

A Theory of Intergenerational Mobility*

Gary S. Becker, *University of Chicago*

Scott Duke Kominers, *Harvard University*

Kevin M. Murphy, *University of Chicago*

Jörg L. Spenkuch, *Northwestern University*

August 2015

Abstract

We develop a model of intergenerational resource transmission that emphasizes the link between cross-sectional inequality and intergenerational mobility. By drawing on first principles of human capital theory, we derive several novel results. In particular, we show that, even in a world with perfect capital markets and without differences in innate ability, wealthy parents invest, on average, more in their offspring than poorer ones. As a result, persistence of economic status is higher at the top of the income distribution than in the middle. Successive generations of the same family may even cease to regress towards the mean. Moreover, we demonstrate that government interventions intended to ameliorate inequality may in fact lower intergenerational mobility—even when they do not directly favor the rich. Lastly, we consider how mobility is affected by changes in the marketplace.

* We received helpful comments from Steven Davis, Edward Lazear, Casey Mulligan, Emmanuel Saez, as well as from seminar audiences at the University of Chicago and the Hoover Conference on Inequality in Memory of Gary Becker. Any errors, however, are our own. Kominers gratefully acknowledges the support of the National Science Foundation (grants CCF-1216095 and SES-1459912) and the Harvard Milton Fund.

1. Introduction

Across countries, inequality and intergenerational persistence are strongly positively correlated—so much so that the relationship depicted in Figure 1 is sometimes referred to as “The Great Gatsby Curve” (Krueger 2012).¹ For the United States, however, the best available empirical evidence suggests there has been little to no change in mobility (see Chetty et al. 2014a), despite a dramatic increase in inequality over the last forty years (Autor et al. 2008, Juhn et al. 1993, Katz and Murphy 1992). One is, therefore, left to wonder how cross-sectional inequality relates to social mobility and whether the intergenerational transmission of resources tends to dampen or exacerbate changes in inequality.

In this paper, we consider how persistence of economic status depends on the distribution of income. Our theory builds on standard models of intergenerational mobility (e.g., Becker and Tomes 1979, 1986, Loury 1981, Solon 2004). However, we depart from traditional approaches by introducing complementarities in the production of children’s human capital. Specifically, we dismiss the (implicit) assumption that all parents are equally good at investing in their children. Complementarities between parental human capital and investments in children directly imply nonlinearities in the production function for human capital, which in turn have far-reaching implications for intergenerational mobility in different parts of the income distribution.

We show that when the return on investments in children increases in parental human capital, then the equilibrium relationship between parents’ and children’s human capital tends to be convex. This convexity results in higher intergenerational persistence among well-to-do families. Credit constraints, on the other hand, may produce high persistence among poor families. The theory thus predicts low intergenerational mobility at both ends of the income distribution, alongside a mobile middle class.

Our model’s key prediction is borne out in United States data. Figure 2, which is based on the findings of Chetty et al. (2014b), shows that children born to parents in the top (bottom) quintile of the income distribution have a 36.5% (33.7%) probability of remaining in the top (bottom) quintile, compared with a 10.9% (7.5%) probability of transitioning to the bottom (top). By contrast, children of parents in the middle of the income distribution have an almost equal probability of ending up in *any* quintile of their generation’s distribution (see also Corak and

¹ Ironically, Jay Gatsby exhibited tremendous social mobility in F. Scott Fitzgerald’s well-known novel.

Heisz 1999, and Mazumder 2005a). In addition to rationalizing these within-country differences, our model explains why countries with very unequal distributions of human capital, such as the United States or many developing nations, have lower rates of intergenerational mobility than countries where fewer people are in the tails of the distribution.²

We also clarify why changes in the marketplace may or may not lead to lower *measured* intergenerational mobility. According to the theory, changes in inequality due to increases in the return to human capital should have virtually no effect on successive generations' *relative* positions in the income distribution, as confirmed by Chetty et al. (2014a). If, however, persistence is measured with respect to labor earnings, then rising returns to education should lead to temporary increases in persistence, followed by long-term declines. The latter result is, again, a direct consequence of complementarities in the human capital production function.

Our analysis highlights an important trade-off in the design of government policies intended to combat inequality and raise intergenerational mobility. Government programs that are complementary to parental inputs crowd *in*—rather than crowd out—investments by parents. Thus, such programs are generally more cost-effective than those that substitute for parental inputs. At the same time, programs that complement parents' investments exacerbate existing disparities in investments in children's human capital. Consequently, they can have the unintended consequence of *lowering* mobility across generations.

In stark contrast to standard analyses, our theory implies that if complementarities in the production of human capital are strong enough, then successive generations of the same family may cease to regress toward the population mean. Even modern societies may therefore exhibit different social strata, with considerable mobility within but not across class boundaries.

By emphasizing complementarities in the production of children's human capital the theory in this paper is closely related to existing models of social interactions and neighborhood effects (e.g., Bénabou 1993, Durlauf 1996a,b, among others). This literature shows how spillover effects within distinct social environments can generate persistent poverty and, thus, low intergenerational mobility (see Durlauf 2006). An important difference between our approach and the literature on poverty traps is that our theory does not rely on group-level influences and

² Indeed, Jäntti et al. (2006) observe that cross-country differences in estimated intergenerational income correlations are driven by the tails of the respective distributions.

residential sorting. Moreover, a key contribution of our paper is to explain why intergenerational persistence is higher at the *top* of the income distribution than in the middle.

The remainder of the paper is organized as follows. Building on first principles in human capital theory, Section 2 develops a model of the intergenerational transmission of resources that links cross-sectional inequality and intergenerational mobility. Section 3 discusses how changes in the marketplace (e.g., increasing returns to education and increased assortative mating), affect mobility. Section 4 explains why government interventions aimed at reducing inequality may have unintended consequences, while Section 5 considers dynasties' evolution in the long run. The last section concludes.

2. A Model of Intergenerational Mobility

This section presents a simple model of intergenerational mobility in economic status. We follow the approach of Becker and Tomes (1979, 1986) but pay especially close attention to complementarities in the production of children's human capital.

Assume two periods of life: childhood and adulthood. Each parent has one child at the beginning of adulthood, which means that parents and children overlap when the latter are young (cf. Figure 3). Adults use the human capital accumulated as children to generate labor income. These earnings can be spent on consumption, investing in the human capital of their children, and possibly leaving bequests.³

Parental preferences are assumed to depend not only on parents' consumption, c , but also on the expected utility of children. A natural formulation is

$$(1) \quad V(I_p) = u(c) + \alpha E[V(I_c)],$$

where $\alpha \in (0,1)$ is parents' degree of altruism toward their children, and I_p and I_c are the monetary resources of parents and children, respectively. For simplicity, we assume that individuals are risk averse, and that the value function V exists and is continuously differentiable.⁴

³ Here, we do not pay much attention to bequests. For an in-depth analysis of the link between human capital investments and bequests, see Becker et al. (2015).

⁴ Existence of V can, for instance, be ensured by bounding the state variable I by some possibly very large values.

We model earnings by assuming that each unit of human capital, H , yields an expected rate of return equal to r , as in

$$(2) \quad E = rH + \epsilon.$$

It is useful to think of r as the return to education, or as the “rental” rate on human capital. The level of r is determined by the stock of human capital in the economy, the stock of physical capital, technological progress, and other factors beyond the control of parents. The “error” term ϵ is distributed independently of H and has a mean of zero. It refers to all other determinants of earnings, such as good or bad luck in the search for jobs, or macroeconomic shocks that affect certain firms or entire regions and industries. Every parent takes r as given and uses equation (2) to determine the optimal investment in his child’s human capital, without knowing the realization of ϵ_c .

A general function for the production of children’s human capital is

$$(3) \quad H_c = F(y, G, A_c, H_p, v_c),$$

where H_c and H_p are the human capital of children and parents, respectively, y is parental investments in children, and G denotes government spending on education. Here, A_c stands for the abilities of children, while v_c records other influences on the formation of human capital. To clearly bring out our main results, we simplify F to be a quadratic function of only y , H_p , and a random term, v_c . Later, we explicitly analyze the effects of government spending and heterogeneity in children’s innate abilities.

A quadratic production function is general enough to include interaction effects between the different determinants of H_c yet specialized enough to produce closed-form solutions. Our main conclusions, however, hold much more generally.

Assuming that H_c is quadratic in y and H_p , we get

$$(4) \quad H_c = F(y, H_p) = \mu + \kappa y + \varphi y^2 + \theta y H_p + \delta H_p + \gamma H_p^2 + v_c.$$

Clearly, we expect that $\kappa > 0$ and $\delta > 0$, i.e. that increases in parental investments and increases in parental human capital both raise the human capital of children. Presumably, there are also

diminishing returns to y —at least eventually—as it becomes harder and harder to instill more knowledge into children with fixed mental capacity. Hence, $\varphi < 0$.⁵

The coefficient θ is especially important for our analysis. It denotes the effect of greater parental human capital on the marginal product of investments in children. Although our framework can deal with any value or sign of θ , considerable evidence suggests that parental human capital and investments in children are complements, i.e. $\theta > 0$ (see, for instance, Lareau 2011 or the studies cited by Heckman and Mosso 2014). Put differently, we assume that highly educated parents are more productive at teaching their children. This assumption reflects the basic idea that education helps parents choose more effective inputs in order to achieve the same outcome, that educated parents may be better at navigating the intricacies of public school systems, or something as trivial as knowledgeable adults being more likely to be able to help children with their schoolwork.

Although we favor the interpretation of high human capital parents being literally “better” at investing, we note that our conclusions would continue to go hold for many other, unmodeled sources of complementarity. For instance, more educated parents might enjoy a higher total return on furthering their children’s human capital because they also derive *intrinsic* utility from their offspring’s academic achievements. Alternatively, the marginal product of human capital investments may be increasing in school and neighborhood quality, or it may depend on the nature of children’s social interactions. The crux of our analysis is that children of well-educated parents are more likely to grow up in environments that act as complements to parental investments.

Human capital theory offers less guidance with respect to the coefficient γ . The marginal product of parental human capital may well be increasing, constant, or decreasing. We, therefore, refrain from signing it *a priori* and note that all of our results go through as long as γ is not too negative.

⁵ If φ were positive, the optimal y with perfect capital markets would either be zero or infinite.

A. Human Capital Transmission with Perfect Capital Markets

First, we consider the case of perfect capital markets. This allows us to simplify the analysis and focus on the basic economic forces that drive our results. We introduce credit constraints in Section 2.B.⁶

By a perfect capital market we mean that parents can borrow as much as they want at a fixed rate $R_k > 0$ and can arrange for the debt to be repaid out of the adult earnings of their children (e.g., by leaving negative bequests).⁷ In a world with perfect capital markets, all parents who care at least a little bit about their children (i.e. for whom $\alpha > 0$) invest in their offspring's human capital until the marginal return is driven down to R_k . That is, they will choose to invest the efficient amount.

Usually, this is taken to imply “perfect” intergenerational mobility, i.e. the earnings of children depend only on innate ability and not on the income of parents (see, e.g., Becker and Tomes 1986). The analysis in this section demonstrates that perfect mobility will generally *not* obtain. In sharp contrast to the results of previous work, we show that while investments continue to be efficient, intergenerational mobility in income does depend on parental human capital and will, therefore, be correlated with parental resources—even when capital markets are perfect and all children are assumed to be equally able.

Incidentally, our results cast doubt on common approaches to testing for credit constraints. As will become clear from the analysis below, a positive correlation between educational attainment and parental income, even after controlling for ability, need not imply that liquidity constraints hinder parental investments in the human capital of their children.

In the model, parents choose consumption level c , investments y , and bequests b_c in order to maximize V subject to the lifetime budget constraint

$$(5) \quad c + b_c/R_k + y = I_p \equiv E_p + b_p,$$

the production function of human capital in equation (4), and the determinants of earnings in (2).

⁶ There remains considerable controversy about the importance of liquidity constraints for educational investments. For conflicting empirical evidence, see the work of Lochner and Monge-Naranjo (2012), Lovenheim (2011), Belley and Lochner (2007), Carneiro and Heckman (2002), and the studies cited therein.

⁷ For an analysis of how human capital investments and bequests interact when the latter are restricted to be positive, see Becker et al. (2015).

Combining the first-order conditions for y and b_c , we find the usual relation determining efficient investment in children's human capital:

$$(6) \quad R_y \equiv rF_y = r(\kappa + 2\varphi y + \theta H_p) = R_k.$$

In words, equation (6) implies that when capital markets are perfect, parents invest in their children's human capital until the marginal return on these investments equals the exogenously given return on capital. Even rich parents would not find it optimal to invest any further, as saving the money and transferring it to children in the form of bequests would yield a weakly higher return. Conversely, investing less is suboptimal because borrowing at rate R_k to invest in children's human capital is always profitable as long as $R_y \geq R_k$ (and as long as the debt can be repaid out of the adult earnings of children).

According to equation (6), the return on human capital investments is falling in y because of the assumption of diminishing returns (i.e. $\varphi < 0$), and it is rising in H_p due to the complementarity between y and H_p (i.e. $\theta > 0$). The latter fact has important consequences for intergenerational mobility in different parts of the income distribution.

We use equation (6) to solve for the optimal investment in children's human capital:

$$(7) \quad y^* = \frac{R_k - \kappa - \theta H_p}{2\varphi}.$$

In equilibrium, parental investments decrease with the return on physical capital (R_k) and the degree of diminishing returns (φ), whereas they increase with the return on human capital (r) as well as parents' own human capital (H_p).

Consistent with the last prediction, college-educated parents in the U.S. not only spend more monetary resources on enhancing their children's human capital (Duncan and Murnane 2011), but also invest considerably more time than less educated parents (Ramey and Ramey 2010 and Guryan et al. 2008). Given that the former face a higher opportunity cost—their time is more valuable—these patterns are indicative of strong complementarities.

By choosing optimal investments that depend positively on parental human capital, parents affect the equilibrium mapping between their own human capital and that of their children. We can see this by using equation (8) to eliminate y from the production function for

H_c . The result is a quadratic relationship between H_c and H_p that differs greatly from the production function in equation (4):

$$(8) \quad H_c = \mu^* + \delta^* H_p + \gamma^* H_p^2 + v_c,$$

where $\mu^* \equiv \mu - \frac{\kappa^2 r^2 - R_k^2}{4\varphi r^2}$, $\delta^* \equiv \delta - \frac{\kappa\theta}{2\varphi}$, and $\gamma^* = \gamma - \frac{\theta^2}{4\varphi}$.

The coefficients δ and γ would affect the linear and quadratic terms in this relation even if parents did not optimize over y . However, optimization affects the equilibrium mapping by giving the interaction term, θ , a significant role in determining δ^* and γ^* . The greater the complementarity between y and H_p , the larger is the (nonlinear) effect of parental human capital on the human capital of children.

The equilibrium relationship between H_c and H_p would be quadratic even if the direct effect of H_p in the production of H_c were linear, i.e. if γ were 0, or if H_c exerted no direct effect at all, i.e. if δ and γ were both equal to 0. The reason is that H_p impacts the marginal return on investment. Importantly, our work shows that as long as long parental human capital does not exhibit diminishing returns that are “too large” relative to the complementarity between y and H_p (i.e. as long as $\gamma > \frac{\theta^2}{4\varphi}$), the equilibrium relationship between the human capital of parents and their children should actually be convex.⁸

Previous work typically assumed that the intergenerational transmission of human capital depends linearly on parental endowments, with an error term that is uncorrelated over time (see, e.g., Becker and Tomes 1979, 1986, Solon 2004). By taking the complementarity between parents’ own human capital and their investments in children into account, our analysis departs from this assumption. Though subtle, introducing complementarities has important consequences for intergenerational persistence in different parts of the distribution.

The substantive implication of a convex relationship between H_c and H_p is that the degree of intergenerational mobility will be greater for children of less educated parents than for those whose parents have high levels of human capital. To see this, note that $\frac{dH_c}{dH_p}$ measures the degree of intergenerational persistence in human capital, i.e. by how much children’s human

⁸ Recall that $\varphi < 0$.

capital covaries with that of their parents. Equation (8) implies that $\frac{dH_c}{dH_p}$ increases in H_p and that it increases faster when θ is large. Hence, intergenerational persistence of human capital rises with parents' own human capital, and the difference between rich and poor families will be more pronounced the more important the complementarity between y and H_p . Thus, whenever labor earnings are individuals' primary source of income, more persistence in human capital among the children of well-educated parents implies less intergenerational mobility in the upper parts of the income distribution than in the middle.

B. Human Capital Transmission in the Presence of Credit Constraints

We now introduce credit constraints into our model and analyze how they impact parental investments in the human capital of children.

We assume that parents' access to credit depends on their earnings: richer parents may be able to borrow more cheaply than poorer ones, while the richest parents have a constant marginal cost of funds. More formally, we assume that the interest rate as a function of earnings, $R(E_p)$, satisfies

$$(9) \quad R'(E_p) \leq 0 \quad \text{and}$$

$$(10) \quad R''(E_p) \geq 0,$$

with $R'(E_p) = 0$ and $R(E_p) = R_k$ for E_p large enough. This formulation is quite flexible. In particular, it allows for the possibility that poor parents may only be able to borrow at interest rates that are prohibitively high. In symbols, we assume that $R(E_p) \rightarrow \infty$ as $E_p \rightarrow 0$.

By modelling liquidity constraints in this straightforward fashion, we can apply most of the intuition developed above. Under the condition that constrained parents spend anything at all on their offspring, i.e. that $y^* > 0$, they will transmit resources by investing in their children's human capital until the marginal return on those funds equals the interest rate. At any interior solution, we therefore have

$$(11) \quad R_y \equiv rF_y = r(\kappa + 2\varphi y^* + \theta H_p) = R(E_p).$$

As in the case of perfect capital markets, parents would not find it optimal to invest more resources in children, since the return on additional spending would be below the cost of funds. Conversely, parents would not want to invest less in their children's human capital. Doing so would reduce the total resources available to the children and parents, which is the relevant criterion as long as adult children can commit to paying back their parents or if educational loans can be repaid out of the adult earnings of children.

In other work we analyze parental investments in the human capital of children when the latter cannot commit to paying back the former (see Becker et al. 2015). Among other things, we show that if children are altruistic towards their parents (perhaps because parents manipulated them to be altruistic), then parents will continue to invest in the human capital of their children until the marginal return on those investments is driven down to the cost of funds. The solution to parents' optimization problem is, therefore, the same as above.

Comparing equations (6) and (11), we see that the main difference between the case with credit constraints and that of perfect capital markets is that $R(E_p) \geq R_k$. In particular, optimal investments by credit-constrained parents continue to be given by

$$(12) \quad y^* = \frac{\frac{R(E_p)}{r} - \kappa - \theta H_p}{2\varphi}.$$

In contrast to the case of perfect capital markets, however, investments in children are inefficiently low. From a social planner's perspective, the opportunity cost of capital equals R_k , which is strictly lower than the return on investing in the human capital of children whose parents are credit constrained. A social planner would, therefore, spend more on these children.

Differentiating equation (12) with respect to H_p shows that

$$(13) \quad \frac{dy^*}{dH_p} = \frac{\frac{1}{r} \frac{dR}{dE_p} \frac{dE_p}{dH_p} - \theta}{2\varphi} = \frac{\frac{dR}{dE_p} - \theta}{2\varphi} > 0.$$

Hence, optimal parental investments in the human capital of children increase in parents' own human capital. Since $(\frac{dR}{dE_p} - \theta)/2\varphi \geq -\theta/2\varphi$, investments increase faster when parents are credit constrained. Intuitively, this is because higher parental human capital not only makes investments in children more productive, but it also lowers the investments' marginal cost.

Using equation (12) to substitute for y in equation (4) and differentiating with respect to H_p gives

$$(14) \quad \frac{dH_c}{dH_p} = \frac{R(E_p)}{2\varphi r} \frac{dR}{dE_p} + \delta^* + 2\gamma^* H_p \quad \text{and}$$

$$(15) \quad \frac{d^2 H_c}{dH_p^2} = \frac{1}{2\varphi} \left[\left(\frac{dR}{dE_p} \right)^2 + R(E_p) \frac{d^2 R}{dE_p^2} \right] + 2\gamma^*.$$

Under the assumptions above, it follows that $\frac{dH_c}{dH_p} > 0$, i.e. that children's human capital is increasing in that of their parents. The sign of $\frac{d^2 H_c}{dH_p^2}$, however, depends on parental income. For parents who are rich enough such that $R(E_p) \approx R_k$, the first summand in equation (15) will be very close to 0. Since $\varphi < 0$, it must be the case that $\frac{d^2 H_c}{dH_p^2} > 0$ for H_p (and, therefore, E_p) large enough. By contrast, for sufficiently poor parents, i.e. for parents for whom $\frac{dR}{dE_p}$ is negative and very large, we have that $\frac{d^2 H_c}{dH_p^2} < 0$. This is because credit constraints lower investments, but their impact reduces more and more as parents' income reaches the level at which $R(E_p) = R_k$. Consequently, the human capital production function will initially be concave in parental human capital and then convex, as in Figure 4.

Thus, we see that the previous result about increased persistence in human capital among children of highly educated parents carries over to a world in which credit constraints depend on parental income. At the same time, relative to a world with perfect capital markets, credit constraints imply more persistence among children of the poor. Our model, therefore, suggests high persistence at both ends of the distribution and more churning in the middle.

As mentioned in the introduction, the predictions of our model are borne out in several careful empirical studies. Corak and Heisz (1999) and Mazumder (2005a), for instance, show that conditional on being raised by parents in the middle of the income distribution, there is almost no relationship between family background and child outcomes. This is true in both the U.S. and in Canada. For children of parents at the extremes of the distribution, however, parental income *is* highly predictive of economic status as adults—with the relationship being stronger in the U.S. Using administrative records on almost 10 million U.S. children born between 1980 and 1982, Chetty et al. (2014b) confirm these results. In their data, children born to parents in the top

(bottom) quintile of the income distribution have a 36.5% (33.7%) probability of remaining in the top (bottom) quintile, compared with a 10.9% (7.5%) probability of transitioning to the bottom (top). By contrast, children of parents in the middle of the income distribution have an almost equal probability of ending up in *any* quintile of their generation's distribution.

Since educational attainment tends to be more equally distributed in Nordic countries than in the U.S. and most of the developing world, our theory is able to explain why socioeconomic inequality and intergenerational persistence tend to be correlated across nations (cf. Figure 1). According to our model, the root cause of low average rates of social mobility lies in an unequal distribution of human capital. This prediction is supported by the findings of Jäntti et al. (2006). Comparing the United States, the United Kingdom, Denmark, Finland, Norway, and Sweden, Jäntti et al. (2006) observe that cross-country differences in estimates of the intergenerational elasticity of income are almost entirely due the tails of the respective distributions. Surprisingly, our model predicts that differences in how well human capital is compensated, i.e. the education premium, matter little. In fact, as we show in Section 3, increases in r may have no, or even positive, effects on mobility.

Our model also allows us to speak to the recent work of Clark (2014). Using rare surnames to identify members of a society's elite (or underclass), Clark (2014) shows that full social convergence often takes centuries rather than a few generations. The convergence rates implied by his estimates are, therefore, significantly lower than those reported in the previous literature. He rationalizes these findings by arguing that fluctuations in realized incomes across generations attenuate traditional measures of social persistence. By contrast, Chetty et al. (2014a,b) report that in their data surname-based estimates of mobility are quite similar to those based on income. One notable difference between the analyses of Clark (2014) and Chetty et al. (2014a,b) is that the former explicitly focuses on dynasties who once belonged to an elite or underclass. According to our theory, rates of persistence measured in the tails of the distribution ought to be higher than those in a society as a whole.

C. Children's Abilities

So far, we have posited that all children are equally able (*ex ante*), thereby neglecting the potentially important interaction between ability and human capital. One useful way to think of

our analysis is to assume that the market does not directly reward ability, but only human capital. That is, we view human capital as a “sufficient statistic” for individuals’ productivity, irrespective of how it was acquired.

Of course, holding fixed parental investments, parental human capital, and other environmental variables, it is natural to think that higher-ability children will also end up having more human capital and, therefore, be more productive. To incorporate ability into our analysis, we assume that parents know their children’s ability (A_c) and take it as given when deciding on their investments.⁹ It is also plausible to assume that the productivity of parental investments is higher when children are more able. This may, for instance, be the case because cognitive ability enables children to pay more attention, or because it lets parents use more efficient methods to convey the same lessons.

If A_c and y are, in fact, complements, then the equation determining the optimal level of parental investments becomes

$$(16) \quad y^* = \frac{\frac{R(E_p)}{r} - \kappa - \theta H_p - \omega A_c}{2\varphi},$$

where $\omega > 0$ denotes the coefficient on the interaction term between y and A_c that now enters the production function for H_c :

$$(17) \quad H_c = F(y, H_p) = \mu + \kappa y + \varphi y^2 + \theta y H_p + \delta H_p + \gamma H_p^2 + \rho A_c + \omega y A_c + v_c.$$

Equation (16) shows that the optimal value of y depends positively on children’s abilities as well as parental human capital. It also implies that the equilibrium relation for H_c , given by equation (8), is not only a function of H_p and its square, but also of A_c , A_c^2 , and the interaction between A_c and H_p . In particular, holding parental ability fixed and assuming that the marginal impact of ability in the production of human capital does not diminish “too fast,” children’s human capital is a convex function of ability.

Nonlinearities in the production function for human capital have important implications for empirical work on intergenerational persistence. For instance, the existing literature

⁹ At first, this assumption may seem at odds with existing evidence showing that cognitive and noncognitive skills are malleable, especially during the early childhood (see, e.g., the review by Heckman and Mosso 2014). This is not necessarily so. In our language, cognitive and noncognitive skills would be classified as human capital, while the term ability is reserved for traits that are truly innate.

sometimes puzzles about why grandparents' educational attainment seems to positively "affect" children, even after controlling for the schooling of parents (see, e.g., Behrman and Taubman 1985, Lindahl et al. 2015, and the studies cited in Solon 2015).¹⁰ Typically, explanations assign grandparents an important role in raising children or appeal to cultural or genetic factors (Zeng and Xie 2014, Solon 2013). By contrast, our theory rationalizes these findings without relying on a direct effect of previous generations. Since the human capital of parents and grandparents is almost certainly positively correlated, it may be that the positive "effect" of grandparents' years of schooling is an artifact of controlling only for H_p , but not H_p^2 .

Our analysis also implies that "reduced form" approaches to inferring the correlation between years of children's education and that of their parents produce upward-biased estimates of the true impact of parental human capital. Since ability and schooling are positively correlated, and since educated parents are more likely to have especially able children, such estimates will, in part, reflect the effect of own ability on human capital accumulation. More surprisingly, the bias may persist even after linearly controlling for IQ and other measures of children's skills. Again, the reason is that the actual relationship between human capital and ability is likely nonlinear and that it may include an interaction between child ability and parental human capital. Unfortunately, the sign and magnitude of the resulting bias is difficult to determine a priori, as it depends on the *residual* covariances in H_p , H_p^2 , A_c^2 and the interaction between A_c and H_p (i.e. after partialling out all other controls).

To be clear, it is possible to consistently estimate the raw correlation between children's and parents' years of education, or other measures of economic status. The point is that such a parameter is difficult to interpret causally, as it confounds persistence due to biology, i.e. inheritability of intelligence or other innate skills, with persistence due to economic factors, such as credit constraints or cronyism. In fact, increases in intergenerational correlations may be the result of a society becoming *more*, rather than less, meritocratic.

To convey this point in as simple an example as possible, imagine two worlds: In the first one, children's educational attainment is determined by their parents' human capital and pure luck. In the second world, the impact of parents' education is exactly as large as in the first one,

¹⁰ According to the model of Becker and Tomes (1979), conditional on variables in the parents' generation, the earnings of grandparents and children should be *negatively* correlated.

but children's innate ability reduces the role of luck. According to most standards, one would think of the latter world as more meritocratic. Nonetheless, as long as ability and parental human capital are positively correlated, the second world will feature higher intergenerational correlation coefficients than the first.¹¹

D. Human Capital and Intergenerational Mobility in Earnings

Our analysis shows that parents have a major influence on the human capital of their children, especially in the tails of the distribution. Yet the human capital of children gets transformed into earnings by market forces that are largely beyond parental control. Although parents take account of the labor market when deciding on their investments in the human capital of children, the family loses some (but by no means all) of its influence in the transition from human capital to earnings.

To clearly bring out the influence that the family does have on the earnings of children, we return to the production function defined in (4) (i.e. the one without an ability term) and combine equations (2) and (8) to obtain

$$(18) \quad E_c = r_c \mu^* + r_c \delta^* H_p + r_c \gamma^* H_p^2 + r_c v_c + \epsilon_c,$$

where subscripts continue to indicate the respective generation. Aside from r_c , the rental rate of human capital in the children's generation, the coefficients in equation (18) are all determined by parameters in the production function for H_c and by the way these parameters affect parental investments in children through equation (7). By using equation (2) to substitute for H_p , the above relationship can be transformed into an equation that describes the intergenerational transmission of earnings:

$$(19) \quad E_c = \mu^{**} + \delta^{**} E_p + \gamma^{**} E_p^2 + \epsilon_c^{**},$$

¹¹ Formally, let $H_c = \beta_1 H_p + v_{c,1}$ and $H_c = \beta_2 H_p + \phi A_c + v_{c,2}$ describe the intergenerational transmission of human capital in the first and second world, respectively, with $\beta_1 = \beta_2$ and $v_{c,1}$ as well as $v_{c,2}$ being independent of H_p and A_c . Since A_c is unobserved and $\text{Cov}(H_p, A_c) > 0$, it follows that $\text{plim } \hat{\beta}_1^{OLS} < \text{plim } \hat{\beta}_2^{OLS}$. That is, there would be less measured persistence in the first world, despite the fact that the true effect of parental human capital is actually identical. Even if ability also reduced the impact of parental human capital in the second world, there would continue to be more "persistence" as long as $\beta_1 - \beta_2 < \phi \frac{\text{Cov}(H_p, A_c)}{\text{Var}(H_p)}$. This condition is more likely to hold the larger the impact of ability on human capital (ϕ) and the higher the correlation between parental human capital and children's ability.

where $\mu^{**} \equiv r_c \mu^* + \left(\gamma^* \frac{r_c}{r_p^2} \epsilon_p - \delta^* \frac{r_c}{r_p} \right) \epsilon_p$, $\delta^{**} \equiv \delta^* \frac{r_c}{r_p} - 2\gamma^* \frac{r_c}{r_p^2} \epsilon_p$, $\gamma^{**} \equiv \gamma^* \frac{r_c}{r_p^2}$, and $\epsilon^{**} \equiv r_c \nu_c + \epsilon_c$.

Unsurprisingly, equations (8) and (19) are virtual mirror images of one another, and our conclusions about the shape of the former carry over to the latter. In particular, the degree of convexity, which measures the increase in persistence among families in the right tail of the distribution, depends crucially on the interaction between parental human capital and their investments in children, i.e. on θ , which enters through γ^{**} .

Equations (18) and (19) highlight another important feature of the intergenerational transmission of income. In the absence of credit constraints, children whose parents have equally high human capital will, on average, end up with equal earnings. Put differently, increases in parental resources due to market luck (ϵ_p) or capital windfalls (b_p) do not result in higher expected earnings of children. Formally, $\frac{dE_c}{d\epsilon_p} = \frac{dE_c}{db_p} = 0$.¹²

The opposite conclusion applies to families who *are* credit constrained. In the presence of credit constraints, market luck does translate into higher expected earnings of the next generation. This effect operates through an increase in H_c caused by a reduction in the cost of borrowing, i.e. $\frac{dR(E_p)}{d\epsilon_p} < 0$. The net result is an increase in earnings equal to $\frac{dE_c}{d\epsilon_p} = r_c \frac{d\mu^*}{d\epsilon_p} > 0$. Consequently, a one-time shock to the earnings of liquidity-constrained parents raises the expected incomes of all subsequent generations until those generations' earnings are high enough such that $R(E) = R_k$, i.e. until they are no longer constrained.

3. How Changes in the Marketplace Affect Intergenerational Mobility

A. Rising Returns to Human Capital

For most of the previous section, our analysis has assumed that families take all macroeconomic parameters as given and that these parameters are constant. Although analytically convenient, the latter assumption is clearly false. The return to education and other human capital increased dramatically in the decades after 1980, especially in the United States (see, e.g., Murphy and Katz 1992, Juhn et al. 1993). In what follows, we use equations (2), (8), and (19) to study how

¹² Note that the coefficients in equation (19) depend on ϵ_p .

changes in the return to human capital impact both cross-sectional inequality and intergenerational mobility.

First, we consider how a sudden change in the return to human capital affects inequality. Holding human capital endowments fixed, equation (2) shows that an increase in r leads to a higher variance in earnings and, therefore, to more cross-sectional dispersion of incomes.¹³ The data for the United States show that inequality has, indeed, increased dramatically during the last decades, especially at the very top of the distribution (Piketty and Saez 2003). Interestingly, labor earnings and increases in the market return to talent seem to be the main drivers of inequality, even among the top 1% (Piketty and Saez 2003, Kaplan and Rauh 2013).

Based on the observation that inequality and intergenerational mobility are strongly negatively correlated across countries, it is often claimed that higher returns to human capital will not only increase cross-sectional inequality but also reduce the degree of intergenerational mobility (see, e.g., Corak 2013, Krueger 2012, Council of Economic Advisors 2012, Solon 2004). Our analysis indicates that such conjectures are *not* true in general. To show this, we consider how the intergenerational transmission of human capital depends on its rental price.

Although the degree of persistence in human capital $\frac{dH_c}{dH_p}$ is higher at the upper end of the income distribution, it turns out to be independent of r . This is because r enters equation (8) only through the constant term, μ^* , and not through the coefficients δ^* and γ^* . Thus, a rise in the returns to education will lead to an increase in children's human capital (which, if some families are credit constrained, need not be the same for all children) but will not affect how much children benefit from an additional unit of parental human capital. Consequently, intergenerational mobility in human capital remains unaffected.

While a quadratic function is a second-order approximation to any more general production function, in higher-order approximations the intergenerational transmission of human capital may very well depend on r . The important implication of the neutrality result in the quadratic case is that in more general formulations the effect of increases in inequality due to higher returns to human capital may go in either direction and is likely small.

¹³ Formally, $Var(E) = r^2 Var(H) + \sigma_e^2$, which is increasing in r .

By contrast, equation (19) indicates that persistence in earnings *does* depend on the rental rate. In the short run, that is, holding the return to human capital in the parents' generation (r_p) fixed, a rise in the return to education for children (r_c) increases the coefficients on parental earnings (E_p and E_p^2) and thereby leads to more persistence in incomes. In the long run, however, i.e. when $r_c = r_p = r$, increasing returns to human capital *lower* γ^{**} , while leaving the average δ^{**} unaffected. Thus, the relationship between the incomes of parents and the incomes of children becomes less convex, and intergenerational income mobility actually *increases*—especially among well-to-do families.

This surprising conclusion is a direct result of convexity in the intergenerational transmission of human capital, i.e. equation (8). To see this, suppose that the equilibrium relation between children's and parents' human capital were, in fact, linear, of the form $H_c = \tilde{\mu}^* + \tilde{\delta}^* H_p + \nu_c$. Then, children's income as a function of parents' income would be given by $E_c = \tilde{\mu}^{**} + \frac{r_c \tilde{\delta}^*}{r_p} E_p + \tilde{\epsilon}$, where $\tilde{\mu}^{**} \equiv r_c \tilde{\mu}^* - \frac{r_c \tilde{\delta}^*}{r_p} \epsilon_p$ and $\tilde{\epsilon} \equiv r_c \nu_c + \epsilon_c$. Consequently, increases in the returns to human capital that affect both generations equally would have no bearing on mobility. The reason is that there are two opposing effects. First, holding H_c fixed, a rise in r_c results in a proportionate increase in children's earnings. This magnifies the consequences of existing disparities in children's human capital and is why intergenerational income mobility declines in the short run. Second, comparing adults with a fixed difference in incomes, a higher r_p implies a proportional narrowing of the gap in their human capital. Hence, the human capital of children born to families with different incomes will be more similar when r_p is higher, which in turn lessens the importance of disparities in parental income. If the intergenerational transmission of human capital is linear, the first and the second effects exactly offset each other. With a convex transmission function, however, the latter effect outweighs the former.

In sum, our analysis suggests that the impact of an increase in the returns to schooling depends on how intergenerational mobility is measured. If mobility is defined as the intergenerational persistence of human capital, then an increase in the returns to schooling should have little to no effect. If, however, mobility is defined in terms of persistence of incomes, then changes in the rental rate of human capital do have measurable effects. In the short run, one would expect a one-generation decline in mobility in response to a permanent rise in the returns to schooling. In the long run, however, earnings mobility should increase whenever

complementarities in the production of children's human capital are strong enough to make the transmission function convex.

Reassuringly, recent empirical evidence for the U.S. is broadly consistent with the predictions of our theory. In a careful study based on a large sample of administrative records for the 1971-1993 birth cohorts, Chetty et al. (2014a) show that the copula of parent and child incomes, i.e. children's percentile ranking in the income distribution as a function of the rank of their parents, has remained approximately constant. Given that the theory predicts the degree of intergenerational persistence in human capital to be independent of r , and that it is human capital that determines children's and parent's *relative* place in each generation's income distribution, this is exactly what one would expect.

Somewhat more speculatively, taking the Chetty et al. (2014a) estimates of the intergenerational elasticity of income at face value suggests a modest decrease in mobility among cohorts born in the 1970s, followed by a small increase for children born after 1980 (see Chetty et al. 2014a, Appendix Table 1). If one believes that the parents of children born in the 1970s did not fully capture the rising returns to education, then our theory predicts an initial decrease in intergenerational income mobility, followed by an increase for generations whose parents would have already benefitted from a higher r .

Also consistent with the predictions of our theory, lagging the rising returns to education in the aftermath of the Civil War by about one generation, intergenerational persistence increased substantially at the beginning of the twentieth century, only to decline between 1920 and 1940 (see Olivetti and Paserman 2015).

B. Changes in Assortative Mating and Family Structure

The last decades have not only witnessed a substantial rise in the returns to human capital but have also seen important changes in the structure of many families. Figure 5 shows that the share of children growing up in single-parent households has increased from less than 10% in 1950 to about 33% in 2010. At the same time, parents in intact families have become more alike in terms of their educational achievement. That is, assortative mating has been rising (see Figure 6). Greenwood et al. (2014) provide further evidence of an increase in assortative mating, and

demonstrate by calibration how such shifts can lead to higher cross-sectional inequality. Below, we discuss the impact of changes in assortative mating on intergenerational mobility.

One channel through which assortative mating may affect intergenerational mobility is through the inheritability of intelligence and other innate traits. Suppose, for instance, that children's ability is a convex combination of that of their parents coupled with some error term, as in $A_c = \tau A_f + (1 - \tau)A_m + \zeta$, with $\tau \in (0,1)$. Increases in assortative mating, i.e. A_f and A_m becoming more similar, would then reduce the degree to which successive generations regress toward the population mean. As earnings depend on children's human capital, and since human capital is a function of ability, more assortative mating would ultimately reduce intergenerational mobility in these variables. The increase in persistence would be greater the more meritocratic a society, i.e. the larger the importance of own ability for human capital and income, but also the stronger the complementarity between ability and parental investments in the production of human capital, i.e. the larger ω in equation (16). The latter is due to the fact that ability not only has a direct effect on children's human capital but also an indirect effect through increasing parental investments. Both work to lower mobility.

A closely related way in which changes in the marriage market may affect intergenerational mobility is through parental human capital, even when holding ability fixed. Although equation (4) does not distinguish between mothers' and fathers' human capital, it is plausible that both partner's H , or some combination thereof, enter the production function for H_c , as in the following CES formulation: $H_p = (\tau H_f^\rho + (1 - \tau)H_m^\rho)^{1/\rho}$. Children of two highly educated parents would, therefore, end up with higher levels of human capital than children of lesser educated couples or children from single-parent homes. Given the complementarity between parental investments and parental human capital, this would follow even if capital markets were perfect. Thus, the aforementioned changes in the marriage market do not only raise contemporaneous inequality across families, but are likely to also reduce social mobility.

This conclusion would continue to hold even if only the human capital of the caregiver, traditionally the wife, entered the production function in equation (4). As the wives of well-to-do men become more educated relative to those of less well-off men, complementarities in the production function cause average parental investments among the former, richer families to diverge from those among the latter, poorer ones.

This is, indeed, what the data show. Figure 7 indicates that since the early 1970s, enrichment expenditures per child grew by about 57% among parents in the bottom quintile of the U.S. income distribution, whereas expenditures grew by more than 150% among top-quintile parents. Ramey and Ramey (2010) and Guryan et al. (2008) report a similar trend for parental time spent with children. Going as far back as 1965, more educated parents have always spent more time on childcare than less educated ones—despite the substitution effect induced by higher wages working in the opposite direction, but consistent with strong complementarities in the production of children’s human capital—and the gap has *widened* since the early 1990s. Today, college-educated mothers spend about 16.5 hours per week on childcare, compared with 12.6 for mothers who have only graduated from high school (Guryan et al. 2008).¹⁴

While the patterns in the data are generally consistent with large impacts of an increase in assortative mating, they need not be caused by changes in assortative mating alone. As mentioned above, the U.S. also experienced a large rise in divorce rates and in the prevalence of single motherhood (Kennedy and Ruggles 2014, Ellwood and Jencks 2004). Since single-parent households are more likely to be poor and, therefore, credit constrained, the divergence in parental monetary investments may, in fact, be due to poor parents’ inability to ramp up investments in their children following the observed rise in the return to human capital. A related explanation would be that single parents are unable to invest as much as two-parent families not because the former are liquidity constrained, but because of a greater need to engage in market work, which acts as a time constraint (Becker 1965).

Whatever the true mechanism for the divergence in parental investments in children’s human capital may be, our theory suggests that it will increase inequality and social persistence for generations to come. Because of the complementarity between parental human capital and investments in children, increases in parental investments among well-to-do families translate into more human capital of the respective children; the higher-human capital children will, in turn, invest more in their own offspring, and so on.

¹⁴ Ramey and Ramey (2010) argue that the widening of the gap is due to increased competition in college admissions, but not everyone agrees (see, for instance, Hurst 2010).

4. Government Interventions

We have shown that countries with a more equal distribution of human capital tend to have higher rates of intergenerational mobility. Yet, we have so far neglected the possibly important role of governments in equalizing the human capital of children. Here, we analyze how government spending, G , is related to y and H_p , and how it affects parental investments. Our positive analysis considers two special cases that together cover a broad range of G 's effects on H_c .

A. Government Spending that Substitutes for Parental Investments

First, we suppose that government spending is a perfect substitute for parental investments, i.e. for y . This may be the case for many types of government spending, such as primary and secondary education, or the provision of textbooks, computers, and school lunches. If G were a perfect substitute for y , the production function in equation (4) would be modified to

$$(20) \quad H_c = \mu + \kappa(y + G) + \varphi(y + G)^2 + \theta(y + G)H_p + \delta H_p + \gamma H_p^2 + v_c.$$

We assume that parents choose y taking G as given.¹⁵

If capital markets are perfect, then parents' optimal level of investment equals

$$(21) \quad y^* = \frac{R_k - \kappa - \theta H_p}{2\varphi} - G.$$

Equation (21) implies that $\frac{dy^*}{dG} = -1$. That is, higher government spending crowds out private investments dollar for dollar. Only if G were large enough that y^* were reduced to exactly zero would additional government spending raise total investments.

From a distributional perspective, government interventions large enough to increase total spending on the children of poor families might very well be desirable. If capital markets were perfect, however, such spending would be excessive from an efficiency perspective. Under perfect capital markets, even families with low levels of parental human capital invest the efficient amount in their children. That is, they invest until the marginal return on y equals R_k .

¹⁵ This is a good assumption as long as government spending does not depend on each family's choice, say because there are many heterogeneous families or because government interventions are fixed over long periods of time.

Thus, government spending that more than replaces private investments would generate a marginal return lower than the market return on capital.

Naturally, the preceding conclusion is sensitive to the assumption of perfect capital markets. As shown in Section 2.B, parents who are credit constrained will generally not invest optimally in their children's human capital. As less-educated parents invest fewer resources than more-educated ones (because their spending is less productive), government spending is more likely to raise total investments for children of parents with very low human capital (and, therefore, low earnings). Consequently, government interventions that are designed to raise the total amount invested in the children of poor families may, in fact, reduce intergenerational persistence *and* increase overall efficiency.

B. Government Spending that Complements Parental Investments

The implications are very different for government interventions that are complementary with private investments, such as certain forms of early childhood education or public funding for elite universities. We analyze the case in which y and G are complements by introducing an additional interaction term into equation (4). The extended production function for human capital then equals

$$(22) \quad H_c = \mu + \kappa y + \varphi y^2 + \theta y H_p + \delta H_p + \gamma H_p^2 + \lambda G + \pi y G + v_c,$$

where λ and π are both assumed to be positive.

When capital markets are perfect and parents take G as fixed, the optimal choice of y is given by

$$(23) \quad y^* = \frac{\frac{R_k}{r} - \kappa - \theta H_p - \pi G}{2\varphi}.$$

Clearly, when G and y are complements, an increase in G raises rather than lowers optimal spending by parents. This has the surprising, but important, implication that higher government spending *increases* the intergenerational transmission of human capital, even when it is the same for all families. Government spending that complements parental investments, therefore, *decreases* intergenerational mobility.

The reason is that higher government spending increases parental investments, which in turn raises the effect of parental human capital on that of children. We can see this by substituting out y^* in the production function for H_c . Doing so gives

$$(24) \quad H_c = \mu^* + \delta^* H_p + \gamma^* H_p^2 + \lambda^* G + \psi^* G^2 + o^* HG + v_c,$$

$$\text{with } \mu^* \equiv \mu - \frac{\kappa^2 r^2 - R_k^2}{4\varphi r^2}, \lambda^* \equiv \lambda - \frac{\kappa\pi}{2\varphi}, \psi^* \equiv \frac{-\pi^2}{4\varphi}, o^* \equiv \frac{-\theta\pi}{4\varphi}, \delta^* \equiv \delta - \frac{\theta\kappa}{2\varphi}, \text{ and } \gamma^* \equiv \gamma - \frac{\theta^2}{4\varphi}.$$

Since $o^* > 0$, more government spending raises the marginal product of parental human capital, which, in turn, increases persistence in the transmission of human capital from parents to children. This effect is larger the more important the complementarities between y and G , i.e. the larger π . Moreover, the impact of government spending on intergenerational mobility is greater when parental human capital makes investments in children more productive, i.e. when θ is large. Only if y and H_p were, in fact, substitutes (i.e. if $\theta < 0$) would government spending that crowds in parental investments lead to a reduction in intergenerational persistence.

Equation (24) also shows that government spending that appears to be “neutral” because it is the same for all children may not really be neutral. For when government spending and parental investments are complements, government spending will implicitly favor families with *more* human capital—simply because G ends up increasing the productivity of H_p .

Although we have assumed that G^2 does not enter the production function for H_c , it does affect the equilibrium rate of transmission of human capital to children. Intuitively, the reason is that optimal parental investments are positively related to G , which produces a convexity in the equilibrium relationship between government spending and children’s human capital.

When G and y are complements, parental investment in children’s human capital will also raise the marginal product of government funds. Thus, governments that are primarily concerned about efficiency would bias their spending toward richer families. That is, instead of G being the same for all families, governments would find it optimal to spend more on families with higher H_p .

Of course, many government interventions—especially in developed countries—are undertaken for the express purpose of lowering inequality within and across generations. When public spending and private investments are complements, governments would have to invest a lot more in the children of poorer families than in those of richer ones in order reduce inequality.

Such targeted interventions entail a second-order welfare loss, even if G can be financed by lump sum taxation. An important question then becomes whether this loss is outweighed by potential gains from correcting the suboptimally low investments of credit-constrained parents.

When government and parental spending are substitutes, government spending need not be skewed toward the poor in order to reduce inequality. From a political point of view, implementing policies that are perceived as “neutral” may be especially attractive. The flipside, however, is that public investments that (partially) substitute for private ones will be more expensive to deliver. That is, compared with investments that are complementary to private spending, G would have to be a lot higher to raise H_c by the same amount.

In this section, we have considered the two polar cases in which government spending on children’s human capital is either a perfect substitute or a complement to parental investments. Our analysis, therefore, emphasizes the trade-offs associated with public financing of human capital investments. Actual government spending likely has both substitutable and complementary components. The general lesson is that the details of how that spending is delivered matter a great deal, even when we ignore the distortionary effects of taxation. Some government programs are likely more efficient than others, but unless they are explicitly targeted toward the poor, they may come at the cost of increasing inequality and reducing intergenerational mobility.

5. Intergenerational Dynamics and the Long-Run Evolution of Dynasties

So far, our analysis has focused on a snapshot of two generations—parents and their children—but the model we have developed also has implications for dynasties’ evolution in the long run. To clearly bring out the contribution of our theory relative to traditional models of intergenerational mobility, we first discuss the case in which children’s status depends linearly on that of their parents, as in Becker and Tomes (1979, 1986) and Solon (2004).

When the transmission of human capital is governed by $H_c = \mu^* + \delta^* H_p + v_c$, the degree of intergenerational mobility is entirely determined by the value of δ^* . For $\delta^* > 1$, successive generations of the same family accumulate ever more human capital, and initial differences between families magnify over time. The empirically relevant case, however, is the one in which $\delta^* < 1$. In fact, virtually all empirical work finds rates of intergenerational transmission below

unity (see, e.g., Clark 2014, Mazumder 2005b, Solon 1992, and the studies cited in Solon 1999 and Black and Devereux 2011). With $\delta^* < 1$, the distribution of human capital in a society is stable, and *all* dynasties regress toward the population mean.

Figure 8 illustrates this prediction by depicting dynasties' expected path over time, i.e. ignoring any short-run fluctuations introduced through ν_c and ϵ_c . Although families' fortunes may diverge over short and medium time horizons due to successive realizations of good and bad luck, in the long run the influence of the first generation vanishes completely.

In what follows we provide an important qualification of this result. In particular, we show that whether different dynasties regress towards a common mean depends critically on how important complementarities in the production of children's human capital are.

The left panel in Figure 9 depicts the case most similar to traditional analyses. Here, parental human capital raises the productivity of investments in children's human capital, but not enough to introduce strong convexities in the intergenerational transmission function. As a result, the mapping from H_p to H_c lies initially above the 45-degree line and then below. As long as the error terms in equations (2) and (4) are sufficiently bounded, and as long as H_c remains below the 45-degree line for all relevant values of H_p , there would exist a stable distribution of human capital and successive generations would regress towards the "steady state" at H^* (which need not coincide with the average level of human capital in the population).

When complementarities in the production of children's human capital are very strong, however, then γ^* in equation (8) might be large enough for the intergenerational transmission function to cross the 45-degree line a second time, as in the middle panel of Figure 9. On average, dynasties whose human capital lies below the second intersection will regress toward H^* . By contrast, dynasties that start out above \tilde{H} will gravitate *away* from the "mean." This produces bifurcation and rules out a stable distribution of human capital. To see the latter point note that as long as the distribution of ν_c is sufficiently unrestricted, a nonzero mass of families would eventually experience enough strings of "good luck" to end up in the region where their human capital grows without bound.

Since the prediction of unbounded growth is driven entirely by the assumption that the production for H_c is quadratic, and since ever-accelerating growth in dynasties' human capital is

clearly unrealistic, it is reasonable to add a small, negative cubic term to the intergenerational transmission function. As shown in the right panel of Figure 9, this cubic would only become relevant for very high levels of parental human capital, and it would reduce the marginal impact of H_p on H_c . Thus, instead of unbounded growth above \tilde{H} , the transmission function would intersect the 45-degree from above, resulting in another “steady state” at H^{**} .

Under some mild restrictions on the distribution of the error terms, there would exist a stable distribution of human capital (which, except for special cases, turns out to be extremely difficult to characterize analytically). Families with parental human capital below \tilde{H} would, on average, move toward H^* , whereas dynasties that start above \tilde{H} can expect to transition to H^{**} . Put differently, with strong enough complementarities in the production of human capital there might exist multiple basins of attraction, and successive generations of different families would no longer regress toward the same “mean.” As a consequence, even modern societies might be divided into two “classes” with considerably mobility within but not across class boundaries.¹⁶

Thus, in line with traditional models, our theory predicts the intergenerational transmission of economic status to be imperfect; yet it allows for a family’s initial position to exert a great deal of influence over the well-being of many future generations.

6. Concluding Remarks

In this paper, we study the link between cross-sectional inequality and intergenerational mobility. Our main contribution is a model of the intergenerational transmission of resources that explains why mobility varies with the distribution of income.

By explicitly considering complementarities in the production of children’s human capital, we show that, even in a world with perfect capital markets and without differences in innate ability, wealthy parents invest more in their offspring than poorer ones. As a result, persistence of economic status is higher at the top of the income distribution than in the middle. Credit constraints, on the other hand, may produce high persistence among very poor families. Our theory thus rationalizes low intergenerational mobility at both ends of the income

¹⁶ Mobility across class boundaries depends now crucially on the distribution of luck, i.e. v_c . With sufficiently many positive realizations of v_c , lower-class families might cross the threshold into the upper class, i.e. \tilde{H} , and thus start regressing toward H^{**} . The opposite applies to dynasties in the upper class.

distribution, all the while allowing for a mobile middle class. Consequently, our theory implies that societies in which human capital is less equally distributed exhibit, on average, lower rates of intergenerational mobility.

At the same time, we demonstrate that government interventions intended to ameliorate inequality may in fact lower intergenerational mobility, even when they do not directly favor the rich. This is because government programs that are complementary to parental inputs crowd in (rather than crowd out) investments by parents. While such programs are generally more cost-effective than those that substitute for parental inputs, they exacerbate existing disparities in investment in children's human capital. To achieve their goal of reducing inequality as well as intergenerational persistence, most forms of government interventions would have to be skewed significantly towards the poor.

We also consider how changes in the marketplace, such as increases in the returns to education, affect measured mobility. Specifically, our theory predicts that changes in inequality due to increases in the return to human capital have virtually no effect on successive generations' *relative* positions in the income distribution. If, however, persistence is measured with respect to labor earnings, then rising returns to education lead to a one-generation increases in persistence, followed by a long-term decline.

In contrast to standard models of intergenerational mobility, our theory predicts that different dynasties need not regress to the same long-run mean. Even modern societies may be class-based, with considerable mobility within but not across class boundaries. Whether convexities in the transmission of human capital from parents to children are, in fact, strong enough to promote the existence of distinct "classes" remains an important, open question for future research.

References

- Autor, D. H., L. F. Katz, and M. S. Kearney (2008). "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics*, 90(2): 300-323.
- Becker, G. S. (1965). "A Theory of the Allocation of Time." *Economic Journal*, 75(299): 493-517.
- Becker, G. S., and N. Tomes (1979). "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy*, 87(6): 1153-1189.
- Becker, G. S., and N. Tomes (1986). "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, 4(3): S1-S39.

- Becker, G. S., K. M. Murphy, and J. L. Spenkuch (2015). "The Manipulation of Children's Preferences, Old Age Support, and Investments in Children's Human Capital." *Journal of Labor Economics*, forthcoming.
- Behrman, J., and P. Taubman (1985). "Intergenerational Earnings Mobility in the United States: Some Estimates and a Test of Becker's Intergenerational Endowments Model." *Review of Economics and Statistics*, 67(1): 144-151.
- Belley, P., and L. Lochner (2007). "The Changing Role of Family Income and Ability in Determining Educational Achievement." *Journal of Human Capital*, 1(1): 37-89.
- Bénabou, R. (1993). "Workings of a City: Location, Education, and Production." *Quarterly Journal of Economics*, 108(3): 619-652.
- Black, S. E., and P. J. Devereux (2011). "Recent Developments in Intergenerational Mobility," (pp. 1487-1541) in D. Card and O. C. Ashenfelter (eds.), *Handbook of Labor Economics, Vol. 4B*. Amsterdam: Elsevier.
- Carneiro, P. and J. J. Heckman (2002). "The Evidence on Credit Constraints in Post-Secondary Schooling." *Economic Journal*, 112(482): 705-734.
- Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014a). "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." NBER Working Paper No. 19844.
- Chetty, R., N. Hendren, P. Kline, and E. Saez (2014b). "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics*, forthcoming.
- Clark, G. (2014). *The Son also Rises: Surnames and the History of Social Mobility*. Princeton, NJ: Princeton University Press.
- Corak, M. (2013). "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives*, 27(3): 79-102.
- Corak, M., and A. Heisz (1999). "The Intergenerational Earnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data." *Journal of Human Resources*, 34(3): 504-533.
- Council of Economic Advisors (2012). *The Annual Report of the Council of Economic Advisors*. Washington, D.C.: U.S. Government Printing Office.
- Duncan, G., and R. Murnane (2011). "Introduction: The American Dream, Then and Now," (pp. 3-25) in G. Duncan and R. Murnane (eds.), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. New York: Russell Sage.
- Durlauf, S. N. (1996a). "A Theory of Persistent Income Inequality." *Journal of Economic Growth*, 1(1): 75-93.
- Durlauf, S. N. (1996b). "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," (pp. 505-534) in W. Barnett, G. Gandolfo, and C. Hillinger (eds.), *Dynamic Disequilibrium Modelling: Proceedings of the Ninth International Symposium on Economic Theory and Econometrics*. Cambridge, UK: Cambridge University Press.
- Durlauf, S. N. (2006). "Groups, Social Influences, and Inequality: A Memberships Theory Perspective on Poverty Traps," (pp. 141-175) in S. Bowles, S. Durlauf, and K. Hoff (eds.), *Poverty Traps*. Princeton, NJ: Princeton University Press.
- Ellwood, D. T., and C. Jencks (2004). "The Spread of Single-Parent Families in the United States since 1960," (pp. 25-65) in D. P. Moynihan, T. Smeeding, and L. Rainwater (eds.), *The Future of the Family*. New York: Russell Sage.

- Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2014). "Marry Your Like: Assortative Mating and Income Inequality." *American Economic Review Papers and Proceedings*, 105(5): 348-353.
- Guryan, J., E. Hurst, and M. Kearney (2008). "Parental Education and Parental Time with Children." *Journal of Economic Perspectives*, 22(3): 23-46.
- Heckman, J. J., and S. Mosso (2014). "The Economics of Human Development and Social Mobility." *Annual Review of Economics*, 6(1): 689-733.
- Hurst, E. (2010). "The Rug Rat Race: Comments and Discussion." *Brookings Papers on Economic Activity*, Spring 2010: 177-184.
- Jäntti, M., B. Bratsberg, K. Røed, O. Raaum, R. Naylor, E. Österbacka, A. Björklund, T. Eriksson (2006). "American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States." IZA Discussion Paper 1938.
- Juhn, C., K. M. Murphy, and B. Pierce (1993). "Wage Inequality and the Rise in the Returns to Skill." *Journal of Political Economy*, 101(3): 410-442.
- Kaplan, S. N., and J. Rauh (2013). "It's the Market: The Broad-Based Rise in the Return to Top Talent." *Journal of Economic Perspectives*, 27(3): 35-56.
- Katz, L. F., and K. M. Murphy (1992). "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics*, 107(1): 35-78.
- Kennedy, S., and S. Ruggles (2014). "Breaking Up Is Hard to Count: The Rise of Divorce in the United States, 1980-2010." *Demography*, 51(2): 587-598.
- Krueger, A. (2012). "The Rise and Consequences of Inequality in the United States." Speech at the Center for American Progress, Washington D.C.
- Lareau, A. (2011). *Unequal Childhoods: Class, Race, and Family Life* (2nd ed.). Berkeley, CA: University of California Press.
- Lindahl, M., M. Palme, S. Sandgren Massih, and A. Sjögren (2015). "Long-Term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations." *Journal of Human Resources*, 50(1): 1-33.
- Lochner, L., and A. Monge-Naranjo (2012). "Credit Constraints in Education." *Annual Review of Economics*, 4(1): 225-256.
- Loury, G. C. (1981). "Intergenerational Transfers and the Distribution of Earnings." *Econometrica*, 49(4): 843-867.
- Lovenheim, M. F. (2011). "The Effect of Liquid Housing Wealth on College Enrollment." *Journal of Labor Economics*, 29(4): 741-777.
- Mazumder, B. (2005a). "The Apple Falls Even Closer to the Tree than We Thought: New and Revised Estimates of the Intergenerational Inheritance of Earnings," (pp. 80-99) in S. Bowles, H. Gintis, and Melissa Osborne Groves (eds.), *Unequal Chances: Family Background and Economic Success*. New York: Russell Sage.
- Mazumder, B. (2005b). "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *Review of Economics and Statistics*, 87(2): 235-255.
- Murphy, K., and L. Katz (1992). "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics*, 107(1): 35-78.
- Olivetti, C., and D. Passerman (2015). "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940." *American Economic Review*, forthcoming.

- Piketty, T., and E. Saez (2003). "Income Inequality in the United States, 1913-1998." *Quarterly Journal of Economics*, 118(1): 1-39.
- Ramey, V., and G. Ramey (2010). "The Rug Rat Race." *Brookings Papers on Economic Activity*, Spring 2010: 129-176.
- Solon, G. (1992). "Intergenerational Income Mobility in the United States." *American Economic Review*, 82(3): 393-408.
- Solon, G. (1999). "Intergenerational Mobility in the Labor Market," (pp. 1761-1800) in D. Card and O. C. Ashenfelter (eds.), *Handbook of Labor Economics, Vol. 3A*. Amsterdam: Elsevier.
- Solon, G. (2004). "A Model of Intergenerational Mobility Variation over Time and Place," (pp. 38-47) in M. Corak (ed.), *Generational Income Mobility in North America and Europe*. Cambridge, UK: Cambridge University Press.
- Solon, G. (2013). "Theoretical Models of Inequality Transmission across Multiple Generations." NBER Working Paper No. 18790.
- Solon, G. (2015). "What Do We Know So Far about Multigenerational Mobility?" mimeographed, Michigan State University.
- Zeng, Z., and Y. Xie (2014). "The Effects of Grandparents on Children's Schooling: Evidence from Rural China." *Demography*, 51(2): 599-617.

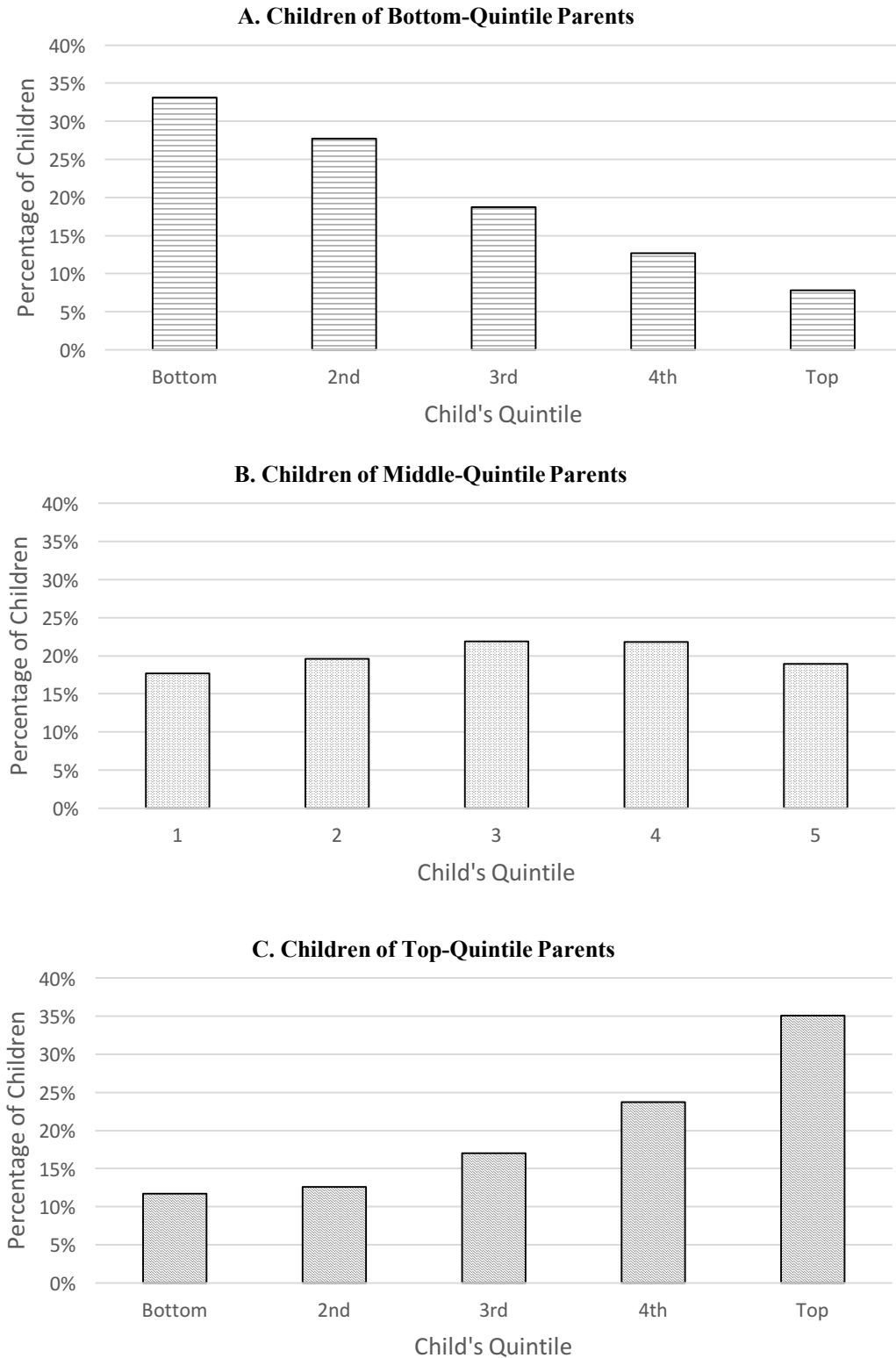
Figure 1: The Great Gatsby Curve



Notes: Figure shows the relationship between inequality, as measured by the Gini coefficient (on the x -axis), and intergenerational persistence, as measured by the intergenerational earnings elasticity (on the y -axis). Higher values respectively indicate more inequality and more persistence.

Sources: Based on Corak (2013).

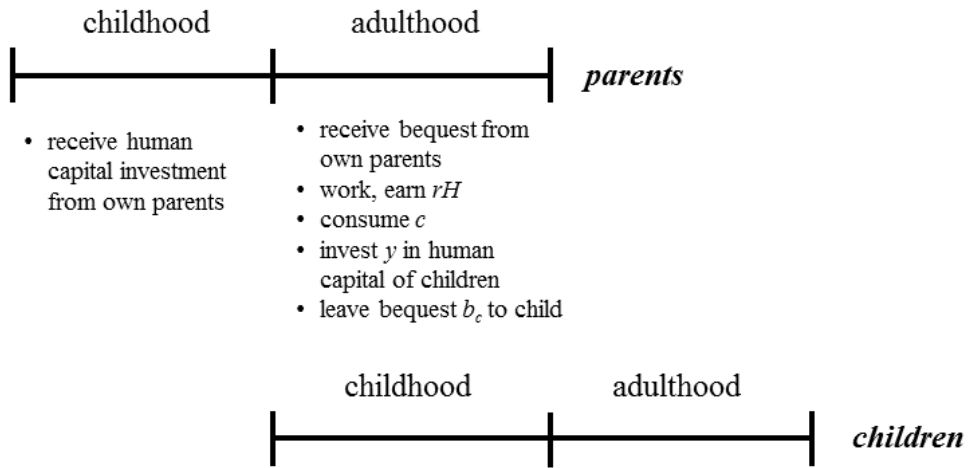
Figure 2: Income Quintiles of Children Conditional on Parental Income, 1980–85 Birth Cohorts



Notes: Panels depict the percentage of children with family income in a particular quintile of their generation's income distribution, conditional on the quintile of their parents.

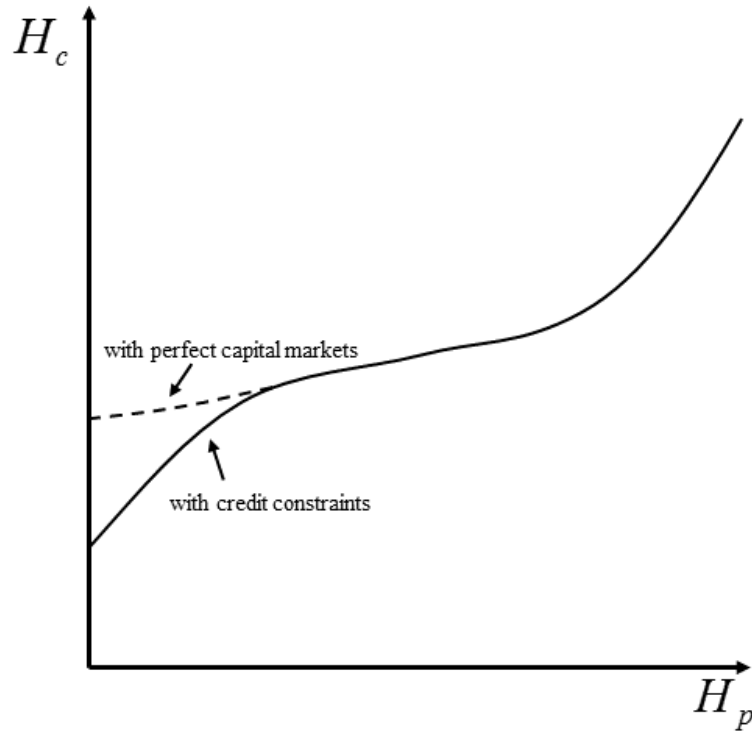
Sources: Based on Chetty et al. (2014b).

Figure 3: Timing



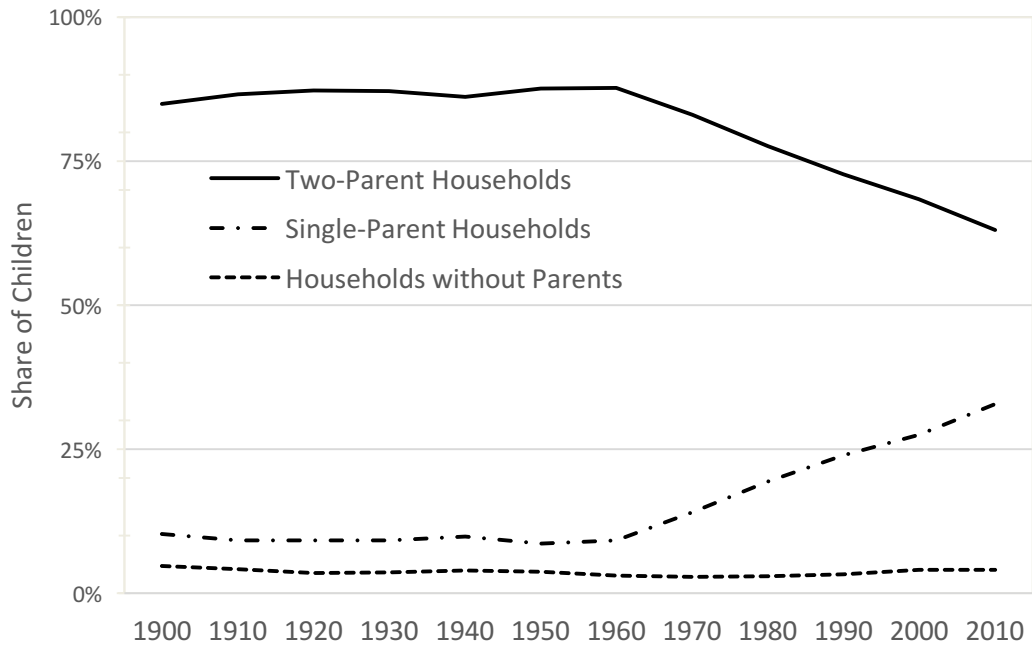
Notes: Figure shows the timing of actions in our model.

Figure 4: Equilibrium Relationship between Children's and Parents' Human Capital



Notes: Figure shows the equilibrium relationship between the human capital of children and that of parents, i.e., the relationship between H_p and H_c after substituting for y in the production function in equation (4).

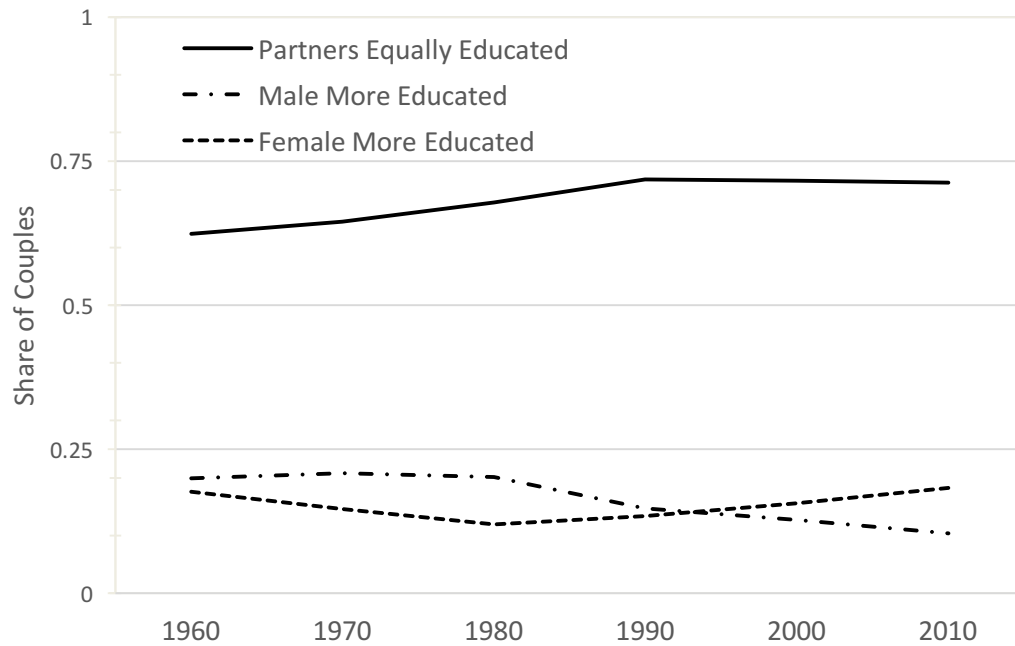
Figure 5: Share of U.S. Children by Family Structure, 1900–2010



Notes: Figure shows the share of U.S. children below the age of 15 living in two-parent households, single-parent households, and households without parents.

Sources: Authors' calculations based on U.S. Census data.

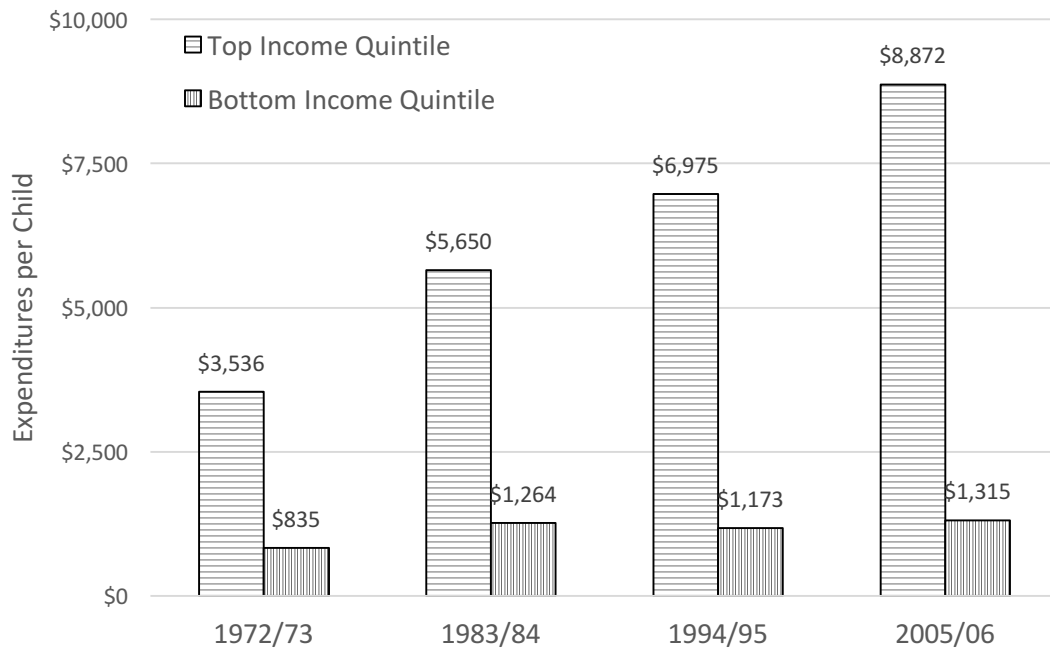
Figure 6: Share of Cohabiting Couples by Relative Educational Attainment, 1940–2010



Notes: Figure shows the share of cohabiting U.S. couples aged 25 to 35 by relative educational attainment. Educational attainment is defined in terms of the following three broad categories: high school dropout, high school graduate, and college graduate.

Sources: Authors' calculations based on U.S. Census data.

Figure 7: Enrichment Expenditures per Child in the U.S., by Parental Income



Notes: Numbers are in 2008 dollars and refer to parental spending on books, computers, child care, summer camps, private school, and other expenditures to promote children's abilities.

Sources: Corak (2013); based on Duncan and Murnane (2011).

Figure 8: Intergenerational Dynamics in Linear Models

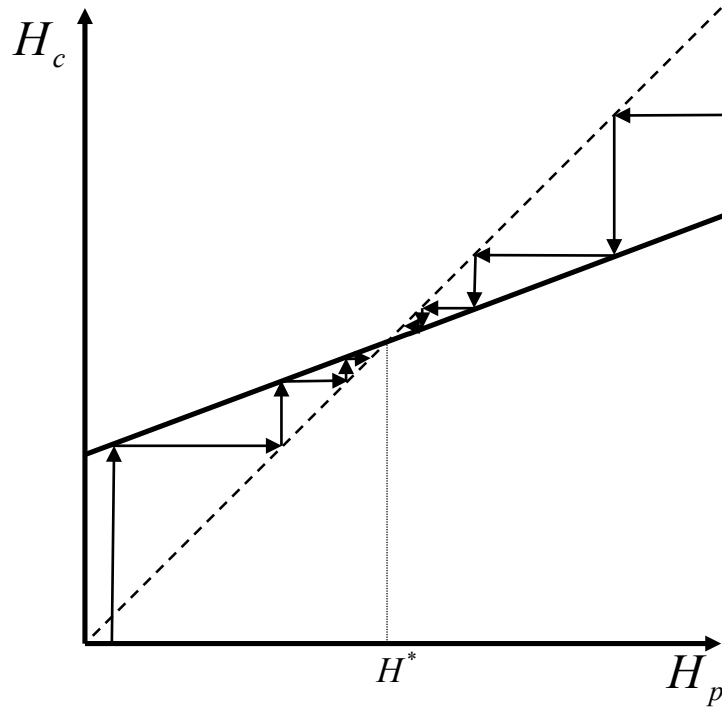


Figure 9: Intergenerational Dynamics when the Production of Children's Human Capital Exhibits Complementarities

