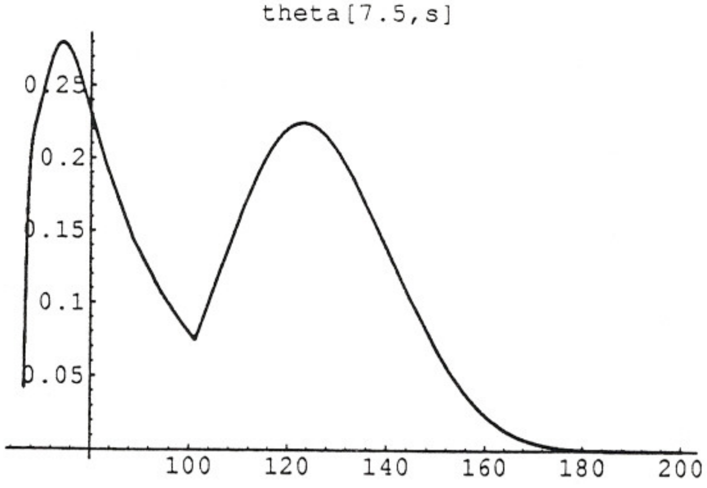


Figure 4c



Figures 4d

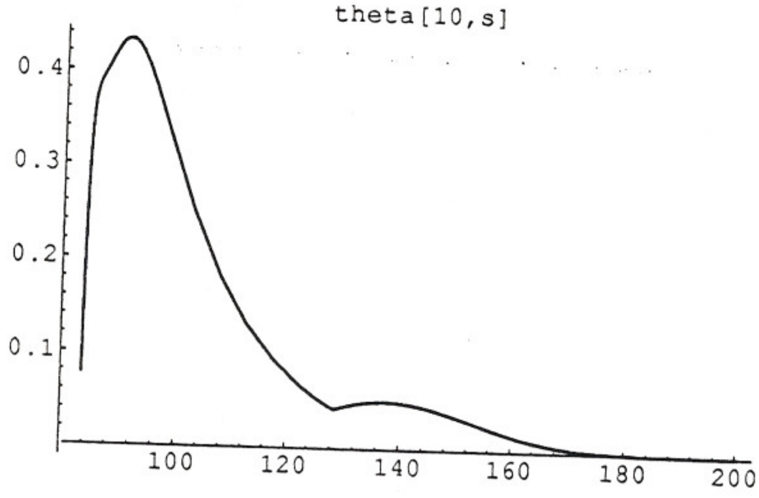


Figure 4e

Figure 5: Densities of talent by years of educatic

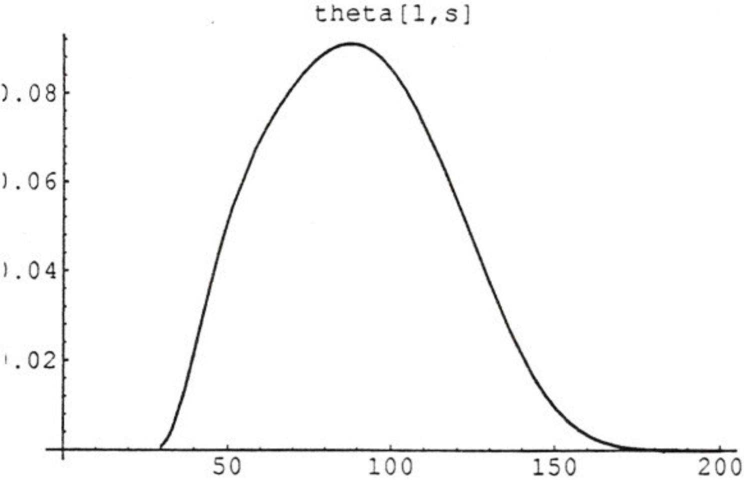


Figure 5a

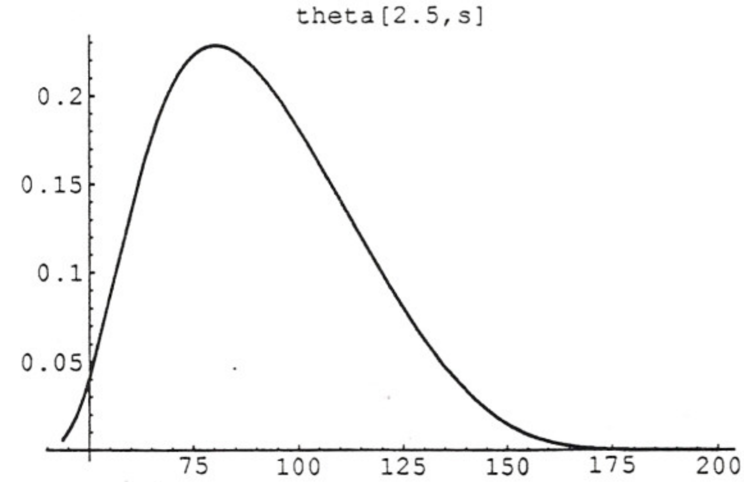


Figure 5b

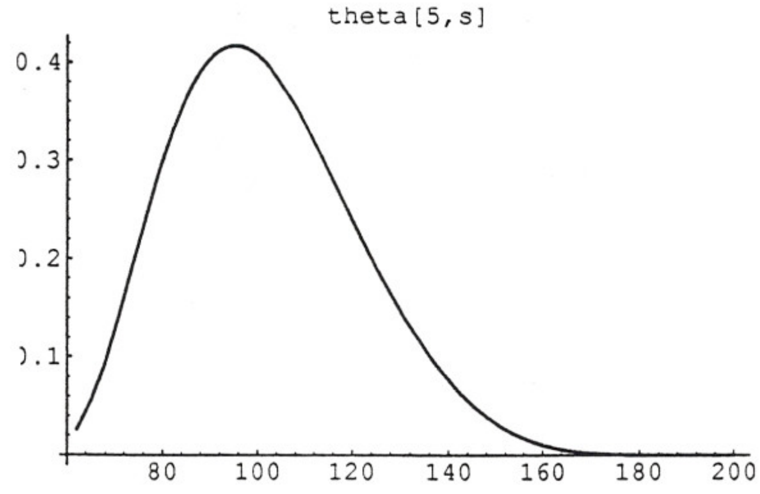


Figure 5c

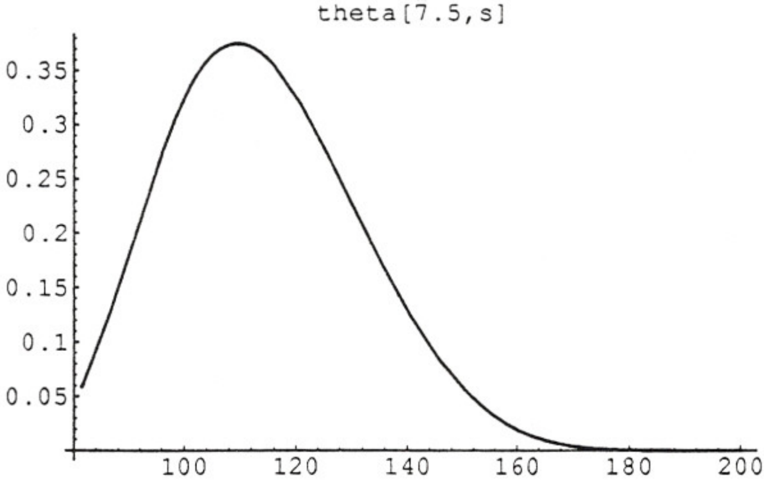


Figure 5d

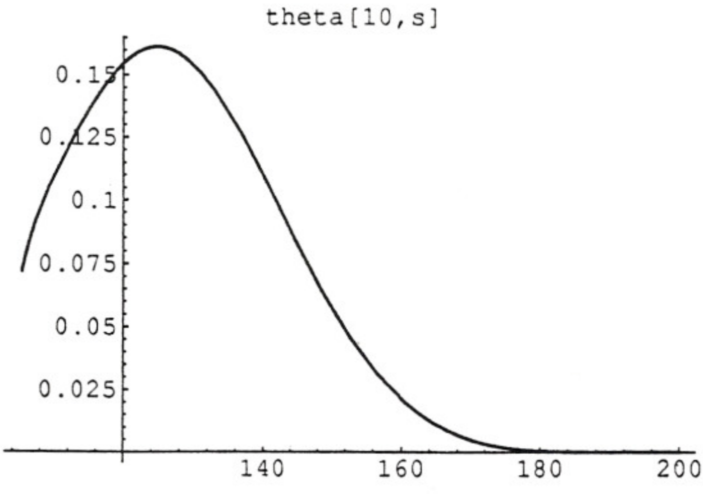


Figure 5e

APPENDIX 1

In this appendix[4], we provide an example of a model that yields a version of the functional equation (2.1) which has been solved in the literature. Let $\Phi(x) = \frac{1}{2}\alpha x^2$ and $u(y,x) = y - \frac{x^3}{3s}$. Then the child's estimate of income as a function of the education level of her parent (i.e., the equation of the tangent line in Figure 1) is:

$$y^e = (\alpha x_0)\, x + k^*,$$

where k^* is a constant which does not matter for us because of the quasi-linearity of utility. Solving the child's optimization problem:

$$\text{choose } x \text{ to maximize } (\alpha x_0)\, x - \frac{x^3}{3s}$$

yields

$$x = (\alpha s x_0)^{1/2}. \qquad \text{(A1)}$$

Thus, the talent level s that a child of a parent with x_0 years of education must have in order to choose to consume x years of education is

$$s(x,x_0) = \frac{x^2}{\alpha x_0}. \qquad \text{(A2)}$$

Let us now assume that talent s is uniformly distributed on [0,1]. From (A1), we see that the largest possible level of education x^{max} must satisfy the inequality:

$$x^{max} \leq (\alpha\, x^{max})^{1/2},$$

which implies that $x^{max} \leq \alpha$. Thus, the density function of the stationary distribution of education must be zero for $x > \alpha$. Note also from (A1) that $x_0 = \frac{x^2}{\alpha s} \geq \frac{x^2}{\alpha}$, which tells us that if $x_0 < \frac{x^2}{\alpha}$, then the child, no matter what her

---

[4]We are grateful to Glenn Loury for discovering that, with the utility function used in this appendix, our equation (2.1) yields the functional equation that he has solved.

talent level, must achieve education no larger than x. Hence (see (2.1)) the function $\alpha(x) = \frac{x^2}{\alpha}$.

We have chosen $\sigma(s) = 1$ (the uniform density), and so $S(s)=s$. Let $F(x)$ be the C.D.F. of the stationary distribution of education. Then equation (2.1) becomes:

$$F(x) = F\left(\frac{x^2}{\alpha}\right) + \int_{\frac{x^2}{\alpha}}^{\alpha} \frac{x^2}{\alpha x_0} f(x_0)\,dx_0. \tag{A3}$$

Differentiation of (A3) w.r.t. x yields:

$$f(x) = \int_{\frac{x^2}{\alpha}}^{\alpha} \frac{2x}{\alpha x_0} f(x_0)\,dx_0. \tag{A4}$$

Our problem is to solve (A4) for the function $f(x)$. This equation is identical to equation (11) of Loury (1981); in that paper, Loury provides an infinite-series representation of the solution.

We have been able to find no other example of a model which yields a solution that can be calculated analytically. In particular, we have no precisely calculable example when the density of talent is not uniform. Therefore, in the text we have relied on approximating the solution $f(x)$ to the analogue of (A4) using the method described in Appendix 2.

2.

APPENDIX NUMERICAL PROCEDURE

This appendix gives the details of how the stationary density functions of the education levels for whites and blacks were calculated. In both cases, the stationary density function —associated with the cumulative distribution function satisfying (2.1)— is the solution of a system of (integral) linear equations of the type:

$$\int k(x,z)f(z)\,dz = f(x) \qquad \forall x \in [0,12],$$

$$\int f(z)dz = 1, \quad f(z) \geq 0 \text{ a.e.},$$

where for whites:

$$k(x,z) = \frac{b((x-d_w(z))/c_w(z))}{c_w(z)},$$

and for blacks

$$k(x,z) = \frac{b((x-d_b(z))/c_b(z))}{c_b(z)},$$

$b(\cdot) = b(\cdot; 7.64, 7.65)$ is the beta density function with parameters $p = 7.64, q = 7.65$ and support $[0, 200]$, and the functions c_w, d_w, c_b, d_b are polynomials in z. This beta density function, for all practical purposes, looks like a normal distribution with mean 100, but it has compact support.

In terms of equation (2.1), k is the derivative of $S(s(x,z))$ with respect to x with $x = a_j(z)s + b_j(z)$ for $j = w, b$. The coefficients a_j, b_j are such that

$$a_j(z)s + b_j(z) \approx \operatorname*{argmin}_z u_j(x; z, s) \qquad \text{for } j = w, b$$

with u_w the utility function used for white and u_b the utility function for black. For fixed z, a numerical optimization procedure was used to find the maximizer $x^\star(z,s)$ of $u(\cdot; z, s)$ as a function of s. Plotting $s \mapsto x^\star(z,s)$ shows that this function is essentially linear. A least square fit was used to obtain the coefficients $a_j(z)$ and $b_j(z)$. Finally, a fifth degree polynomial was used to fit the functions $z \mapsto a_j(z)$ and $z \mapsto b_j(z)$. For $j = w$, one obtains

$$c_w(z) = 0.00894 + 0.0183z + 0.00273z^2 - 0.001582z^3 + 0.00022z^4 - 0.00001z^5,$$

and

$$d_w(z) = -0.47527 - 1.33323z + 0.30322z^2 - 0.03356z^3 + 0.00176z^4 - 0.00003z^5.$$

Except for very special cases, one cannot find closed form solutions to the linear system of integral equations. What we are going to do is discretize the interval over which f is defined, and find an approximate solution. The last step is to fit these points (by interpolation or curve fitting) to obtain an expression for f that can be manipulated.

Let $[z_1, z_2, ..., z_n]$ be a discretization of the interval $[0, 12]$ on which f is defined. We choose n points, usually equally spaced, but not necessarily. The unknow function f is then replaced by an unknow n-vector: $[f(z_1), f(z_2), ..., f(z_n)]$. To simplify notation, let us also write v_i for $f(z_i)$, $i = 1, ..., n$, and $v = (v_1, ..., v_n)^T$ (i.e., a column n-vector).

The integral equations,

$$\int k(x, z) f(z)\, dz = f(x) \quad \forall x$$

can also be written as

$$\int f(z)\, dK(x, z) = f(x) \quad \forall x$$

where

$$K(x, z) = \int^z k(x, \tau)\, d\tau.$$

Thus for each x, we are dealing with a Rieman-Stieltjes integral with the measure induced by the distribution $K(x, \cdot)$. Using for x the same mesh as for the z variables, and approximating these integrals by a finite sum, one obtains the following n equations:

$$\sum_{j=1}^{n} f(z_j) \Delta K(z_i, z_j) = f(z_i), \quad i = 1, \dots, n,$$

where

$$\begin{aligned} a_{i,j} &:= \Delta K(z_i, z_j) = K(z_i, \tfrac{1}{2}(z_j + z_{j+1})) - K(z_i, \tfrac{1}{2}(z_{j-1} + z_j)) \text{ for } j = 2, \dots, n-1, \\ a_{i,1} &:= \Delta K(z_i, z_1) = K(z_i, \tfrac{1}{2}(z_1 + z_2)) - K(z_i, 0), \\ a_{i,n} &:= \Delta K(z_i, z_n) = K(z_i, 12) - K(z_i, \tfrac{1}{2}(z_{n-1} + z_n)) \end{aligned}$$

so that

$$\begin{aligned} a_{i,j} &= \int_{\frac{1}{2}(z_{j-1}+z_j)}^{\frac{1}{2}(z_j+z_{j+1})} k(z_i, \tau)\, d\tau, j = 2, \dots, n-1, \\ a_{i,1} &= \int_0^{\frac{1}{2}(z_1+z_2)} k(z_i, \tau)\, d\tau, \\ a_{i,n} &= \int_{\frac{1}{2}(z_{n-1}+z_n)}^{12} k(z_i, \tau)\, d\tau, \end{aligned}$$

Similarly, one discretizes the equation $\int f(z)dz = 1$ as follows

$$\sum_{j=1}^{n} f(z_j)\Delta(z_j) = 1,$$

with

$$b_j := \Delta(z_j) = \tfrac{1}{2}(z_{j+1} - z_{j-1}),\ j = 2, \ldots, n-1,$$

$$b_1 := \Delta(z_1) = \tfrac{1}{2}(z_1 + z_2), b_n := \Delta(z_j) = 12 - \tfrac{1}{2}(z_n - z_{n-1}).$$

Writing I for the identity matrix, the discretized version of the problem we need to solve is:

$$\begin{aligned} (A - I)v &= 0 \\ b\ v &= 1 \\ v &\geq 0 \end{aligned}$$

This is a system of $n+1$ equations and n inequalities in n variables. We are going to solve this as follows:

$$\begin{aligned} &\text{minimize } \kappa \\ &\text{such that } \kappa \geq (A - I)_i v,\ i = 1, ..., n \\ &\qquad\qquad\ \kappa \geq (I - A)_i v,\ i = 1, ..., n \\ &\qquad\qquad\ 1 = bv, \quad v \geq 0 \end{aligned}$$

where $(A - I)_i$ is the i-th row of the matrix $(A - I)$, i.e. $(A - I)_i v$ is the (inner) product of the i-th row of $(A - I)$ and v.

This problem is a linear programming problem, for which standard techniques are available, i.e., if we know the coefficients of the matrix $(A - I)$, we can feed this in a standard package. The package we used was MINOS and we chose $n = 50$.

This yields a discrete approximation to the density f. An analytic expression was obtained by fitting an eighth degree polynomial to these points, cf Figures 3a and 3b.

Benabou, R. 1991. "Workings of a city: Location, education, and production," MIT discussion paper

Durlauf, S. 1992. "A theory of persistent income inequality" Stanford University, mimeo

--. 1993. "Neighborhood feedbacks, endogenous stratification, and income inequality," Stanford University, mimeo.

Jencks, C. Rethinking social policy: Race, poverty, and the underclass

Loury, G. 1981. "Intergenerational transfers and the distribution of income," Econometrica 49, 843-867

Streufert, P. 1991."The effect of underclass social isolation on schooling choices," University of Wisconsin at Madison, mimeo

Willis, R. J. 1986. "Wage determinants: A survey and reinterpretation of human capital earnings functions," in O. Ashenfelter and R. Layard, Handbook of Labor Economics, Amsterdam: North-Holland Press

Wilson, W. The truly disadvantaged:The inner city, the underclass, and public policy, Chicago: University of Chicago Press, 1987