ECONOMIC RESEARCH REPORTS

Sorting and Long-Run Inequality

By

Raquel Fernandez
and
Richard Rogerson

RR# 99-22 **December 1999**

C.V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, NY 10003-6687

Sorting and Long-Run Inequality

Raquel Fernández
New York University
CEPR and NBER

Richard Rogerson
University of Pennsylvania

September 1999
This Version: December 1999

Abstract

Many social commentators have raised concerns over the possibility that increased sorting in society can lead to greater inequality. To investigate this, we construct a dynamic model of intergenerational education acquisition, fertility and marital sorting and parameterize the steady state to match several basic empirical findings. Contrary to Kremer's (1997) findings of a basically insignificant impact of sorting on inequality, we conclude that increased marital sorting will significantly increase income inequality. Three factors are central to our findings: a negative correlation between fertility and education, a decreasing marginal effect of parental education on children's years of education, and wages that are sensitive to the relative supply of skilled workers.

We thank Michael Kremer, Nezih Guner, John Knowles, Victor Rios-Rull, and participants at the NBER Summer Institute Inequality Workshop, LACEA, Latin American Econometric Society Meetings, SITE and SED Conference, and seminar participants at NYU, Rochester, University of Pennsylvania, Southampton University, Warwick, LSE, and the Board of Governors for very helpful comments. We thank Dan Levy for his research assistance. The authors gratefully acknowedge support from the NSF and the first author also thanks the CV Starr Center for financial support.

1. Introduction

Many social commentators claim that American society is becoming more stratified, in the sense that individuals are tending to interact more with others who are similar to themselves, and less with others who are different. These interactions include who one works with, who one marries, who one goes to school with, and who one has as neighbors.[1] Thus, increased stratification or sorting may take place along the dimensions of skills, income, education, aptitude, race and ethnicity.

Why might increased sorting matter? It has been hypothesized that increased sorting may reduce redistribution, increase negative activities such as crime, or reduce positive peer effects in school. Some observers (e.g. Wilson (1987) and Reich (1991)) have argued that increased sorting may have significant consequences for the degree of inequality in society, with more sorting leading to greater inequality. A recent and provocative paper by Kremer (1997), however, argues that the quantitative effects of even very large increases in sorting–whether marital or residential–are likely to be negligible, at least as concerns the distribution of income and education.

The objective of this paper is to investigate in greater depth the effects of increased marital sorting on inequality. In order to do so, we examine a model of intergenerational education acquisition and marital sorting and parameterize it to match several basic empirical findings. We find that increased sorting may significantly increase income inequality.

Our model is very simple. Individuals are either skilled (college educated) or unskilled (high-school educated). They meet, marry and have children. Unable to borrow against future human capital, families decide how many of their children to send to college based on their family income, their children's abilities, and the expected wage differential for skilled relative to unskilled labor. The distribution of education determines wages, and together they determine the distribution of

[1] Whether stratification or what we will generally call "sorting" has actually increased is a separate question and one that we will not investigate here. Kremer and Maskin (1996) present some evidence in support of the argument that sorting by skill level in the workplace has increased. Evidence on marital sorting appears mixed. Based on years of schooling, the evidence leads to the conclusion that sorting has not increased; however, the decreased probability that certain educational barriers will be crossed (e.g. high school graduate married to college graduate), suggests greater sorting (Mare 1991). Sorting at the neighborhood level appears not to have increased (Kremer 1997), whereas annecdotal evidence of increased tracking in schools and the proliferation of magnet schools suggests that sorting in schools may be increasing.

income. We solve for the steady states of the dynamic model; perhaps not surprisingly, the existence of borrowing constraints generates multiple steady states.

The degree of sorting in this model is reflected in the fraction of the population that gets perfectly (as opposed to randomly) matched with a marriage partner. An increase in the degree of sorting can, in theory, either increase or decrease the skilled fraction of the population, depending on a number of factors that we discuss in our analysis. In our calibrated model we find that if marital sorting increases, then a smaller fraction of children will become skilled. This drives down wages for unskilled workers and increases those of skilled workers and also increases the degree of wage inequality. If as a result of lower wages, borrowing constraints become tighter for low-income families, the effect on wage inequality is further magnified.

Contrasting our findings with that of Kremer's, we find three factors, all absent in Kremer's analysis (and, we argue, present in the data), to be central to our results. In particular, a negative correlation between fertility and education, a decreasing marginal effect of parental education on children's years of education, and a process of wage determination that is sensitive to the relative supply of skilled to unskilled workers all contribute to our qualitative and quantitative conclusions.

In addition to the paper by Kremer, our work is related to several others in the literature. Benabou (1996a), Caucutt (1997), Cooper (1997), Durlauf (1995), Epple and Romano (1996) and Fernandez and Rogerson (1996, 1997) examine the effects of neighborhood and school sorting generated either endogenously by education policies or exogenously via increased neighborhood stratification. Banerjee and Newman (1993), Benabou (1996b), Fernandez and Rogerson (1998), Galor and Zeira (1993), Loury (1981), and Ljungvist (1993) examine the effects of the existence of borrowing constraints on the dynamic evolution of the economy and income inequality. The effects of endogenous fertility on income distribution (and vice versa) have recently been the subject of analysis in Dahan and Tsiddon (1998), Greenwood, Guner, and Knowles (1999), and Kremer and Chen (1999) among others.

The outline of the paper follows. In the next section we describe the model and its steady states. In section 3 we analyze the effects of changes in sorting. In section 4 we use data to parameterize the model, and in section 5 we use our parameterized model to assess the effects of a large increase in sorting. Section 6 reviews Kremer's analysis and contrasts it with our own. Section 7 examines the robustness of our results to alternative parameterizations and Section 8 concludes.

2. The Model

To examine the effects of marital sorting on the process of intergenerational education transmission and income inequality requires a dynamic model that incorporates marriage, fertility, education and the determination of income. The interaction of these factors easily yields a non-tractable model (see Greenwood, Guner and Knowles (1999) for a computational approach to this problem) so, wherever possible, we choose to model these decisions in as simple a way as possible, keeping many elements exogenous (in particular fertility and marriage decisions) in order to highlight the interactions that are central to our analysis.[2]

The story our model tells is a simple one. In each period the adult population is characterized by a distribution of education or skill levels. We assume that individuals are either skilled or unskilled and that a competitive labor market determines the relative wages of these workers. These individuals meet and marry, with their marriage partners determined via an exogenous matching process that exhibits positive assortative matching. Couples have children and, based on the number of children, their aptitudes, family income, and expected wages, they decide the education levels of their children. This generates the next generation's distribution of education (skill levels). A more formal description follows.

Marriages

Consider a population at time t whose number is given by n_t and some division of that population into skilled workers, n_{st}, and unskilled workers, n_{ut}, where:

$$n_t = n_{st} + n_{ut} \tag{2.1}$$

For our purposes, skill levels will be synonymous with an educational attainment. All college-educated workers are skilled (s); all others are unskilled (u).

Each individual is matched with another, resulting in a "marriage" according to the following mechanical process. In order to capture the degree of sorting in the economy, we allow some fraction of marriages, say θ, to be perfectly matched, i.e., a skilled worker marries another skilled worker or an unskilled worker marries another unskilled worker. The remaining fraction of the population is matched in a random fashion. Thus marriages will belong to one of three categories: skilled marries skilled (denoted by h for high type), skilled marries unskilled (denoted by m for mixed or middle type), and unskilled marries unskilled (denoted by l for

[2] As we argue in the conclusion, we believe that most plausible ways of endogenizing fertility and marital decisions will reinforce our conclusions.

low type). These categories will also correspond to the relative position of couples in the income distribution.

Given the degree of assortative matching θ and the distribution of the population at time t into skilled and unskilled, the total number of high type marriages at time t, h_t, is given by:

$$h_t = \lambda_{ht} \frac{n_t}{2}$$

where

$$\lambda_{ht} \equiv \theta\beta_t + (1-\theta)\beta_t^2 \tag{2.2}$$

and β is the ratio of skilled workers in the population, i.e.,

$$\beta_t = \frac{n_{st}}{n_t} \tag{2.3}$$

The number of marriages that are of the middle type at time t is given by

$$m_t = \lambda_{mt} \frac{n_t}{2}$$

where

$$\lambda_{mt} \equiv 2(1-\theta)\beta_t(1-\beta_t) \tag{2.4}$$

whereas the number that are low type is given by:

$$l_t = \lambda_{lt} \frac{n_t}{2}$$

where

$$\lambda_{lt} \equiv \theta(1-\beta_t) + (1-\theta)(1-\beta_t)^2 \tag{2.5}$$

Of course, $\lambda_{ht} + \lambda_{mt} + \lambda_{lt} = 1$ and $h_t + m_t + l_t = \frac{n_t}{2}$.

Children

Fertility undoubtedly depends on parental education, income, culture and technology among other things. We simplify matters by assuming that fertility is determined entirely by the educational backgrounds of the parents. Thus, fertility is only a function of marriage type and can be denoted by f_j, $j = h, m$, or l (so that all families of marriage type j have the same number of children).[3]

[3] It would not be difficult to relax this assumption, but this would offer little additional insight.

A child can be of two "aptitude" types which we denote by either high or low, the significance of which will be made clear shortly. The probability that a given child is of high aptitude, γ_j, $j = h, m, l$ is allowed to differ across family types but not across families within the same category.[4] Realizations are independent across children. The probability, therefore, that a family with a total number of children f_j has $n \leq f_j$ children of high aptitude is $\gamma_j^n(1-\gamma_j)^{f_j-n}\binom{f_j}{n}$ where $\binom{f_j}{n}$ is the binomial coefficient (equal to the number of combinations of f_j things taken n at a time).

Education

A family's decision to send a child on to college is determined by the child's aptitude, family income, and expected wages. If a child with high aptitude obtains a college education, we assume she receives one unit of skilled human capital, whereas a low aptitude child who goes on to college is assumed to obtain zero units of skilled human capital.[5] The quantity of unskilled human capital that a child obtains is assumed to be independent of her aptitude level, i.e., all individuals who obtain only a high-school education have the same level of human capital. The aptitude (and education) of a child is assumed to be perfectly observable to all.

We assume that the cost of sending a child to high school is zero whereas a positive (constant) cost, ν, must be incurred before obtaining a higher education. To render the decision of whether to send a child to college as simple as possible, we assume that, subject to obtaining a minimum per capita consumption level of $\overline{c}$, a family would always desire to send a high-aptitude child to college if the net return from doing so, $w_s - \nu$, exceeded the return from high school, w_u. More formally, if a family of type j with n high aptitude children sends $r \leq n$ of them on to college, and has per capita consumption equal to c, we assume they receive utility

[4] In this sense perhaps the term aptitude is a misnomer since, strictly speaking it is not genetically determined (otherwise we would have to keep track of whether a couple included 0, 1, or 2 high-aptitude individuals). It is best thought of as a high or a low ability to obtain marketable skills from college. This ability is assumed to depend on parental education and hence differs across family types.

[5] We could easily assume that a low aptitude child ends up with $\varphi < 1$ units of skilled human capital. This would multiply the number of potential steady states we have to examine but not add any new factor of interest to our analysis.

$$U_j = \begin{cases} (c-\overline{c}) & \text{for } c < \overline{c} \\ (c-\overline{c}) + \frac{r}{(2+f_j)} w_s + \frac{(f_j - r)}{(2+f_j)} w_u, & \text{otherwise} \end{cases} \tag{2.6}$$

where w_s and w_u are next period's wages for skilled and unskilled workers respectively.[6]

Wages

We assume a constant returns to scale aggregate production function given by:

$$\begin{aligned} F(n_s, n_u) &= n_u F(n_s/n_u, 1) \equiv n_u F(\frac{\beta}{1-\beta}, 1) \equiv n_u f(\beta) \\ f' &> 0, \quad f'' < 0 \end{aligned} \tag{2.7}$$

Assuming a competitive labor market, it follows that wages are determined only by the value of β:

$$w_s(\beta) = (1-\beta)^2 f'(\beta) \quad \text{and} \quad w_u(\beta) = f(\beta) - \beta(1-\beta) f'(\beta) \tag{2.8}$$

where the assumptions in (2.7) imply that skilled wages are decreasing in the ratio of skilled to unskilled workers and the opposite for unskilled wages. Note that no family would want to send their child to college if the fraction of skilled workers exceeds $\overline{\beta}$, where $\overline{\beta}$ is defined by:

$$w_s(\overline{\beta}) = w_u(\overline{\beta}) + \nu \tag{2.9}$$

We assume henceforth that $\overline{\beta}$ is strictly positive.Note, furthermore, that $\overline{\beta}$ would be the fraction of the population that would attend college if there were no borrowing constraints and, on aggregate, the fraction of high aptitude children exceeded $\overline{\beta}$.

Budget Constraints

The utility maximization problem of a family of type j with n high-aptitude children is given by the maximization of U_j as specified in (2.6) subject to a budget constraint. Note that in the absence of any impediments to borrowing against future income, all high-aptitude children would attend college as long as

[6] We use this linear utility function for simplicity only; we could specify a concave utility function. It would be equally simple to incorporate discounting of children's future income or differential weights on family members' consumptions.

$\beta < \overline{\beta}$ in the subsequent period. With borrowing constraints, however, household income is an important determinant of the number of children that a family can afford to send to college.

In what follows, we assume that families are unable to access credit or insurance markets.[7] For interpretational purposes, however, we think that it is important to note that these borrowing constraints need not be thought of as constraining directly the capacity of a family to send a child to college (which is debatable as some colleges are close to free).[8] Instead, in a richer model the inability to borrow against a child's future income could serve to constrain a family's residential choice and consequently the quality of the high school their children can attend. This would then affect both the amount of human capital obtained from high school attendance and the probability that the child attends college.

Thus utility maximization is subject to a household-income budget constraint:

$$\begin{aligned} (2+f_j)c + r\nu &\leq I_j(\beta) \\ 0 &\leq r \leq n \end{aligned} \tag{2.10}$$

where

$$I_j(\beta) = \begin{cases} 2w_s(\beta) & \text{for } j = h \\ w_s(\beta) + w_u(\beta) & \text{for } j = m \\ 2w_u(\beta) & \text{for } j = l \end{cases} \tag{2.11}$$

Note that a higher fraction of skilled workers implies lower wages for skilled workers and higher ones for unskilled workers. Hence, an increase in β implies tighter budget constraints for high-type families and looser ones for low-type families. Whether the budget constraint for middle-type families is loosened or tightened depends on whether the increase in the wage of unskilled workers is greater

[7] It is, of course, not necessary to shut down capital markets altogether in order to obtain the result we desire—that the maximum number of children a family can afford to send to college is a function of family income. It is simple to write down micro-foundations (e.g., moral hazard or imperfect enforcement technology) for this or less extreme assumptions (see, for example, Ljungvist (1993), Banerjee and Newman (1993), or Galor and Zeira (1993)). Note also that families would want to pool risk since the number of high aptitude children each has is stochastic.

[8] Although, of course, there are subsistence costs to be met, etc. Indeed Behrman, Pollak, and Taubman (1989) argue that unequal access to financing for college can help explain differences in educational attainment.

than the accompanying decrease in skilled wages, i.e., on whether $n_s - n_u$ is positive.

2.1. Steady States

It is straightforward to show that if β_t is the fraction of the population that is skilled in period t, then next period's value of β is uniquely determined. The dynamic evolution of this economy will of course depend on the fertility of each family type, the fraction of children of each type that are of high aptitude, wages, minimum required consumption, and the cost of college.

Though the economy will follow a unique path starting from any initial condition, in general this economy may have multiple steady states. To see why this is the case, note that the fraction of skilled workers in the economy determines the income level for each marriage type, which in turn determines who can afford to attend college. A higher fraction of skilled workers implies a higher wage for unskilled workers, thereby tightening constraints for high-type families, loosening constraints for low-type families and having an ambiguous effect for middle-type families. Thus, a low initial proportion of skilled workers can be reinforcing if as a consequence of low unskilled wages a large fraction of families find themselves constrained. Similarly, a high initial proportion of skilled workers can be reinforcing if as a consequence of high unskilled wages a small fraction of families find themselves constrained. This positive feedback effect can give rise to multiple steady states.

Finding the *potential* steady-states of the system is simple. Suppose that families of type j can afford to send a maximum of z_j children to college (and find it desirable to do so). To solve for the fraction of their children that type j families will send to college in aggregate, $\Gamma_j(z_j)$, requires finding the distribution of high-aptitude children over type j families and evaluating which of these are constrained. In particular, we need to calculate

$$\Gamma_j(z_j) \equiv \sum_{s=1}^{z_j} \binom{f_j}{s} \gamma_j^s (1-\gamma_j)^{f_j - s} \frac{s}{f_j} + \sum_{s=z_j+1}^{f_j} \binom{f_j}{s} \gamma_j^s (1-\gamma_j)^{f_j - s} \frac{z_j}{f_j} \tag{2.12}$$

where the first summation finds the fraction of children that go to college from families of type j that are not constrained (as the number of high-aptitude kids they have is less than z_j) and the second summation does the same for constrained families of type j.

We can now solve for the potential steady states of the system by examining the fixed point generated by the dynamic system described in equation (2.13) below for each possible vector of $\mathbf{z} = (z_h, z_m, z_l)$ such that $z_j \leq f_j \ \forall j$. Hence the potential steady states of the economy, $\widehat{\beta}$, are the solutions to:

$$\beta_{t+1}(\theta) = \frac{n_{st+1}}{n_{t+1}} = \frac{\sum_j \Gamma_j(z_j) f_j \lambda_{jt}(\beta_t; \theta)}{\sum_j f_j \lambda_{jt}(\beta_t; \theta)} \tag{2.13}$$

such that $\beta_{t+1} = \beta_t = \widehat{\beta}$.

Of course, not all these *potential* steady-states will be actual steady states. First, parents must wish to send their children to college, i.e., $\widehat{\beta} \leq \overline{\beta}$. Second, a choice of $\mathbf{z}$ corresponds to an assumption about the extent to which borrowing constraints bind, i.e., about the maximum number of children each family type can afford to send to college. In equilibrium the assumed value of $\mathbf{z}$ must in fact be consistent with the family budget constraints implied by the steady-state levels of skilled and unskilled wages corresponding to the value of $\widehat{\beta}$. Hence, $\mathbf{z}$ and $\widehat{\beta}$ must jointly satisfy, $\forall j$:

$$(2 + f_j)\overline{c} + z_j \nu \leq I_j(\widehat{\beta}) \tag{2.14}$$

and either of the equations below:

$$\begin{array}{rcl} (2 + f_j)\overline{c} + (z_j + 1)\nu & > & I_j(\widehat{\beta}) \\ \text{or} \qquad z_j & = & f_j \end{array} \tag{2.15}$$

Furthermore, we will restrict our attention to locally stable steady states, so an additional constraint to be met is $\frac{\partial \beta_{t+1}}{\partial \beta_t} \Big|_{\beta_t = \widehat{\beta}} < 1$.

3. Changes in Sorting

How will a change in the degree of sorting (i.e., in the level of θ) affect the steady-state level of β? In answering this question it is useful to distinguish two cases: one in which the change in sorting does not affect the maximum number of children any family type can afford to send to college and the other in which it does. In the first case what we will call the "bindingness" of borrowing constraints is not affected; in the second case it is.[9]

[9] In a model with a continuous income distribution, there would always be a change in the bindingness of borrowing constraints for some families (as long as some of them were constrained

Assume initially that the bindingness of constraints is not affected (i.e., assuming that $\mathbf{z}$ does not change and hence that the Γ_j's are constant). Using the implicit function rule on (2.13) yields:

$$\frac{d\widehat{\beta}}{d\theta}=\frac{\sum_j f_j \frac{\partial \lambda_j(\widehat{\beta};\theta)}{\partial \theta}(\Gamma_j-\widehat{\beta})}{\sum_j f_j \lambda_j(\widehat{\beta};\theta)+\sum_j f_j \frac{\partial \lambda_j(\widehat{\beta})}{\partial \widehat{\beta}}(\widehat{\beta}-\Gamma_j)} \tag{3.1}$$

Taking the derivatives of the λ_j's (given by (2.2), (2.4) and (2.5)), evaluating at $\beta=\widehat{\beta}$, and substituting into the expression above yields:

$$\begin{aligned}\frac{d\widehat{\beta}}{d\theta} &= \frac{\widehat{\beta}(1-\widehat{\beta})[f_h(\Gamma_h-\widehat{\beta})-2f_m(\Gamma_m-\widehat{\beta})+f_l(\Gamma_l-\widehat{\beta})]}{D} \\ &= \frac{\widehat{\beta}(1-\widehat{\beta})[(f_h\Gamma_h-2f_m\Gamma_m+f_l\Gamma_l)-\widehat{\beta}(f_h-2f_m+f_l)]}{D}\end{aligned} \tag{3.2}$$

where $D=\sum_j f_j\lambda_j(\widehat{\beta};\theta)+\sum_j f_j \frac{\partial \lambda_j(\widehat{\beta};\theta)}{\partial \beta}(\widehat{\beta}-\Gamma_j)$. It is easy to show that local stability requires:

$$\frac{\sum_j f_j \frac{\partial \lambda_j(\widehat{\beta};\theta)}{\partial \beta}(\Gamma_j-\widehat{\beta})}{\sum_j f_j\lambda_j(\widehat{\beta};\theta)}<1 \tag{3.3}$$

implying that D is positive.

Note that one way to think about what an increase in sorting does is that for every two middle-type marriages it destroys, it creates one high and one low-type marriage. With this in mind, note that an interpretation of (3.2) is that increased sorting increases the steady-state fraction of the population that attends college if the result of substituting two middle types by one high and one low type on net increases the number of children that attend college by more than what would result from that same substitution and all three types sending a fraction $\widehat{\beta}$ of their children to college.

in the initial equilibrium). Thus this second channel would always be present. In our discrete model, whether constraints become more binding depends on the cost of college relative to family income. Thus, small changes in sorting may not affect the extent to which families are constrained.

It is easy to evaluate (3.1) or (3.2) in a few special instances. Consider first the case where the Γ_j's are constant across family types, i.e., $\Gamma_j = \Gamma$, $\forall j$. Then, by (2.13), $\widehat{\beta} = \Gamma$, and a change in the degree of sorting has no effect on the economy (e.g., if all family types send 10% of their children to college, the steady-state fraction of the population that attends college is 10%, irrespective of the degree of sorting).

Next consider the case where the Γ_j's are not identical but where fertility is constant across family types, i.e., $f_j = f$. In such case the sign of (3.1) is given by the sign of $\Gamma_h + \Gamma_l - 2\Gamma_m$. The intuition behind this is simple given the earlier observation: Since fertility is the same across family types, the effect of increased sorting depends on whether the fraction of children sent to college on average by two middle-type marriages ($2\Gamma_m$) is smaller than the combined fraction of children that go to college on average in one high and one low type family ($\Gamma_h + \Gamma_l$).

Another case for which it is relatively easy to derive an expression is for when $\Gamma_h + \Gamma_l - 2\Gamma_m = 0$ and $f_h + f_l - 2f_m = 0$. In this case, after manipulating (3.2), it is easy to see that the sign of the effect of an increase in θ is given by the sign of $(\Gamma_h - \Gamma_l)(f_h - f_l)$. This is an interesting case since it implies that if both fertility and the probability of attending college are linear in parents' average years of education, the effect of increased sorting is to decrease the fraction of the population that attends college if children of high-type parents have a greater probability of attending college and if the fertility of low-type parents is greater than that of high types.

Lastly, it is useful to note from (3.2) that a sufficient condition for increased sorting to impact negatively on $\widehat{\beta}$ is for $f_h\Gamma_h - 2f_m\Gamma_m + f_l\Gamma_l \leq 0$ and $f_h + f_m - 2f_l \geq 0$ (with at least one inequality strict). The first expression captures whether the number of children that on average attend college is increased or decreased by substituting two m couples by an h and an l. Thus, it indicates by how much the population that attends college would increase given this substitution. The second expression captures the amount by which the population as a whole is increased or decreased by substituting two m couples by an h and an l. Obviously, a decrease in the population attending college will, ceteris paribus, serve to reduce $\widehat{\beta}$, as will an increase in the overall population (since it dilutes further the gain/loss of the first term). As we shall see further on, our parameterization implies that both inequalities hold strictly and hence that increases in sorting decrease the fraction of the population that goes to college.

As mentioned previously, the degree of sorting can also affect the steady-state level of $\widehat{\beta}$ via its effect on the tightness of borrowing constraints. To see this, suppose that, keeping $\mathbf{z}$ constant as before, an increase in θ decreases $\widehat{\beta}$. This smaller proportion of skilled workers is associated with lower unskilled wages and higher skilled wages. The change in wages will increase family income for high types and decrease it for low types, and thus may lead to less binding constraints for the first group and tighter ones for the second. Should this happen, the original equilibrium values of the z_j's and hence of the Γ_j's would no longer be feasible and $\widehat{\beta}$ will fall even further. That is, a change in the degree of sorting can affect the feasibility (in steady-state equilibrium) of different values of $\mathbf{z}$.

4. Parameterizing the Model

In this section we parameterize our model. We choose parameters so that the cross-section data generated in a steady state of the model are consistent with similar cross-section relationships that hold in actual US data. This ensures that the reduced-form relationships implied by this steady state of the model are "reasonable".

Recall that in the model there are three types of marriages–high, middle and low–which differ in both the average education and the average income of the couple. Each type of marriage j is further characterized by two numbers: f_j, the number of children per couple, and γ_j, the fraction of kids (on average) from that marriage type that have the aptitude to benefit from skill acquisition. These two profiles are central to our analysis, so much of our discussion will focus on them.

Fertility

It is empirically well-established that fertility rates are negatively correlated with both income and education. Even for the US, the magnitude of these differences are fairly large, especially between the lower and upper end of the distributions. For example, using data from the CPS for 1995, (Table 103 from the 1997 Statistical Abstract of the US) the cross-sectional fertility rate for women aged 15-44 is roughly one and a half times larger in the lowest quartile of the household income distribution than it is for the top quartile of the household income distribution.[10] Similar magnitudes are found in the relationship between parental education and fertility. Using data from the PSID we find that couples in which

[10] Of course, annual income measures do not necessarily reflect where a household lies in the distribution of lifetime income. An alternative comparison using lifetime income is computed by Knowles (1999) using PSID data. His numbers indicate that fertility in the lowest quintile is roughly one and a half times that for the highest quintile.

neither parent finished high school (roughly the bottom fifth of the education distribution) have approximately 1.5 times the number of children (2.65 versus 1.82) as do couples in which both parents have at least some college (roughly the top fifth of the sample).[11] The relationship is not linear, however. Couples in which both parents have high-school educations have only marginally higher fertility (1.88) than do couples in which both parents have college degrees (1.79).

The evidence above suggests modelling the fertility of low-type couples as 1.5 times that of high-type couples, which is what we elect to do.[12] We choose fertility rates of 2, 2, and 3 for high, middle and low types respectively for our benchmark model.[13] As will be seen further on, this choice yields a steady state in which low-type marriages constitute roughly 25% of the population and hence represent the bottom quartile of both the income and education distribution. Thus, this parameterization generates the facts cited above as well as respecting the fact that fertility differentials become fairly small outside the bottom quartile of the distribution. Nonetheless, we also carry out a sensitivity analysis to our choice of fertility profile. As we report, even abstracting completely from any fertility differences we still obtain large quantitative effects resulting from changes in the degree of sorting.

Aptitude

We have no direct measure of the fraction of children from marriages of different types that have an aptitude for skill acquisition. The PSID, however, contains data on the relationship between the educational attainments of parents and their children. We create a sample by selecting all individuals over 25 in the 1993 PSID whose parents were in the PSID in 1968. For the purposes of this exercise we split this sample into skill categories by counting all individuals with high school or below as unskilled, and all individuals with some college or above as skilled. We find that the fraction of children from high-type families that become skilled is .81, whereas the values for middle and low-type families are .63 and .30 respectively.

[11] Some studies argue that maternal education is a more important determinant of fertility than is average parental education. Again, using data from the PSID we find that fertility rates of women with less than high school, high school and some college and above are 3.6, 2.9 and 2.5 respectively. Once again, the ratio of top to bottom is about 1.5.

[12] In making this choice there is an implicit assumption that recently educated cohorts, which have roughly 55% of their members going beyond high school, will have a distribution of fertility across quartiles that resembles those of older cohorts.

[13] As we have chosen to model all families of the same type as having the same number of children, we are restricted to integer choices for our fertility rates. It would be easy to modify this but we chose not to so so as it would introduce considerably more notation.

Thus, in the steady state of our model the Γ_j's must match these values.

As is evident from equation (2.12), Γ_j is a function of γ_j and z_j. Thus, the probability that a child from a particular marriage type is of high aptitude (i.e., the γ_j) can be deduced from the value of Γ_j in conjunction with an assumption about the maximum number of children that family type can afford to send to college (i.e., the z_j) in the steady state. Table 1 illustrates this mapping by showing the values of the γ_j's implied by various assumptions regarding the tightness of borrowing constraints subject to the constraint that each (γ_j, z_j) pair yield the aggregate Γ_j found in the data.

The first column of Table 1 corresponds to a case in which no one is constrained—all high-aptitude children become skilled. In this case the values of Γ_j's and γ_j's must coincide. The second column assumes that only low-type families are constrained and that these can afford to send at most two kids to college. Lastly, the third column assumes that low-type families can afford to send only one child to college.[14]

Table 1
Aptitude Profiles Under Various Scenarios

Aptitude	$z_h = z_m = 2, z_l = 3$	$z_h = z_m = z_l = 2$	$z_h = z_m = 2, z_l = 1$
γ_h	.81	.81	.81
γ_m	.63	.63	.63
γ_l	.30	.31	.33

Our analysis, for the most part, is independent of which of these scenarios we take to represent the steady state. If the maximum number of children that different family types can afford to send to college remains unchanged when the degree of sorting increases, then as equation (3.2) indicates, the effect of sorting depends only on the Γ_j's which are given by the data; the mix of γ_j's and $\mathbf{z}$ used to generate them is irrelevant. It is only when we allow the equilibrium steady-state value of $\mathbf{z}$ to be affected by the increased degree of sorting that the exact specification might matter. But even in that case all that matters to our results,

[14] Note that our benchmark specification does not allow us to consider the case of either h or m type families being constrained (or more generally that all variation across family types is a result of borrowing constraints), since the assumption that high and middle families have two kids implies that any constraint would have less than 50% of them attending college. Since in the data more than 50% of children from these family types attend college, this option is not feasible. One way to get around this would be to assume (more realistically) that the number of children a family has is drawn from a distribution that differs according to family type.

as will be seen in the next section, is that there be a large group of individuals that are affected by a tightening of the borrowing constraints.

We choose the second column of the table for our benchmark specification; i.e., we assume that borrowing constraints do not affect middle and high marriage types, but that low types are able to send at most two of their kids to college. This is actually a relatively mild constraint. Since only families who have three high-aptitude children are constrained, this implies that only 3% of low-type families are affected by the borrowing constraints in the steady state. Of course, as outlined in the previous section, it is necessary to check that our assumptions on credit constraints are consistent with wages, consumption requirements and the cost of skill acquisition. We leave this for later in the analysis.

Sorting

Next we assign a value to θ, the fraction of marriages which involve perfect sorting. In our model this is the correlation between the education levels of spouses. We use our sample from the PSID to obtain an estimate of this correlation for the US, yielding $\theta = .6$.

Steady-State Determination of β

Given the values assigned thus far we can solve for $\widehat{\beta}$ (the fraction of population that goes to college in the steady state). Doing so, we obtain $\widehat{\beta} = .55$. This turns out to be a close match with the data. According to data from the CPS for 1996, among individuals aged 25-34, roughly 55% have at least some college. Moreover, this fraction is basically the same for those individuals aged $35-44$ and $45-54$.[15] We take this as an indication that the implications of our parameterized model for educational attainment in the steady state are reasonable.

Production Function

It remains to specify the production function. We choose a constant elasticity of substitution production function:

$$y = A[bn_s^{\rho} + (1-b)n_u^{\rho}]^{\frac{1}{\rho}}$$

Note that the ratio of skilled to unskilled workers can be written as $\frac{\beta}{1-\beta}$, and that the relative wage of skilled to unskilled workers is given by $\frac{w_s}{w_u} = \frac{b}{1-b}(\frac{\beta}{1-\beta})^{\rho-1}$. As is well known, the ratio of skilled to unskilled wages has varied considerably over the last 30 years in the US.[16] Recall that our two skill groups are those with at

[15] Source: Statistical Abstract of the US, 1997, Table 245.
[16] See, e.g., Katz and Murphy (1992).

least some college and those with high school or less. Based on the data in Katz and Murphy (1992), we match a ratio of 1.9 for our benchmark case. This value is at the upper end of what has been observed in the US, so in our robustness check we redo our analysis assuming a ratio of 1.4 and find that it has no impact on our results.

There is a literature that attempts to estimate the degree of substitutability between skilled and unskilled labor that we can use to provide an estimate for ρ. This literature suggests that a reasonable elasticity of substitution is 1.5, which implies $\rho = .33$.[17] Using this value of ρ and matching the above-mentioned value for the skill premium implies $b = .6865$. As we will see shortly, the elasticity of substitution is a key parameter for our analysis—if we use a value that is substantially larger, it becomes much harder for our model to generate large effects from changes in sorting. Lastly, for ease of interpretation of our results, we choose a value of A to scale steady-state unskilled wages to some "reasonable" value, which we set to be $30,000$. This is purely an issue of normalization.[18]

Other Properties of the Steady State

Having assigned parameter values, we can solve for the steady state in which educational attainment is dictated by the observed values of the Γ_j's as discussed previously. We now report some additional properties of this steady state. The model produces a distribution for individual income, with mass at two points, corresponding to the skilled and unskilled wage rates. The standard deviation of log income in the steady state equals .32. Distributions of annual income in the US typically imply a value of around .6 for this figure. Alternatively, the lifetime income distribution generated by Fullerton and Rogers (1993) using PSID data yields a value around .4. Since we are relying entirely on the skill premium to generate our variation in income it is not surprising that we produce less variation than is found in the data.

Lastly, we can also compute the standard deviation and mean of the educational attainment distribution. We assume that a high-school education corresponds to $e = 11.3$ and a college education corresponds to $e = 15.0$, our choice of numbers given by the average educational attainments of children with high school or less and those with college or more in our PSID sample. The resulting standard deviation and mean of the steady state educational attainment distri-

[17] See Krusell et al. (1997) and the references therein for more detail.

[18] It is not clear what the "best" normalization is, since in our model these are lifetime earnings. Rescaling of this variable, of course implies that the parameters $\bar{c}$ and ν need to be scaled accordingly as well.

bution are equal to 1.84, and 13.3. For our sample of children from the PSID the corresponding values are 2.56 and 12.9. Given our restriction to two levels of education it is not surprising that we generate less variation than the data. The fact that our mean is somewhat higher is related to the fact that it is the steady-state value. Even if recent cohorts have had relatively constant educational attainments, average educational attainment continues to increase as older generations are replaced with younger generations.

5. The Effects of Increased Sorting

We now use the parameterized model to assess the effects associated with an exogenous increase in the degree of sorting in the marriage market. Our objective is to examine whether the concern that some writers have expressed—namely, that increased sorting will lead to increased inequality—has any significant quantitative support (in contrast with Kremer's findings). To this end, we consider a large change in the degree of sorting: from $\theta = .6$ to $\theta = .9$.[19]

Table 2 displays the main results. The first column gives the values for the original steady state (i.e., $\theta = .6$). The second column shows what the steady state would look like if θ were to increase to .9 and the tightness of borrowing constraints were unchanged (i.e., all high-aptitude children from middle and high-type families could afford to attend college but among low-type families at most two children per household could be sent to college). The third column reports the new steady-state values ensuing from the θ change, but assumes that the wage change associated with this increase tightens constraints for low-type families to the point that they can afford to send at most one of their children to college.[20] Below we discuss each case in turn.

[19]Note that Kremer (1997) also considers a large change in the degree of sorting.

[20]Later in this section we show that this outcome is consistent with choices for $\bar{c}$ and ν.

Table 2
Effects of Increased Sorting on Steady State

	$\theta = .6, \Gamma_l = .30$	$\theta = .9, \Gamma_l = .30$	$\theta = .9, \Gamma_l = .22$
$mean(e)$	13.3	13.2	12.8
$std(e)$	1.84	1.85	1.82
$cv(e)$	.138	.140	.142
$\widehat{\beta}$	.55	.52	.42
n_s/n_u	1.236	1.076	.713
w_s/w_u	1.900	2.085	2.746
w_u	30,000	28,120	23,340
$std(log\ y)$	.319	.367	.500

Case 1: Set of Borrowing Constrained Individuals Remains Constant

We begin by comparing the first two columns. This amounts to examining the effects of increased sorting holding the pattern of college attendance fixed. The first three rows report the mean, standard deviation, and coefficient of variation for the steady-state distribution of educational attainment. The effect of the increase in sorting is to cause a small decrease in the mean of the distribution (less than one percent), an even smaller increase in the standard deviation (less than one half of one percent) and a roughly one percent increase in the coefficient of variation.[21]

The decrease in mean educational attainment results from the decrease in the fraction of the population that goes to college; the fourth row shows that the fraction of the population that becomes skilled falls to 52% in the new steady state. It is important to underscore that this fall in β implies a decrease in the ratio of skilled to unskilled workers of almost 13%, as shown in row 5. As the next row indicates, this change in relative labor supply induces an increase in the skill premium of 10%.[22] The next to last row shows that the standard deviation of the log income distribution increases by about 15%.

[21] Note that since our educational attainment distribution is over two levels, its variance is maximized when the population is evenly distributed across them. Hence, whether a change in $\widehat{\beta}$ results in an increase or decrease in the standard deviation of education depends entirely upon whether the starting value was above or below .5. Having said this, we think that what is most relevant to notice is that the change in the standard deviation is very small, rather than the direction in which it changes.

[22] Recall that the elasticity of substitution in the production function is 1.5, which implies that the percent change in n_s/n_u will be 1.5 times as large as the percent change in w_s/w_u for small changes. For large changes this expression continues to hold exactly in logs, but only aproximately in ratios.

Case 2: Set of Borrowing Constrained Individuals Changes

In the case above we assumed that borrowing constraints did not become more binding as a result of the increase in the degree of sorting. As shown in row 7, this resulted in a decrease in the wage of unskilled workers by almost $2000, and hence a decrease in low-type family income of almost $4000. This wage decrease makes it possible that low-type families will be able to send fewer children to college than previously and hence that the equilibrium steady-state value of $\mathbf{z}$ used in column two is no longer a feasible one. Consequently, in the third column we assume that as a result of the θ increase, in the new steady state low-type families can afford to send a maximum of one child to college, rather than two. This constraint affects families with two and three high-aptitude children (respectively 20 and 3 percent of low-type families approximately), causing the fraction of children from low-type marriages that go to college to drop from .30 to .22 (as indicated by the reported values of Γ_l). Thus, the new steady-state equilibrium is now given by $\widehat{\beta} = .42$ (and the equilibrium shown in the second column is eliminated since at $\beta = .52$, z_l should equal one rather than two).

As the table shows, the tightening of borrowing constraints has a dramatic effect on how the θ increase affects the income distribution. In particular, although the change in the mean level of education is still relatively small (a bit under 4%), this masks a dramatic change in the fraction of the population that goes to college which falls by 42% implying a drop of more than 40% in the ratio of skilled to unskilled workers. The skill premium (w_s/w_u) also increases by more than 40%, to 2.75, and the standard deviation of distribution of log income increases by more than 50%!

We next verify that the structural change in college attendance decisions is a feasible equilibrium outcome. As discussed in Section 3, this requires showing that certain inequalities are satisfied. In what follows let $w_u(\widehat{\beta}_i)$ be the unskilled wage rate when the equilibrium value of skilled to unskilled workers, $\widehat{\beta}$, is given in column i of Table 2.

For column 1 to represent an equilibrium steady state, it must be that type l families can send two but not three children to college. This requires (i) $2\nu + 5\bar{c} < 2w_u(\widehat{\beta}_1) < 3\nu + 5\bar{c}$.[23] For the allocations in column 2 to be infeasible because at those wages l-type families cannot afford to send two children to college, requires (ii) $2\nu + 5\bar{c} > 2w_u(\widehat{\beta}_2)$. Lastly, to ensure that the outcome in column 3 is an equilibrium requires checking that it allows type l families to send one child and

[23] Of course, if low-type families can afford to send two children to college, so can higher-type families.

type m families to send two to college. That is, it requires (iii) $\nu + 5\bar{c} < 2w_u(\widehat{\beta}_3)$ and (iv) $2\nu+4\bar{c} < w_u(\widehat{\beta}_3)+w_s(\widehat{\beta}_3)$. Inequalities (ii) and (iii) imply $\nu > 2(w_u(\widehat{\beta}_2) - w_u(\widehat{\beta}_3))$, so that $\nu > 9,402$ given the numbers in Table 2. There are many combinations of ν and $\bar{c}$ that satisfy these inequalities. For example, $\nu = 11,000$, and $\bar{c} = 7,000$. A last inequality to check is that high-ability individuals prefer to go to college over high school. It is easy to verify that this is indeed the case.

We do not attach too much significance to the magnitudes of ν and $\bar{c}$. The simple choices that we made about utility functions and the fact that we abstract from life-cycle income dynamics and the timing of college attendance make us reluctant to do so as does our unwillingness to interpret the borrowing constraints literally as the ability to afford college. The main point of the above paragraph is to establish the logical consistency of our argument that the change in sorting can lead to a change in the extent to which credit constraints bind. It is perhaps not surprising that this can be done, given that we have not imposed any discipline on our choices of $\bar{c}$ and ν.[24]

6. Discussion and Comparison With Kremer

To summarize the main result of the preceding section, we find in our calibrated model that a large increase in the degree of sorting may be expected to produce substantial changes in inequality. This is true independently of whether one believes that borrowing constraints play any role in the economy, as the results in column two of Table 2 demonstrate. These effects are significantly magnified if borrowing constraints are tightened as indicated by the last column of Table 2.

Our results support a conclusion very different from that reached by Kremer (1997). Whereas he concluded that a large increase in sorting would have little effect on steady-state inequality given a reasonable parameterization, we have concluded quite the opposite. In this section we analyze what lies behind this difference. Having identified the factors that generate such different conclusions,

[24]Having said this, however, we do offer one check of "reasonableness" for the value of ν. Specifically, we can compute the annual rate of return to spending on education. This obviously depends on how many years one assumes there are between the expenditure and the return in the form of higher wages, since the expenditure takes place in the first period of life but yields higher wages in the second period. If one interprets a period to be a generation, then 20 years may be reasonable. On the other hand, if one wants to look at the time between college expenditures and the midpoint of a typical working life, then a slightly smaller period length may be appropriate. In any case, for the steady state in column 1, the annual rate of return lies between 5 and 10 percent as we vary the number of years between 10 and 20.

we then examine the robustness of our results to different specifications. We first turn to a brief review Kremer's analysis.

6.1. Review of Kremer '97

Kremer posits an intergenerational model of marriage, fertility and educational attainment in which a child's educational attainment e can be written as a linear function of parental and neighborhood average education. For expositional purposes we consider the argument in the simplest context, and hence abstract from neighborhood effects.

The model assumes that all individuals marry and have two kids. A child's educational attainment is determined by the following linear relationship:

$$e_{i,t+1} = \kappa + \alpha \frac{(e_{i,t} + e_{i',t})}{2} + \varepsilon_i$$

where $e_{i,t+1}$ is the education level for the child, $e_{i,t}$ and $e_{i',t}$ are the education levels of the two parents, and ε is a normally distributed random shock that is *iid* across families, with mean 0 and standard deviation equal to σ_ε. An exogenous (assortative) matching of individuals takes place such that ρ_m is the correlation between the education levels of parents.

Assuming that parameter values are constant over time, it is straightforward to characterize the steady-state distribution of education. Specifically, it will be normally distributed, with mean and standard deviation given by:

$$\mu_\infty = \frac{\kappa}{1-\alpha}$$

$$\sigma_\infty = \frac{\sigma_\varepsilon}{[1 - \alpha^2(1+\rho_m)/2]^{.5}}$$

Kremer's objective was to determine how changes in sorting among marriage partners (i.e., ρ_m) would affect the level of inequality in the steady state. His main measure of inequality was the standard deviation of educational attainment and he argued that since there is a linear relationship between educational attainment and log of income in the cross-section, that this measure of inequality would probably be a good proxy for inequality in log of income as well. We shall return to this point later.

The effect of an increase in ρ_m on the steady state distribution of education can be read off of the above equations. Because of the assumption of linearity, there is no effect of ρ_m on the mean of the distribution of education, but its standard deviation is increasing in ρ_m, i.e., increases in the degree of sorting among parents will increase the standard deviation of education. Obviously, this model is at least qualitatively consistent with the view that increased sorting leads to increased inequality.

Kremer's main contribution, however, was to show that while the model supported this view qualitatively, there was little support for the view that this effect was important quantitatively. It is easy to see that in this model the percent change in the standard deviation of income due to a change in the sorting parameter ρ_m is determined solely by the magnitude of the parameter α. Using data from the PSID (the same source that we used to parameterize our model) he obtained an estimate of α of about .4 and $\rho_m = .6$.[25] In this case, an increase in ρ_m from .6 to .9 would result in only a 1.4% increase in the standard deviation of education.[26]

Next Kremer argued that even if his estimate of α were somewhat off, his conclusion would survive due to the insensitivity of his result to modest changes in α. The easiest way to see this is by asking how large α would need to be in order that an increase in ρ_m from .6 to .9 to result in a 10% increase in the standard deviation of log income. The answer is given by solving the equation:

$$\frac{1.1}{[1-\alpha^2(1+.6)/2]^{.5}} = \frac{1}{[1-\alpha^2(1+.9)/2]^{.5}}.$$

which yields .78 as the required value of α.

Kremer's paper is mainly about the effect of increased neighborhood and marital sorting on the distribution of education. If, however, one takes the view (as Kremer does in his introduction) that log earnings are approximately linear in years of education, and that the coefficients in this relationship are invariant to changes in the distribution of education, then the same conclusion applies to inequality in income; a large increase in sorting will not significantly affect income inequality in the US.

[25] When neighborhood effects were included the sum of coefficients on parental and neighborhood education was about .55. This does not change his conlusions.

[26] In fact had ρ_m increased from an original value of .1 to .9, this still would only result in a 3.7% increase in the standard deviation of log income.

6.2. Discussion

It should clear from our review of Kremer's work that his main finding is really about the small impact of sorting on the level of inequality in the *skill* distribution. Our results in Table 2 do not contradict this finding, especially if we assume that the increase in sorting does not affect the bindingness of borrowing constraints. To further demonstrate that there is no inconsistency between our results and his we perform his analysis on data generated from our model. Specifically, using data generated by the steady-state of our calibrated model (i.e., column 1 in Table 2), we take a random sample of 1200 families and run a regression of a child's educational attainment ($e_{i,t+1}$) on a constant and the average educational attainment of its parents ($\bar{e}_{i,t}$).[27] As noted previously, we assume that a high-school education corresponds to $e = 11.3$ and a college education corresponds to $e = 15.0$.[28] We do this 100 times and average across the trials. The result of this exercise is:[29]

$$e_{i,t+1} = 6.69 + .51\bar{e}_{i,t}$$

It follows that if Kremer had performed his exercise using data generated from our model he would still have reached the same conclusion; i.e., he would have concluded that the coefficient on average parental education is too small to generate large effects on the standard deviation of educational attainment.[30]

What gives rise to our very different conclusion about income inequality is that in our model there is an interaction between changes in the skill distribution and the price of skill. This interaction is governed by three elements that are absent in Kremer's analysis but which are central to generating this effect on the price of skill: (i) The existence of a nonlinear relationship between parental years of education and those of their children, (ii) A negative correlation between fertility and parental education, and (iii) Wage rates that are sensitive to changes in the skill distribution. As we shall see, it turns out that if we had only incorporated any one of these three elements, we would have reached the same conclusion as

[27] We note that since our model is not linear, this regression is not suggested by our model.

[28] The specific values chosen here affect the constant term in the regression but have very little effect on the coefficient on parent's education.

[29] Running this regression on our sample from the PSID yields a coefficient of .37. This discrepancy is accounted for by the fact that in our model we compress the education distribution to two levels, thereby increasing the correlation between the education levels of parents and their children.We have verified this via simulation.

[30] Note that since our model is not linear, it does not lead one to run this regression. We run this regression simply to illustrate how Kremer's analysis would look in our set-up.

Kremer.[31] But, allowing for the interaction of all three factors (especially (i) and (iii)) leads to a very different conclusion.

We begin with a discussion of the third factor. The distribution of labor earnings can be thought of as depending on the interaction of two factors. One is the distribution of skill (in our model, education) across individuals, and the second is the price of skill (i.e., the skill premium). As stated in our discussion of Table 2, the impact of sorting on the level of inequality in the skill distribution is small. In fact, were wages not responsive to the distribution of skills, the change in the standard deviation of log income would have been around one-half of one percent. What drives our results is that a large change in sorting produces a large change in the skill premium, even if it seemingly does not produce "large" effects on mean educational attainment. As can be seen from a comparison of columns one and two in Table 2, a less than one percent decrease in the mean of the education distribution is associated with an almost thirteen percent decrease in the relative supply of skilled labor. This translates to a ten percent increase in the wage premium, leading to a significant change in the distribution of income.

To better understand how various elements interact to yield the increase in the skill premium, note first that in our model the impact of a change in θ on w_s/w_u can be decomposed into two distinct effects. The first concerns how a given change in θ affects $\widehat{\beta}$, and the second with how a given change in $\widehat{\beta}$ affects w_s/w_u. This decomposition is useful because college attendance and fertility profiles are only relevant for the first effect, whereas the elasticity of substitution in the production function is only relevant for the second.[32]

Consider now the roles of fertility differences and of the function relating parental education to children's education in generating the change in $\widehat{\beta}$. Recall from the discussion in Section 3 that a sufficient condition for increased sorting to impact negatively on $\widehat{\beta}$ is for $f_h\Gamma_h - 2f_m\Gamma_m + f_l\Gamma_l \leq 0$ and $f_h - 2f_m + f_l \geq 0$ (with at least one strict inequality). Our parameter values strictly satisfy both

[31] In fact, Kremer considers a markov model in section IV of his paper, and finds little effect of sorting on the standard deviation of education. This is obviously consistent with our findings. We have also rewritten Kremer's model to account for differential fertility and used numerical techniques to compute the steady-state distribution of education. Once again, changes in sorting have little effect on the standard deviation of this distribution.

[32] In particular, the change in $\log w_s/w_u$ equals the inverse of this elasticity times the change in $\log \widehat{\beta}/(1-\widehat{\beta})$. Moreover, how this change in w_s/w_u is split between changes in each of the two wages is entirely determined by the elasticity and the initial value of $\widehat{\beta}$. Had we assumed $\rho = 1$, i.e., a linear production function, there would be no effect of sorting on wages and, as discussed previously, we would have found very small effects from increased sorting on inequality.

inequalities, guaranteeing that increased sorting will decrease the fraction of the population that attends college. The magnitude of the respective contributions of our fertility profile and the concavity of the intergenerational education transmission function will be discussed in the next section on robustness.

One can ask under what conditions our model would give rise to the conclusion that changes in sorting do not have significant effects on the income distribution (without shutting down the effect of changes in skill distribution on wages). A simple condition is given by the combination of a linear relationship between parents' education and children's' (i.e., $2\Gamma_m = \Gamma_h + \Gamma_l$) and no fertility differentials (i.e., $f_j = f$ for all j). But these are precisely the assumptions made by Kremer in his paper–all parents have two kids and the child's years of education are linear in average parental years of education. Thus, had we adopted Kremer's assumptions our model would not have generated any effect from increased sorting on the steady-state value of β, and hence no effect on wage rates or inequality either. Moreover, the fact that wage rates would not have changed would necessarily imply that the bindingness of borrowing constraints would be unaffected and consequently there would be no scope for any change in college attendance decisions via this channel either.

Lastly, our analysis also suggests that one exercise caution in interpreting regressions of child's educational attainment on parental educational attainment. In our discussion of Kremer's work, this regression coefficient was denoted α and was treated as a structural parameter that would not be affected by changes in sorting. However, as should be clear from our model, the degree to which education is heritable may differ across family types for a variety of reasons including the presence of borrowing constraints. The degree of sorting, as evidenced in the last column of Table 2, affects the bindingness of borrowing constraints and hence the degree to which parents' education is passed on their children.[33]

7. Robustness

In this section we report how the findings from our benchmark model are affected by changes in our parameterization. We restrict our attention to the results generated under the assumption that the bindingness of borrowing constraints

[33] Running a linear regression for the steady state in column three of Table 2 (using the same procedure described earlier), we obtain .59 rather than the .51 obtained for the scenarios in columns one and two, although the true "heritability" of education is unchanged as reflected in the γ_j's.

are unchanged by the degree of sorting. Our main finding is that our result of a quantitatively important increase in income inequality arising from (large) changes in the degree of sorting is robust to reasonable variation in the model's parameterization.

We begin by considering how alternative profiles for fertility affect our results. An issue with our choice of fertility profile in the benchmark model is that for other purposes we generally interpret low-type marriages as those in which neither parent has gone beyond high school. Our sample from the PSID has completed fertility profiles only for those individuals that we have designated as parents (recall that these are individuals in the PSID with children older than 25 in '93). Thus, our parents are from fairly old cohorts and the fraction of this group with high school or less is in fact quite large (over 55%). The fertility differential across educational classes for these cohorts is consequently lower: 2.26 versus 1.86 for all other couples–a difference of a bit over 21 percent (rather than the 50% that we have used). While it is true that the bottom quintile of the sample does have roughly 50% higher fertility, this quintile would be comprised of those that have less than a high school education. As we report below, however, even abstracting completely from any fertility differences we still obtain large quantitative effects resulting from changes in the degree of sorting.

We examine two alternatives for the fertility profile (f_h, f_m, f_l)—(2,2,2) and (2,2,4)—to the profile (2,2,3) used in our benchmark model.[34] In each case we recalibrate our model to match the same statistics as before, and perform the same comparative statics exercise as in the movement from column 1 to column 2 in Table 2. That is, leaving the bindingness of constraints unchanged, we examine the effects of an increase in θ from .6 to .9. As can be seen in Table 3 below, the basic message is the same for both of the alternative fertility profiles. Even with no fertility differences the increase in income inequality is still substantial, albeit somewhat less than in Table 2 (10% versus 15%). For the case in which low-type families have four kids the increase is slightly more than 12%. One might have expected larger effects for this case, but the fact that the initial steady-state value is substantially lower as a result of the recalibration translates into a smaller change in the ratio of skilled to unskilled workers.

[34] We also note that using fertility rates from the cross section may understate the impact of fertility differences for some purposes. The model does not allow for the fact that lower income families have their first child some five years before richer families. This would increase the relative size of the poorer group in steady state by more than what would be predicted based solely on differences in the number of children. See Knowles (1999) for details.

Table 3

Effect of Alternative Fertility Profiles

$f_h = f_m = 2, \Gamma_h = .81, \Gamma_m = .63, \Gamma_l = .30$

	$f_l = 2$		$f_l = 4$	
	$\theta = .6$	$\theta = .9$	$\theta = .6$	$\theta = .9$
mean(e)	13.7	13.6	13.1	13.0
std(e)	1.78	1.80	1.85	1.84
cv(e)	.130	.132	.142	.141
$\hat{\beta}$	.64	.62	.49	.46
n_s/n_u	1.781	1.634	.945	.835
w_s/w_u	1.900	2.018	1.900	2.063
w_u	30000	28640	30000	28466
std($\log y$)	.308	.341	.321	.361

Next we examine how our findings are affected by changes in the profile of Γ_j's used in the calibration. In Table 4 we examine the effect of varying Γ_m from its value of .63 in our benchmark model by decreasing the degree of concavity in the relationship between parental and children's education to the point where it is linear ($\Gamma_m = .555$). In each case the production function parameters are recalibrated to match the same statistics as before.

Table 4

Effect of Alternative Γ_m's

$f_h = f_m = 2, f_l = 3, \Gamma_h = .81, \Gamma_l = .30$

	$\Gamma_m = .63$		$\Gamma_m = .58$		$\Gamma_m = .555$	
	$\theta = .6$	$\theta = .9$	$\theta = .6$	$\theta = .9$	$\theta = .6$	$\theta = .9$
mean(e)	13.3	13.2	13.3	13.2	13.2	13.2
std(e)	1.84	1.85	1.85	1.85	1.85	1.85
cv(e)	.138	.140	.139	.140	.140	.140
$\hat{\beta}$	.55	.52	.54	.51	.53	.51
n_s/n_u	1.24	1.08	1.15	1.06	1.11	1.05
w_s/w_u	1.9	2.09	1.9	2.01	1.9	1.97
w_u	$30,000$	$28,120$	$30,000$	$28,820$	$30,000$	$29,253$
std($\log y$)	.319	.367	.320	.349	.321	.340

As can be seen, as we move closer to the linear case, the increase in the standard deviation of log income caused by an increase in sorting becomes smaller, but even in the linear case this measure increases by more than 6 percent. Increasing the degree of concavity in this profile, on the other hand, would have the effect of increasing the impact of a change in sorting.

Given the importance in our analysis of a non-linear relationship between children's and parents schooling, we think it is of interest to document these beyond the markov transition probabilities reported earlier. Table 5 below presents several regression results that incorporate higher-order terms in Kremer's original regression.[35] These regressions are based on our sample of parents and children from the PSID.

Table 5
Children's Education as a Function of Parent's Education
Dependent variable is years of education for the child.
(Standard errors are in parentheses)

	(1)	(2)	(3)
constant	8.71	12.30	16.51
	(.247)	(.608)	(1.285)
$\overline{e}$	.378	$-.347$	-1.798
	(0.021)	(0.114)	(0.407)
$\overline{e}^2$	–	0.034	0.184
		(0.005)	(0.041)
$\overline{e}^3$	–	–	-0.005
			(0.001)
N	1385	1385	1385
R^2	0.185	0.208	0.216

Column (1) in this table is the equivalent to column (5) in Table II in Kremer, with basically identical results. What columns (2) and (3) show, however, is that there is strong support for the notion that this relationship is nonlinear. In every specification, all terms are significant at the one percent level. Note that in the cubic specification the second derivative changes from positive to negative at $\overline{e}$=12.26. Hence, up to this point there are "increasing returns" to parental education in terms of "producing" child's education, but beyond this

[35] Although Kremer runs a regression that includes the square of parental average education, he also includes the square of average neighborhood education and an interactive term between parental and neighborhood effects. In that regression all variables are statistically insignificant, including average parental education.

point there are "decreasing returns" to parental education.[36] The fact that there are increasing returns in the lower part of the distribution suggests that increased sorting within this part of the distribution may actually increase mean educational attainment within this group. Our analysis abstracts from this issue since it is concerned with the degree of sorting between the top and bottom parts of the income distribution rather than the within group sorting. There we find a concave relationship between children and average parental years of education.

Our conclusion is also not sensitive to the choice of the value for the wage premium. Although the extent of income inequality in the steady state is affected by this ratio, using values for the wage premium anywhere in the range of 1.4 to 1.9 has virtually no impact on the extent to which the increase in sorting increases the steady-state standard deviation of log income.

Lastly, we consider how are results are affected by considering alternative values for the elasticity of substitution between skilled and unskilled workers. In our benchmark model we assumed a value for this elasticity equal to 1.5 (i.e., $\rho = .33$). Here we report how our conclusions are affected by assuming values of 1.0 ($\rho = 0$) and 2.0 ($\rho = .5$). While the range of estimates in the literature seems to be relatively tightly bunched around 1.5, we consider a relatively large interval for our sensitivity analysis to indicate the effect of this key parameter on our results. As above, we focus on how this change would affect the results assuming that the bindingness of constraints is unaffected (i.e., a move from the first to the second column in Table 2).

Table 6 contains the results, with the first column repeating the findings from Table 2 in order to facilitate comparisons. As the change in $\widehat{\beta}$ (and hence all changes in the distribution of education) is not affected by the value of this elasticity, we only include information on wages and inequality. As expected, the change in the standard deviation of log income is decreasing in this elasticity, but even when $\rho = 0$ the resulting change is still substantial—more than 11%. We conclude again that our results are robust to changes in this key elasticity.

[36] We have also run regressions by splitting the sample into two groups: parents with average education less than or equal to 12 years, and parents with average education greater than or equal to 12 years. These results confirmed the above finding concerning the switch in returns to scale.

Table 6
Effects of Increased Sorting for Alternative Values of ρ

	$\rho = .33$	$\rho = .5$	$\rho = 0$
$\%\Delta w_s/w_u$	9.74	7.37	14.89
$\%\Delta w_u$	-6.27	-4.70	-9.27
$\%\Delta std(log\ y)$	15.05	11.29	22.26

8. Conclusion

This paper investigated the effects of increased assortative matching in marriage. We constructed a dynamic model of education acquisition and parameterized it to US data. We conclude that large increases in sorting are likely to have quantitatively significant effects on the degree of income inequality.[37] Our conclusion is independent of the existence of imperfect borrowing markets. If borrowing constraints exist and are tightened as a result of the increase in sorting, the effects of the sorting increase on the degree of inequality are magnified.

Several factors contribute to our obtaining this conclusion. In particular, a negative correlation between fertility and education, a decreasing marginal effect of parental education on children's years of education, and a process of wage determination that is sensitive to the relative supply of skilled to unskilled workers all play a role in our qualitative and quantitative analysis.

Our model interpreted borrowing constraints as high-aptitude individuals unable to borrow to cover the cost of obtaining a college education. We do not take this interpretation literally. An alternative formulation would be to assume that a child's aptitude is determined jointly by parental educational attainment and the resources that they devote to the child's development (for example the quality of K-12 education the child obtains). If parents are unable to borrow against their child's future income to provide them with greater schooling resources, parental income is again a factor determining investment in a child's future education. This alternative interpretation does not require borrowing constraints to be operative at the time a person decides whether to attend college. Children who grow up in poor families will be less likely to attend college, not because they cannot obtain a loan to finance their college education, but because they have had lower quality K-12 educations and are less able to benefit from a college education.

[37] A recent study by Dahan and Gaviria (1999) using household survey data for several Latin American countries finds a positive relationship between sorting and inequality.

The model we constructed assumed for simplicity that both fertility and the matching process were exogenously determined. It would be of interest in future work to endogenize these variables.[38] If, as is reasonable to assume, lower family income leads to greater fertility and greater wage differentials lead to more effort to match with higher-income individuals, we conjecture that these would serve to reinforce our conclusions.[39]

[38] Becker (1973) is the classic static model of marriage. See, for example, Burdett and Coles (1997) and Cole, Mailath, and Postlewaite (1992) for models that endogenize the degree of marital sorting and Fernandez and Gali (1999) for a model that incorporates borrowing constraints into the matching process.

[39] See Fernandez (1999) for a model that endogenizes fertility and sorting.

References

[1] Banerjee, A. and A. Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101 (1993), 274-98.

[2] Becker, G., "A Theory of Marriage," *Journal of Political Economy*, 81 (1973), 813-846.

[3] Behrman, J., R. Pollak, and P. Taubman, "Family Resources, Family Size, and Access to Financing for College Education," *Journal of Political Economy*, 97(2), (1989), 398-419.

[4] Benabou, R., "Heterogeneity, Stratification and Growth", *American Economic Review,* 86(3) (1996a), 584-609.

[5] ________, "Equity and Efficiency in Human Capital Investment," *Review of Economic Studies*, 63 (1996b). 237-264.

[6] Burdett. K., and M. Coles, "Marriage and Class," *Quarterly Journal of Economics*, 112 (1) (1997), 141-68.

[7] Caucutt, B., "Education Policy When There are Peer Group Effects," mimeo, (1997).

[8] Cole, H., G. Mailath, and A. Postlewaite, "Social Norms, Savings Behavior and Growth," *Journal of Political Economy* 100 (1992), 1092-1125.

[9] Cooper, S., "A Positive Theory of Income Redistribution," *Journal of Economic Growth* 3 (1998), 171-195.

[10] Dahan, M. and A. Gaviria, "Sibling Correlations and Intergenerational Mobility in Latin America," mimeo, 1999.

[11] Dahan, M. and D. Tsiddon,"Demographic Transition, Income Distribution and Economic Growth," *Journal of Economic Growth*, 3 (1998), 29-52.

[12] Durlauf, S., "A Theory of Persistent Income Inequality," *Journal of Economic Growth* 1 (1995).

[13] Epple, D., and R. Romano, "Ends Against the Middle: Determining Public Service Provision When There are Private Alternatives," *Journal of Public Economics* 62 (1996a), 297-325.

[14] Fernandez, R. "Sorting, Inequality and Growth," work in progress, 1999.

[15] Fernandez, R and J. Gali, "To Each According to ...? Tournaments, Markets and the Matching Problem with Borrowing Constraints," *Review of Economic Studies*, 66(4) (1999), 799-824.

[16] Fernandez, R. and R. Rogerson, "Income Distribution, Communities and the Quality of Public Education," *Quarterly Journal of Economics*, 111(1) (1996), 135-164.

[17] _______, "Education Finance Reform: A Dynamic Perspective," *Journal of Policy Analysis and Management*, 16(1) (1997) 67-84.

[18] _______, "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education Finance Reform," *American Economic Review*, 88 (1998), 813-833.

[19] Galor, O. and J. Zeira, "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60 (1993), 35-52.

[20] Greenwood, J., N. Guner and J. Knowles, "More on Marriage, Fertility, and the Distribution of Income," mimeo, (1999).

[21] Knowles, J. "Marriage, Fertility and Income Distribution," mimeo, (1999).

[22] Kremer, M. "How Much Does Sorting Increase Inequality," *Quarterly Journal of Economics* CXII (1997), 115-139.

[23] _______ and D. Chen, "Income Distribution Dynamics with Endogenous Fertility," *American Economic Review Papers and Proceedings*, (1999), 155-160.

[24] _______ and E. Maskin, "Wage Inequality and Segregation by Skill," mimeo, (1996).

[25] Krusell, P., L. Ohanian, V. Rios-Rull, G. Violante, "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Federal Reserve Bank of Minnesota Staff Report* #239, (1997).

[26] Ljungqvist, L., "Economic Underdevelopment: The Case of a Missing Market for Human Capital," *Journal of Development Economics*, 40 (1993), 219-239.

[27] Loury, G., "Intergenerational Transfers and the Distribution of Earnings," *Econometrica,* 49 (1981), 843-867.

[28] Mare, R. "Five Decades of Educational Assortative Mating," *American Sociological Review*, LVI (1991), 15-32.

[29] Reich, R., "Secession of the Successful," *New York Times Magazine*, Jan. 20 (1991).

[30] Wilson, W.J., "*The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy,*" Chicago, IL.: University of Chicago Press, (1987).