The Causal Effect of Parental Human Capital on Children's Human Capital *

Sang Yoon (Tim) Lee
Department of Economics
University of Mannheim

Nicolas Roys
Department of Economics
University of Wisconsin-Madison

Ananth Seshadri
Department of Economics
University of Wisconsin-Madison

December 2014

Preliminary; comments welcome.

Abstract

We present a structural model of life-cycle human capital accumulation to isolate the direct effect of parents' human capital on children's human capital. Identification of this spillover term comes from the model delivering a relationship between parental human capital, and both the schooling *and* earnings of the child. We demonstrate identification is achieved by comparing the earnings levels of individuals with the same schooling level but different parental schooling levels. A generalized version of the model with taste shocks for schooling is estimated using HRS data, and we find substantial evidence of strong parental spillover effects. We conduct a policy experiment to examine the impact of compulsory schooling laws. These laws have been used as an instrument to isolate the causal effect parental schooling on children's schooling. We find that a large parental spillover term is consistent with both a large OLS coefficient from regressing child schooling on parent schooling, as well as a (close to) zero IV coefficient. Nonetheless, the reform has a positive impact on earnings. This is because much of schooling variation is explained by taste shocks, and higher parental human capital has a level effect, reducing the need for children to spend more time in school. Contrary to some recent debates that put less emphasis on nurture, we conclude that parental spillovers can explain more than half of the human capital transmission from parents to children.

*We thank Steven Durlauf, Jim Heckman and Chris Taber for very helpful conversations. Junjie Guo and Wei Song provided outstanding research assistance.

1 Introduction

Parents have a large influence on their children's outcomes. The correlation between parent and child schooling is as high as 0.5, and even after controlling for the child's schooling, parental schooling or earning levels have a positively significant effect on children's earnings. Does this merely represent correlation in unobserved heterogeneity across generations (selection)? Or does it also partly reflect human capital spillovers from parent to child? These are important questions from the perspective of education policy. After all, if intergenerational correlations merely reflect selection, government subsidies aimed at improving human capital would only impact one generation. On the other hand, in the presence of intergenerational externalities, the returns from such public investments are reaped by all succeeding members of a dynasty, resulting in long-lasting effects. Redistributive education policies may also go beyond reducing inequality within a single generation and have positive impacts on intergenerational mobility.

Furthermore, intergenerational spillovers form an integral part of models of human capital accumulation such as Becker and Tomes (1986). Being able to identify and estimate the magnitude of these spillovers is consequently important not just for better understanding the effects of socioeconomic policy, but also to gain a better understanding of the mechanics through which inequality is formed and persists over generations. How important is nature (transmission of unobserved prenatal conditions) as opposed to nurture (postnatal transmission of human capital) in determining the next generation's economic status?[1] Do forced increases in the schooling of parents affect the schooling of children? These are the questions we seek to answer in this paper.

We begin with a very simple life-cycle model of human capital accumulation owing to Ben-Porath (1967), which encompasses both schooling and learning on-the-job. In this model, individual earning profiles are determined by their initial level of human capital, and the speed, or innate learning ability, at which the individual accumulates human capital.[2] We augment this model in several aspects. First, we posit that an individual's initial level of human capital when he begins schooling at age 6 is also a function of his parental human capital, which is what represents the parental spillover effect in our model. Second, we assume that the initial level of human capital is also affected by innate ability, i.e., the speed at which human capital is accumulated after age 6. Thus, the initial level and speed of human capital accumulation are affected by his own ability, which is unobserved but correlated across generations, while parental human capital, which is observed, only has a direct effect on the child early on in life.

The simple model delivers an analytical expression for the schooling and earnings of a child as a function of parental human capital. The resulting schooling and earnings equations closely resemble those that have been estimated in the empirical literature. In particular, we demonstrate

[1] The purpose of our paper is to measure the size of spillovers, and we estimate the correlation between the observed parental human capital of an individual and the unobserved prenatal conditions. While such prenatal conditions need not necessarily reflect the biological transfer of genes, in this paper we will refer to such conditions as innate abilities or nature.

[2] The Mincer regression is a special case of this set-up when post-schooling time allocation declines linearly until retirement.

that the identification problem of separating nature from nurture may be rectified by using information on earnings and schooling jointly. Intuitively, since schooling and earnings are both functions of parental human capital but the direct spillover occurs early on in life, the magnitude of the spillover can be identified once we know the structural relationship between own schooling and own earnings. Since this is precisely what the Ben-Porath model delivers, we are able to separately identify nurture as opposed to the previous literature that focuses on either the child's schooling or earnings as the outcome variable, but not both simultaneously. The solution to the model also gives us insights into the identification problem—depending on how we control for own schooling in a Mincer regression, the coefficient on parental variables can reveal nature or nurture depending on the underlying relationship between own schooling and earnings.

In the model, all else equal, children with high human capital parents spend less time in school. This is because there is less need to accumulate additional human capital before entering the labor market. It is still the case that these children ultimately attain higher levels of human capital and hence higher earnings, but it takes them less time to reach that level so that they have shorter schooling periods. Since in the data, these children in fact spend more time in school, it must be that high human capital parents, who also tend to have high innate abilities, also pass on their abilities, which overcomes the negative level effect. Now, suppose that we observe two individuals with the same schooling level but different levels of parental human capital and earnings. Through the lens of our model, this reveals the relative magnitudes of their innate abilities as a function of the relative magnitudes of their parents' human capital. Then, the earnings differences between these two individuals can be completely accounted for by parental human capital differences, from which we can recover the size of parental spillovers. Once this is done, the contribution of nature is identified by simply observing how children's schooling levels vary across different parental schooling levels.

This identification scheme relies on parental human capital spillovers having a constant effect over the children's life-cycle earnings, as well as schooling levels. Is such an assumption empirically reasonable? Reduced form evidence suggests that children who attain the same years of schooling, but whose moms have different years of schooling, have parallel earning profiles with a constant gap. The parallel gap points toward the existence of a parental spillover that only affects how much human capital the child accumulates before entering the labor market (schooling), and then remaining constant once controlling for the child's schooling level. Moreover, this gap is indeed similar across different child schooling levels. We show analytically that the spillover is precisely picking up this gap, which is quantitatively large. If our model were true, educating a mom for an extra year is equivalent to having a mom with an extra year of education—i.e., the treatment and selection effects are similar. Furthermore, five extra years of mom's education has the same reduced form effect as one additional year of own schooling, suggesting that 20% of parental human capital spills over to child's earnings.

Of course, because schooling is also endogenous, this does not tell us the exact magnitude of parental human capital on children's earnings is. To bring our model closer to the data and quantify the magnitude of nature (the fact that high ability parents have high ability children) and nurture (parental spillovers), we extend the basic model and estimate it to data on individual schooling and earnings, and parental schooling. We explicitly account for three sources of heterogeneity—unobserved innate ability and a taste for schooling, and observed parental human capital. We let both ability and taste be correlated with parental human capital. This is to capture the fact that high human capital parents can have high human capital children both because they pass on their high ability, or because of intergenerational human capital spillovers, which was already present in the simpler model. By letting tastes for schooling also be correlated with parental human capital, we let the relationship between schooling and parents freely vary from the relationship between earnings and parents. For example, highly educated parents may send their children to school longer for reasons unrelated to higher future earnings, or simply motivate the children to emulate their parents regardless of future economic outcomes. In fact, our estimates suggest that most of schooling differences can be explained by this preference heterogeneity, and more importantly, that nurture can explain approximately 55-60% of intergenerational earnings transmissions.

By no means are we the first to estimate the contribution of nature and nurture in the transmission of socioeconomic status across generations. The general consensus in the literature is that genetics play a large role, consistently above 50% in the determination of children's schooling or IQ. Some notable works include Behrman and Taubman (1989), who use data on twin parents and extended family relationships to decompose the variance of observed years of schooling into the variance of genetic and environmental variables, and Behrman and Rosenzweig (2002) who, also using data on twin parents, finds that mothers' education has a zero or even marginally negative impact on children's schooling. More recently, Plug and Vijverberg (2003), using data on biological and adopted children to separately identify the coefficient on genetic variables, find a larger role for nurture, although it is still significantly below 50%. In contrast, Black et al. (2005) using administrative data from Norway and using compulsory schooling reforms as an instrument, again find a near-zero IV coefficient of parental schooling on children's schooling.[3]

Underlying these studies is the idea that both nature and nurture effects from the parent should increase schooling. As we have already emphasized, however, it may well be that large spillovers (a large nurture effect) *decrease* schooling, while increasing earnings. Since we in fact do observe that children of higher human parents stay longer in school, this means that previous studies that do not allow for such a negative relationship underestimate both the effect of nature and spillovers. The goal of our paper is to compare these nature-nurture effects when taking account of this relationship.

In our model, the level effect is a reduced-form representation of all parental effects before the child reaches age 6. This is related to the large literature initiated by Cunha and Heckman (2007), who also emphasizes that the early childhood environment is hard to separate from prenatal and/or genetic transmissions. At the same time, however, it is puzzling that the early childhood literature finds large nurture effects, while the earlier literature finds only little. Since this

[3] See Sacerdote (2011) for an extensive review.

literature predominantly takes children's schooling outcomes as the dependent variable, it is intuitive that in the process of doing so, they miss the level effect and consequently attribute a larger role to nature. We emphasize that it is important to compare both earnings and schooling outcomes to recover this effect, which is a contribution of our model over previous studies that rely on a single measure for child outcomes.[4]

A related literature uses compulsory schooling reforms as a natural experiment. Black et al. (2005) find that the fact that some parents were forced to go to school for one or two more additional years in Norway had almost no impact on the schooling of their children. They interpret this as evidence suggesting that "the apple does not fall far"—that is, most of the correlation in schooling across generations is due to selection (correlation in innate abilities across generations). We argue that such experiments also miss the spillover effect by using only one measure of children's outcomes. Even when the intergenerational schooling response is zero or even negative, we may observe a significant effect on the next generation's earnings. In addition, since we also find that heterogeneous tastes for schooling can account for a large amount of schooling outcome variations, the intergenerational effects on schooling outcomes are even more dampened unless they also alter the distribution of taste heterogeneity. To be precise, we use our model economy to conduct a similar thought experiment in which we were to force all parents to go to school for a compulsory number of years. The model is consistent with a high OLS estimate of parents' schooling on children's schooling as well as a near zero IV estimate, despite significant parental spillovers on the child's earnings.

The rest of the paper is organized as follows. Section 2 posits a simple model of human capital accumulation and adds to it a parental spillover. The solution to this model is derived analytically and the quantitative implications of this model explored. In section 3 we describe the HRS data and argue that our model assumptions are consistent with the empirical evidence. Section 4 presents a more comprehensive model which features multiple sources of heterogeneity which is estimated to the data. We discuss how we use data on parent and child schooling as well as child earnings to estimate the parental spillover term. Section 5 examines counterfactual predictions of the model including a hypothetical compulsory schooling reform. Section 6 concludes.

2 A Simple Model with Parental Spillovers

In this section, we lay out a simpler version of the model than is later to be estimated, and characterize its solution to gain intuition on how the parameters are identified. The model is a simple variant of a life-cycle human capital accumulation model owing to Ben-Porath (1967). The objective of an individual is to maximize present-value discounted lifetime earnings. This model is the

[4]Our structural model has implications not only for schooling outcomes but also earnings, which is what allows us to identify the level effect. Conversely, the level effect would also be identified if we had a measure of children's human capital at age 6. However, such measures (young children's cognitive test scores, vocabulary, etc.) are hard to link to their later economic ability earn. Nonetheless, although we are looking at two cognitive outcomes and not any non-cognitive outcomes, our argument that the childhood environment can have different impacts on different outcome variables are in line with Cunha et al. (2010).

workhorse for modern labor economics and has been used widely in various contexts—the effects of rising skill premia, cross country income differences, the study of inequality within a country and so forth. For completeness, we first present the model without intergenerational spillovers. Then we alter it to incorporate parental effects of which solution is compared with Mincer regressions in section 3.

An individual begins life at age 6, retires at age $R > 6$ and dies at age $T \geq R$, all exogenous to the individual. At age 6, an individual is described by the state vector (h_0, z), which denotes his initial stock of human capital and (learning) ability, respectively. At every age $a \in [6, R)$, he makes time and good investments $(n(a), m(a))$ into human capital accumulation to maximize the present discounted value of net income. His problem is

$$\max_{n(a),m(a)} \left\{ \int_6^R e^{-r(a-6)} \left[wh(a)(1-n(a)) - m(a) \right] da \right\}$$

subject to

$$\begin{aligned}
&\dot{h}(a) = z[n(a)h(a)]^{\alpha_1} m(a)^{\alpha_2}, \qquad a \in [6, R), \\
&n(a) \in [0,1], \qquad m(a) \geq 0, \\
&h(6) = h_0
\end{aligned}$$

where α_1 and α_2 are the returns to time and good investments, respectively. The market wage w and discount rate r are constant and taken as given by the individual. Since the market wage multiplies human capital to generate earnings, human capital is understood as an individual's "earning abilities" as opposed to z, the learning ability.

In this framework, the time path of $n(a)$ decreases with age $n(a)$ (Ben-Porath, 1967). Assuming decreasing returns to scale, i.e. $\alpha_1 + \alpha_2 < 1$, if the initial stock of human capital is low enough, the time allocation decision is constrained at 1 for a few years and then declines. The period of time that n equals 1 is typically labeled the schooling period. All else equal, individuals with higher ability levels (z) make more human capital investments and stay longer in school, while individuals with a higher initial stock of human capital (h_0) stay less time in school.

The above decision problem is a finite horizon problem. When the individual retires, his stock of human capital depreciates completely. There is no transmission of human capital from one generation to the next. The above framework cannot generate the correlation in schooling across generations without resorting to exogenous transmission in abilities, z. Furthermore, if each subsequent cohort begins life with the same initial stock of human capital, parents do not have any influence on the earnings profiles of children. This leads us to think further about intergenerational linkages. Clearly, parents have an influence on both the learning ability of the child (z) as well as the amount of learning that happens before school entry (h_0).

We now augment the model with spillovers from parents to children. We do this in the simplest possible manner—we assume that initial human capital is a function of one's own learning ability z and the parent's human capital h_P. By assuming that abilities are correlated with h_P, we face the identification problem of whether the child's schooling or earnings are correlated with the parent because of z, which is nature, or h_P, which is nurture. We further assume that individuals do not internalize this externality.[5] Consequently, we are simply solving a life-cycle decision problem in which parental human capital is taken as given by the child as an exogenous state variable. We do this primarily to gain insight into the precise mechanisms at work and understand how the model parameters can be recovered from the data. Relaxing this assumption can be accomplished easily by adding parental preferences over the utility or human capital of their children. Given that the bulk of investments are made early in life and markets are assumed to be complete, the incorporation of dynastic utility does not have much of a qualitative impact on the decision problem.

2.1 The Individual's Problem

We include intergenerational spillovers into the Ben-Porath framework. As above, individuals maximize their present discounted value of net income. All that we change is the initial condition:

$$h(6) = z^{\lambda} h_P^{\nu} \tag{1}$$

with parental human capital, h_P, given. This is simply a relabeling of the initial stock of human capital $h_0 \equiv z^{\lambda} h_P^{\nu}$. Then the individual state is still (h_0, z), and can be solved using exactly the same tools used to solve the standard model. To be clear, while the standard framework takes the initial stock of human capital as parameters, we want to estimate how it interacts with learning ability and parental human capital.

The parameter λ relates own learning ability to the initial stock of human capital—while z determines the speed of learning after age 6, we expect it to also affect learning prior to age 6, the extent of which is what this parameter is intended to capture. In other words, while z is transmitted across generations prenatally, λ captures the postnatal effect of the transmission. The intergenerational spillover is captured by ν, the degree to which a higher human capital parent transmits more human capital to her child during the formative first 6 years of life. This is a rather standard assumption in the literature on intergenerational transmission, for instance Becker and Tomes (1986), and can be considered a reduced form characterization of the importance of early childhood.

Let $V(a, h)$ denote the value function for an individual of age a and human capital level h. The problem faced by an individual at age 6 given $h(6) = h_0$ can be written

$$V(6, h_0) = \max_{\{n(a), m(a)\}} \left\{ \int_6^R e^{-r(a-6)} g\left(h(a); n(a), m(a)\right) da \right\}$$
$$\dot{h}(a) = f\left(h(a); n(a), m(a)\right), \qquad n(a) \in [0,1], \qquad m(a) > 0.$$

[5] In other words, individuals do not invest more in their own human capital in anticipation of that investment spilling over to subsequent generations. If they do, nurture would have even a larger effect.

where the objective function and law of motion are given as

$$g(h; n, m) = -whn - m$$
$$f(h; n, m) = z(nh)^{\alpha_1} m^{\alpha_2}.$$

This is a continuous time deterministic control problem with state h and controls (n, m). The terminal time is fixed at R but the terminal state $h(R)$ must be chosen. Since the objective function is linear, the constraint set strictly convex, and the law of motion strictly positive and concave (since $\alpha \equiv \alpha_1 + \alpha_2 < 1$), the optimization problem is well-defined and the solution is unique. The Hamilton-Jacobi-Bellman equation is

$$rV(a,h) - \frac{\partial V(a,h)}{\partial a} = \max_{n,m} \left\{ g(h; n, m) + \frac{\partial V(a,h)}{\partial h} f(h; n, m) \right\}.$$

As usual, the HJB equation can be interpreted as a no-arbitrage condition. The left-hand side is the instantaneous cost of holding a human capital level of h at age a, while the the right-hand side is the instantaneous return. The first order conditions for the controls are

$$whn \leq \alpha_1 z(nh)^{\alpha_1} m^{\alpha_2} \cdot V_h, \qquad \text{with equality if } n < 1 \tag{2}$$
$$m = \alpha_2 z(nh)^{\alpha_1} m^{\alpha_2} \cdot V_h, \tag{3}$$

where V_h is the partial of $V(a, h)$ with respect to h. These conditions simply state that the marginal cost of investment, on the left-hand side, is equal to the marginal return. The envelope condition gives (at the optimum)

$$r \cdot V_h - V_{ah} = w(1-n) + \frac{\alpha_1 z(nh)^{\alpha_1} m^{\alpha_2}}{h} \cdot V_h + z(nh)^{\alpha_1} m^{\alpha_2} \cdot V_{hh}, \tag{4}$$

where V_{xh} is the partial of V_h with respect to $x \in \{a, h\}$. This "Euler equation" states that at the optimum, the marginal cost of increasing human capital must be equal to the marginal return. Equations (2), (3) and (4) along with the law of motion

$$\dot{h} = z(nh)^{\alpha_1} m^{\alpha_2} \tag{5}$$

characterize the complete solution, given the initial state $h(6) = h_0$ and terminal condition $V_h = 0$, the appropriate transversality condition for a fixed terminal time problem. We solve this problem in Appendix A and here only present the important results.

Proposition 1: Optimal Schooling Choice *Define* $\alpha \equiv \alpha_1 + \alpha_2$ *and*

$$F(s)^{-1} = \frac{\alpha_1^{1-\alpha_2} (\alpha_2 w)^{\alpha_2}}{r} \cdot \left[1 - \frac{(1-\alpha_1)(1-\alpha_2)}{\alpha_1 \alpha_2} \cdot \frac{1 - e^{-\frac{\alpha_2 rs}{1-\alpha_2}}}{1 - e^{-r(R-6-s)}} \right]^{\frac{1-\alpha}{1-\alpha_1}} \cdot \left(1 - e^{-r(R-6-s)} \right).$$

The optimal choice of schooling S is uniquely determined by

$$F'(S) > 0, \qquad F(S) \geq z^{1-\lambda(1-\alpha)} h_P^{-\nu(1-\alpha)} \tag{6}$$

with equality if $S > 0$.

Proof. See Appendix A. □

The higher the innate ability z of an individual, the higher the optimal choice of his schooling (as long as $\lambda(1-\alpha) < 1$). The effect of parental human capital h_P depends on its spillover to the initial stock of human capital (ν) and correlation with z. In particular, note that the spillover effect on schooling is *negative*. This implies that, even if we were to observe that increasing parental human capital had no effect on children's schooling, it would not imply that there are no spillovers, but that the spillover countervails the correlation between z and h_P.

PROPOSITION 2: POST-SCHOOLING HUMAN CAPITAL *For any S, human capital at the end of schooling,* h_S, *satisfies*

$$wh_S = C_1(S) \cdot z^{\frac{1}{1-\alpha}}$$

where

$$C_1(s) = \left[\frac{\alpha_1^{1-\alpha_2} \alpha_2^{\alpha_2} w^{1-\alpha_1}}{r} \cdot \left(1 - e^{-r(R-s)}\right) \right]^{\frac{1}{1-\alpha}}.$$

Proof. See Appendix A. □

The above proposition tells us that, once the length of schooling is known, the human capital level of a child is affected only by his own learning ability z. In other words, the parental effects of ν is subsumed in the length of schooling. His initial stock of human capital, h_0, has no effect on the *amount* of human capital accumulated except through the *length* of schooling, S.

Using Proposition 2, we can also describe the dependence of children's earnings profiles on parental human capital. Assume that a fraction π_n of time investments n, and π_m of goods investments m, are subtracted from the value of the human capital to obtain measured earnings. In other words, we are assuming that the individual pays for the job training costs in the form of lower contemporaneous wages.

COROLLARY 1: EVOLUTION OF EARNINGS PROFILES *For an individual who attains S years of schooling, for all* $a \in [6+S, R)$,

$$e(a) = wh(a)\left[1 - \pi_n n(a)\right] - \pi_m m(a) = \left[C_1(S) + C_2(a; S)\right] \cdot z^{\frac{1}{1-\alpha}}$$

for any $(\pi_n, \pi_m) \in [0,1]^2$, *where*

$$C_2(a;s) = \left(\frac{\alpha_1^{\alpha_1}\alpha_2^{\alpha_2}w^{1-\alpha_1}}{r}\right)^{\frac{1}{1-\alpha}} \cdot \left\{ r\left[\int_{6+s}^{a} q(x)^{\frac{\alpha}{1-\alpha}}dx\right] - (\alpha_1\pi_n + \alpha_2\pi_m)\left(1 - e^{-r(R-a)}\right)\right\}$$

for $a \geq s - 6$.

Proof. See Appendix A. □

By virtue of Corollary 1, what fraction of job training costs are paid for by the firm only depends on age, as long as it is assumed to be constant. More importantly, Proposition 1 and Corollary 1 imply that both λ and the correlation between parental human capital and a child's learning abilities are identified once we have data on schooling, earnings and parental human capital.

Corollary 2: Identifying Nature Effects 1. *Suppose we observe children with the same level of parental human capital* $h_P = \hat{h}_P$. *Then for any* $a \in [6+S, R)$,

$$e(a|h_P = \hat{h}_P) \propto [C_1(S) + C_2(a;S)] \cdot F(S)^{\frac{1}{(1-\alpha)[[1-\lambda(1-\alpha)]}} \tag{7}$$

so if α *is known,* λ *is recovered from a Mincer regression controlling for parental human capital.*

2. *Suppose that* $\log z = \rho_{zh_P} \log h_P + \eta$, *and that we have a large, representative sample for same aged individuals in which we observe* (h_P, S, e) *(but not* z*). If we run the regressions*

$$\begin{aligned} \log F(S) &= a_0 + a_1 \log h_P + \epsilon_s \\ \log e(a) &= b_0 + b_1 \log\left[C_1(S) + C_2(a;S)\right] + b_2 \log h_P + \epsilon_e, \end{aligned} \tag{8}$$

we recover the estimates

$$\hat{a}_1 = \rho_{zh_P}\left[1 - \lambda(1-\alpha)\right] - \nu(1-\alpha), \qquad \hat{b}_2 = \frac{\rho_{zh_P}}{1-\alpha}.$$

Proof. Suppose we observe two individuals with different levels of schooling and age a earnings but identical levels of parental human capital, denoted by (S_1, S_2), (e_1, e_2), and (h_{P_1}, h_{P_2}), respectively. Let (z_1, z_2) denote their *unobserved* innate abilities. Then by Corollary 1,

$$\frac{e_1}{e_2} = \frac{[C_1(S_1) + C_2(a;S_1)]}{[C_1(S_2) + C_2(a;S_2)]} \cdot \left(\frac{z_1}{z_2}\right)^{\frac{1}{1-\alpha}},$$

so ability differences are identified by earnings differences and schooling differences. Furthermore by Proposition 1

$$\frac{F(S_1)}{F(S_2)} = \left(\frac{z_1}{z_2}\right)^{1-\lambda(1-\alpha)} \qquad \Rightarrow \qquad \frac{e_1}{e_2} = \frac{[C_1(S_1) + C_2(a;S_1)]}{[C_1(S_2) + C_2(a;S_2)]} \cdot \left[\frac{F(S_1)}{F(S_2)}\right]^{\frac{1}{(1-\alpha)[1-\lambda(1-\alpha)]}}$$

so λ is identified. The second part of the Corollary follows trivially from Proposition 1 and **??** after plugging in the assumed relationship between z and h_P.

$\square$

Put simply, earnings and schooling differences identify ability differences and hence the magnitude of λ, *conditional on parental human capital*. This is because ν has no role to play in identifying nature's effect on initial human capital, λ. However, even when λ is known, however, regressing children's schooling on parental outcomes does not identify whether the child-parent relationship reflects selection of spillovers when ν is large. On the other hand, a Mincer regression that controls for schooling *linearly* (because the functions (C_1, C_2) are close to exponential, their logarithm is close to linear), and age, would reveal that the coefficient on parental human capital captures only nature and completely miss nurture.

Despite the fact that we are working with a stylized model, the result that we can simply consider the Mincer coefficient on parental human capital as nature may be surprising, since it does not pose any identification problems at all.[6] We will see in Proposition 3, however, that conditional on schooling—not linearly—it is the difference in earnings profiles *across individuals* that identifies ν from the data, even though ν has no impact on earnings once the length of schooling period is determined for any given individual. In other words, identifying nature and/or nurture effects depends on how we control for the relationship between own schooling and earnings.

2.2 Recovering Parental Spillovers

One of the advantages of our framework is that it is possible to measure the parental spillover by using data on individual earnings, own schooling and parental human capital.[7] A robust finding in empirical studies is that even after controlling for observables, mothers' education has a statistically significant relationship (not causation) with children's schooling and earnings. Our spillovers are a structural representation of this. In this section, we demonstrate that our theory generates clear-cut predictions from which we can estimate the spillover terms ν using micro data.

Recall from Proposition 1 that learning ability z has a positive influence on schooling as long as $\lambda(1-\alpha) < 1$. Controlling for ability (which is unobserved), i.e., the fact that children of higher parental human capital will tend to have higher learning abilities (through correlation of z across generations), a higher parental human capital level implies a higher initial stock of human capital for the child (though ν), which decreases schooling. We now argue that one can use information on individual earnings, schooling and parental human capital (inferred from parental schooling) to estimate the relative importance of ν relative to θ.

[6] In an environment in which z is transmitted intergenerationally, ρ_{zh_P} is positive only to the extent that high ability parents who tend to have high human capital also have high ability children.

[7] In the data, we observe parental schooling, but not earnings. On the other hand, we do observe earnings for the parent generation. To estimate the model, we link parental earnings to schooling using a separate Mincer regression on the parent generation to capture the parental human capital distribution.

COROLLARY 3 *Suppose we observe children with the same level of schooling, $S \neq 0$. Then for any $a \in [6+S, R)$, earnings depend only on parental human capital through ν and not on innate abilities, i.e.*

$$e(a|S = \hat{S}) \propto h_P^{\frac{\nu}{1-\lambda(1-\alpha)}} \tag{9}$$

so if α is known and since λ can be identified from Corollary 2, ν can be identified from regressing earnings on parental human capital, controlling for own schooling.

Proof. Suppose we observe two individuals with the same level of schooling and age a earnings but different levels of parental human capital, denoted by (S_1, S_2), (e_1, e_2), and (h_{P_1}, h_{P_2}), respectively. Denote the ability of these two individuals by (z_1, z_2). Since the years of schooling are the same for these two individuals, by Proposition 1 we must have

$$\left(\frac{z_1}{z_2}\right)^{1-\lambda(1-\alpha)} = \left(\frac{h_{P_1}}{h_{P_2}}\right)^{\nu(1-\alpha)} \qquad \Rightarrow \qquad \frac{z_1}{z_2} = \left(\frac{h_{P_1}}{h_{P_2}}\right)^{\frac{\nu(1-\alpha)}{1-\lambda(1-\alpha)}} \qquad \Rightarrow \qquad \frac{e_1}{e_2} = \left(\frac{h_{P_1}}{h_{P_2}}\right)^{\frac{\nu}{1-\lambda(1-\alpha)}}$$

where the last relationship follows from Corollary 1. So ν is identified. □

Corollary 3 shows that our model has clear implications for the level and steepness of the age-earnings profiles of two individuals with the same years of schooling but parents with different levels of human capital: the steepness of the profiles should be identical while parental human capital differences manifest as level differences. Given that we have already shown that λ is identified, the earnings differentials between individuals with the same years of schooling but different parental human capital identifies ν. If there is no earnings differential whatsoever across these individuals, $\nu = 0$. On the other hand, sizable earnings differentials across such individuals will be associated with large values for ν.

The next natural questions are whether such level differences are empirically reasonable, and how to control for age and schooling in the data to separately identify ρ_{zh_P} from ν, i.e., the nature and nurture effects. In section 3 we present evidence that justifies our assumptions and present raw evidence on the magnitude of ν, and in section 4 we present a generalized model with taste shocks so that we do not overestimate the structural effects from the stylized model. Before this, however, we compare our model with previous studies.

2.3 Comparison to Previous Studies

By Proposition 1, for $S \in [6, R-6)$ we can write

$$\log F(S) = [1 - \lambda(1-\alpha)] \cdot \log z - \nu(1-\alpha) \cdot \log h_P.$$

Now suppose that $F(S)$ can be approximated by

$$F(S) = \exp\left[\tilde{a}_1(S - \tilde{a}_0)\right].$$

Then assuming α is a known parameter, we can write

$$
\begin{aligned}
& S = -\tilde{a}_0 + \frac{1-\lambda(1-\alpha)}{\tilde{a}_1} \cdot \log z - \frac{\nu(1-\alpha)}{\tilde{a}_1} \cdot \log h_P \\
\Rightarrow \qquad & S = a_0 + a_1(\lambda) \cdot \log z + a_2(\nu) \cdot \log h_P. \qquad (10)
\end{aligned}
$$

If $\log z$ is observable, (10) is a testable equation given data on $(S, \log h_P)$ for a cross section of individuals. The classical challenge with running such a regression, of course, is that $\log z$ is unobserved and correlated with $\log h_P$, which creates an endogeneity problem.

In this section, we use equation (10) to compare our model with the previous empirical literature which will highlight what we gain from a structural estimation. Most previous studies overcome the endogeneity problem by using special datasets: i) twins and families, ii) adopted vs biological children, and iii) compulsory schooling reforms as natural experiments. The problem is that all previous studies look at a unidimensional child outcome as a dependent variable, which is usually schooling. Our structural model points to the fact that more information may be gained by looking at the joint distribution of schooling *and* earnings of the child.[8]

Twins and Extended Families Given (10), we can write a variance decomposition similar to traditional twin studies[9] as

$$
\begin{aligned}
\mathbb{V}[S] &= [a_1(\lambda)]^2 \, \mathbb{V}[\log z] + [a_2(\nu)]^2 \, \mathbb{V}[\log h_P] + 2a_1(\lambda)a_2(\nu)\mathbb{C}[\log z, \log h_P] \\
&= [a_1(\lambda)]^2 \sigma_z^2 + [a_2(\nu)]^2 \sigma_{h_P}^2 + 2a_1(\lambda)a_2(\nu)\rho_{zh_P}\sigma_z\sigma_{h_P},
\end{aligned}
$$

where σ_z is the standard deviation of abilities, σ_{h_P} the standard deviation of parental human capital and ρ_{zh_P} the correlation between children's innate abilities and parental human capital. The more recent literature has already pointed out the drawbacks to such a decomposition, one being that even if it were to work perfectly, relative contributions to the total R^2 does not in fact tell us how nature and nurture affect children's outcomes (unless there is an underlying structure, as in our model).[10] Another concern is that they impose $\rho_{zh_P} = 0$, while especially for twins and families, this correlation is potentially high, making it hard to interpret the decomposition. Such studies infer the unobserved variation in parental ability from the observed variation in child and parent outcomes, typically schooling. Hence any observed variation in child outcomes unexplained by observed variations in parental outcomes can be biased toward being attributed to unobserved genetic variation.

Adopted vs Biological Children The earnings equation (10) resembles the reduced form models studied in Plug and Vijverberg (2003) or Björklund et al. (2006). The main challenge for, and

[8] Some studies, of course, do use the child's earnings or income as the dependent variable, but then they do not use this jointly with information on the child's schooling.

[9] E.g., Behrman et al. (1977). Behrman and Taubman (1989) extends this to account for relatives.

[10] In virtually all the non-twin study papers we cite, i.e. Plug and Vijverberg (2003); Black et al. (2005); etc.

usually, also the main contribution of such studies, is how to control for nature, which is unobservable. The former study exploits the difference between adopted and biological children, while the latter uses information on twin parents. In contrast to such studies, we instead posit a structural relationship between own schooling and earnings, and exploit data on both to obtain estimates for ν separately (Proposition 3). To clarify why our approach is desirable, suppose that all log variables are centered at zero and that

$$\log z = \rho_z \log z_P + \eta_z \tag{11a}$$

$$\log h_P = \rho_P \log z_P + \eta_P. \tag{11b}$$

Equation (10) becomes

$$S = a_0 + \underbrace{a_1(\lambda)\rho_z \log z_P}_{\text{(a)}} + \underbrace{a_2(\nu)\rho_P \log z_P}_{\text{(b)}} + \underbrace{a_2(\nu)\eta_P}_{\text{(c)}} + a_1(\lambda)\eta_z. \tag{12}$$

In this equation, both (a) and (b) come from the parent's genetic ability, but (a) represents how much of the child's schooling can be explained by genetic persistence (nature), while (b) represents how much of human capital spillovers (nurture) can be explained by the parent's genetic ability. Similarly, both (b) and (c) are nurture, but (c) is the part of nurture that cannot be explained by parental ability alone. In other words, (a) is purely prenatal and (c) postnatal, but (b) is an interaction of both.

The basic idea in the adoption literature is to capture the pure nature effect by exploiting the fact that for adopted children, ρ_z must be zero, or at least smaller than for biological children. For example, Plug and Vijverberg (2003), proxy for parental ability using IQ test scores (z_P in our model). Then the difference in the coefficients on parental IQ between biological and adopted children identifies nature. But the empirical counterpart to h_P in their framework is parental income. To obtain a measure for ϵ_P, they purge parental income of the parent's own ability,[11] which they use in the estimation. The nature-nurture gap (the ratio between the coefficients for biological and adopted children, or a_1 in the above equation) drops by half once parental income is purged as opposed to when it is not. This indicates that not appropriately accounting for the effect of the parent's ability on parental income produces upwardly biased estimates of nature, which is apparent from the above equation because otherwise (b) would be incorrectly attributed to nature while in fact it is nurture.

We further argue that IQ is only a noisy measure of nature which should partially be categorized as nurture, which would also bias the estimate upward.[12] Relatedly, note that a_2 is also nurture, but is ignored because they cannot separate whether it comes from parental ability or human capital. The presence of an interaction such as (b) highlights how sensitive a nature-nurture

[11] This is achieved by regressing parental income on parental IQ, parental schooling and grandparent's schooling. In our context, this means they are also controlling for the human capital component of income, so what they actually get can be interpreted as random shocks to parental human capital.

[12] Also emphasized in Sacerdote (2002).

decomposition can be depending on what measures for parental nature or nurture effects are taken. Indeed, a linear decomposition may not be possible to begin with. For example, Björklund et al. (2006) explicitly match adopted children to their biological parent which allows them to explicitly account for (a) and (b) separately, and find not only a larger effect for nurture, but that the interactive component (b) is significant.

Compulsory Schooling Reforms as a Natural Experiment Finally, our structure can yield large values for nurture as measured by ν while remaining consistent with studies that find that increased parental schooling following the advent of compulsory schooling laws had no effect on the next generation's schooling, such as Black et al. (2005).[13] In the context of our model, previous studies that use compulsory schooling laws as a natural experiment run a Two Stage Least Squares (2SLS) of the form

$$\begin{aligned} S &= b_0 + b_1 \cdot \log h_P + \mathbf{b}_2 \cdot [X\ X_P]' + \epsilon \\ \log h_P &= b_0^P + b_1^P \cdot \text{REFORM} + \mathbf{b}_2^P \cdot [X\ X_P]' + v, \end{aligned}$$

where S is the child's years of schooling, $\log h_P$ is parental human capital, and REFORM is a dummy that equals 1 after implementation of the law, and 0 before. The vectors $[X, X_P]$ are individual characteristics of the child and parent, respectively. The 2SLS structure is supposed to take REFORM as an instrument for $\log h_P$. If this is a valid instrument, we can recover parental spillovers from the estimates of (b_1, b_1^P) by looking at the changes in $(S, \log h_P)$ after the reform. Assuming that the reform increases parental schooling by Δ_{S_P} years, children's schooling should increase by Δ_S years according to

$$\Delta \log h_P = b_1^P \Delta_{S_P} \quad \Rightarrow \quad \Delta_S = b_1 \cdot \Delta \log h_P = b_1 b_1^P \Delta_{S_P}. \tag{13}$$

Hence $b_1 b_1^P$ is an estimate of a parental spillovers terms of years of schooling.

We already saw in the previous subsection that the fact that increasing parental education has no effect on children's education may be an outcome of the nature and nurture countervailing each other, so that the estimated b_1 need not be large. At the same time, individual schooling decisions are not only affected by economic earnings abilities (z), so that a weak response in children's schooling does not necessarily translate into parental human capital having no effect on their children. Indeed, in our counterfactual analysis, we find that much of schooling variation can be explained by preference heterogeneity. For these reasons, we also find that a large OLS coefficient is consistent with a small IV coefficient even when the structural spillover parameter ν is large. Most importantly, the effect of the counterfactual reform is found in the children's counterfactual earnings, not schooling.

[13] In contrast to Norway, Oreopoulos and Page (2006) find a significant effect in the U.S.. But their dependent variable is whether children younger than 15 repeat a grade, which is difficult to compare to our model.

3 Data Analysis

We showed above that it is possible to measure parental spillovers by using individuals' earnings and schooling data. A robust finding in empirical studies is that even after controlling for observables, a mother's education has a significant impact on children's schooling and earnings.[14] In the previous section we argued that such a significant coefficient can be capturing either nature or nurture, depending on how the relationship between schooling and earnings are controlled for, even when the structural spillover is large. This section presents the data we use to estimate the model along with a preliminary empirical analysis to compare the predictions of the simple model with available empirical evidence. This sheds light on how the spillover parameter is identified in the estimated model of section 2.

3.1 Data

The Health and Retirement Study (HRS) is sponsored by the National Institute of Aging and conducted by the University of Michigan with supplemental support from the Social Security Administration. The HRS is a national panel study with a sample (in 1992) of 12,652 persons in 7,702 households. It over-samples blacks, Hispanics, and residents of Florida. The sample is nationally representative of the American population 50 years old and above. The baseline 1992 study consisted of in-home, face-to-face interviews of the 1931-41 birth cohort and their spouses, if they were married. Follow up interviews have continued every two years after 1992. As the HRS has matured, new cohorts have been added.

A large fraction of HRS respondents gave permission for researchers to gain access, under tightly restricted conditions, to their social security earnings records. Combined with self reported earnings in HRS, these earnings records, although top-coded in some cases, provide almost the entire history of earnings for most of the HRS respondents. We impute earnings histories for those individuals with missing or top-coded earnings records assuming the following individual log-earnings process

$$
\begin{aligned}
\log w_{i,0}^{*} &= x_{i,t}'\beta_0 + \varepsilon_{i,0} \\
\log w_{i,t}^{*} &= \rho \log w_{i,t-1}^{*} + x_{i,t}'\beta_x + \varepsilon_{i,t}, \qquad t \in \{1, 2, ..., T\} \\
\varepsilon_{i,t} &= \alpha_i + u_{i,t}
\end{aligned}
$$

where $w_{i,t}^{*}$ is the latent earnings of individual i at time t in 2008 dollars, $x_{i,t}$ is the vector of characteristics at time t, and the error term $\varepsilon_{i,t}$ includes an individual specific component α_i, which is constant over time, and an unanticipated white noise component $u_{i,t}$. We employed random-effect assumptions with homoskedastic errors to estimate above model separately for men with and without a college degree. Scholz et al. (2006) gives details of the above earnings model and its coefficient estimates, along with a description of the procedure used to impute earnings for

[14] One exception is Behrman and Rosenzweig (2005), once they control for twin parents.

Table 1: Summary Statistics by Years of Schooling

	HSG	CLG	All
Schooling	12.00	16.00	12.21
			(3.27)
Mother's schooling	9.38	10.95	9.45
	(2.87)	(3.07)	(3.42)
Father's schooling	8.76	10.82	8.96
	(3.29)	(3.58)	(3.85)
% White	86.66	90.99	84.39
% Black	11.14	6.46	13.04
Earnings at age 30	36.55	37.84	33.73
	(17.03)	(18.22)	(17.55)
Earnings at age 40	48.28	56.13	47.30
	(22.04)	(28.89)	(24.27)
Earnings at age 50	46.68	69.35	50.69
	(30.11)	(53.51)	(42.51)
Sample size	1823	588	5677

- HSG=12, CLG=16 years of schooling.
- Standard deviations in parentheses.
- Earnings inflated to 2008, measured in \$1000.

individuals who refuse to release or who have top-coded social security earnings histories.

For the purpose of this paper, we start with 30,548 individuals in the RAND HRS Version J (RAND, June 2010). We keep 6967 male respondents born between 1924 and 1941. We drop 32 individuals with missing information on years of schooling. Another 127 individuals are dropped because we can not get any of their earnings histories between age 25 and age 55 even with the above model. We drop another 1131 observations for whom we have no earnings data either between 24 and 26 or between 49 and 51, the trough and peak of average earning profiles. This leaves us with 5677 observations. Table 1 describes this sample by years of schooling, and Table 9 in the appendix describes it by lifetime earnings, which is defined to be the sum of earnings from age 25 to age 55.[15]

3.2 Mincer Regression

We first augment a standard Mincer regression with parental schooling. We estimate the regression model

$$\log w_i = \beta_0 + \beta_1 S_i + \beta_2 S_{P,i} + f(\text{EXP}_i) + \beta_X X_i + \epsilon_i \tag{14}$$

[15]We use the HRS because it gives us the longest information on individuals' earnings histories while still providing information on parental schooling, and also because these older individuals and their parents were less affected by compulsory schooling. However, the evidence we present here is also consistent with available data from the PSID or NLSY79.

Table 2: Mincer Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
S (EDUC)	0.0657***	0.0599***	0.0609***	0.0592***	0.0591***	0.0869***	0.0600***
	(113.27)	(86.57)	(86.44)	(79.90)	(79.87)	(46.15)	(78.15)
Mother S_P		0.0127***			0.0116***	0.0157***	0.0130***
		(21.63)			(15.55)	(10.83)	(20.12)
Father S_P			0.0092***		0.0026***		
			(17.03)		(3.93)		
Parental S_P				0.0068***			
				(20.73)			
EXP						0.116***	
						(64.90)	
EXP^2						-0.0018***	
						(-74.38)	
EXP$\times S$						-0.0013***	
						(-17.34)	
EXP$\times$Mother S_P						-0.0001	
						(-1.95)	
Observations	171194	150684	146077	140380	140380	150684	127666
R^2	0.156	0.154	0.151	0.151	0.152	0.152	0.133

- OLS regressions of (log) earnings on years of schooling and years of schooling of the parents controlling for potential experience, race and cohort effects. t-stats shown in parentheses.
- Age 25-55, cohort born 1926-1941, Male, earnings below first percentile are trimmed.

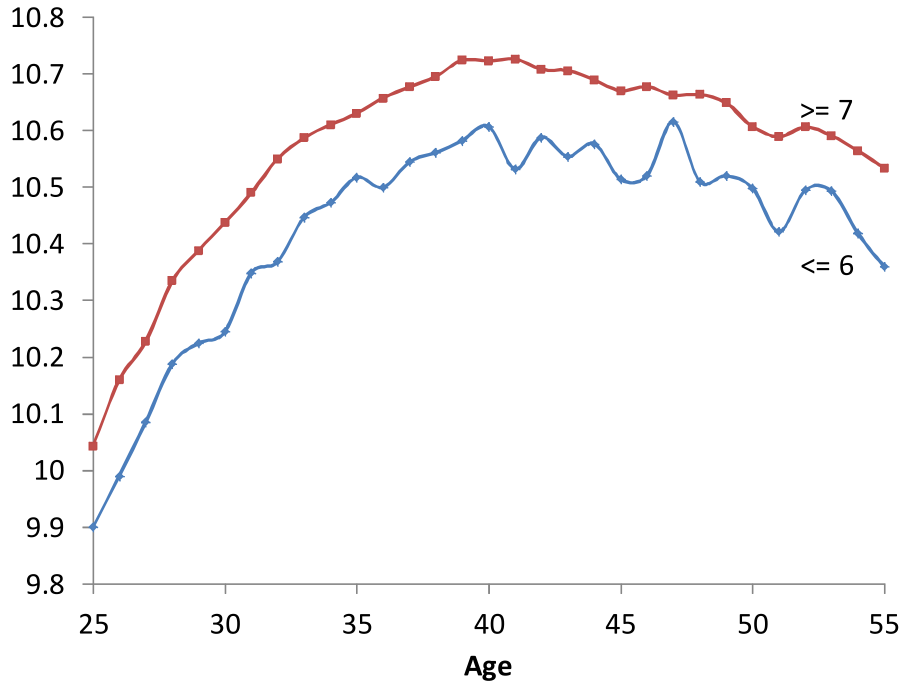
where $f(\cdot)$ is a flexible function we specify in various different ways below, and w_i, S_i, $S_{P,i}$, and EXP$_i$ denote, respectively, earnings, years of education of individual i, years of education of individual i's parents, and potential experience (age). The vector X_i is a vector of demographic control variables (race and cohort dummies) and ϵ_i is an error term. We estimate different specifications of (14) using data from the HRS, and tabulate the results in Table 2.

We consider three measures for $S_{P,i}$: mother's years of education, father's years of education, and the sum of both parent years of schooling. Our theory is silent on which of these measures is more appropriate. However, many studies have found that parental inputs have its strongest impact on children's human capital early on (the ν effect), e.g. Del Boca et al. (2012), who also find that mothers spend more time with their children at an early age. This leads us to suspect mothers should play the dominant role, which is confirmed in our results in Table 2.

The first specification (1) is a standard Mincer regression and uses dummies for each potential experience level observed in the data. The return to schooling is estimated to be 6.57%. This is in the lower range of the estimated returns to schooling for more recent cohorts, which is in line with the increased return to education over the last century (see Goldin and Katz (2007)). The returns slightly decrease to 5.99% when we include mother's years of schooling in the regression (column 2). The coefficient on mother's education is 1.27% and is statistically significant. This suggests that an additional year of schooling for the child has about the same effect on earnings as would being born to a mother with five additional years of schooling.

The results are quite similar when we measure parental human capital with the father's year of education (column 3) but with an attenuated coefficient. The rate of return of paternal education

Figure 1: Evidence of Spillovers.



Earnings profiles of high school graduates (12 years of schooling) by mother's schooling level: 1926-1941 birth cohort. The y-axis is annual earnings in 2008 USD.

is 0.92% and is statistically significant. The coefficient drops further when we measure parental human capital as the sum of the schooling of both parents (column 4). If we include both (column 5), mother's education is very slightly reduced from 1.27% to 1.16% while the coefficient on fathers drop significantly from 0.93% to 0.26%. In column (6) we replaced the experience dummies with a quadratic in experience, and also added two interactions terms: one between parental education and experience and another between education and experience. The interaction between parental education and experience is insignificant. The results are quite stable when the sample is restricted to white males only (column 7).

Overall, we find that mothers have a stronger impact on sons than fathers and that the estimated effect of parental education, or β_2, is about 1.5%. These results are in line with those reported by previous empirical work on this topic (e.g. Card (1999)). The question is, does this represent intergenerational persistence of abilities or spillovers?

Nature in our model indicates would be captured by the β_2 coefficient (Corollary 2). To capture nurture, we should not assume a log-linear relationship between the child's earnings and schooling but condition completely on the child's schooling level. Figure 2 depicts the average age-earnings profile for male high school graduates controlling for cohort effects. We split individuals into two subsamples depending on whether or not the mother has more than 6 years of schooling. The estimated profiles support the importance of parental human capital since we observe higher earnings for individuals with more educated mothers. It appears essentially as a

Table 3: Auxiliary Parameters

Schooling			Educational	Average Earnings at age			
Mom S_P	Own S	$\overline{S}$	Attainment (%)	23-27	28-32	33-37	38-42
$\leq$5	$\leq$11	6.35	60.04	1.00	1.41	1.71	1.85
	$\geq$12	13.12	39.96	1.26	1.91	2.32	2.51
6-7	$\leq$11	8.02	42.48	1.10	1.66	1.98	2.11
	$\geq$12	13.53	57.52	1.27	1.91	2.37	2.65
8	$\leq$12	10.72	66.23	1.35	1.99	2.42	2.62
	$\geq$13	15.21	33.77	1.31	2.12	2.71	2.99
9-11	$\leq$12	11.00	63.73	1.37	1.98	2.38	2.57
	$\geq$13	15.24	36.27	1.30	2.15	2.65	2.96
12	$\leq$12	11.21	45.41	1.52	2.24	2.62	2.77
	$\geq$13	15.38	54.59	1.36	2.29	2.87	3.34
$\geq$13	$\leq$12	11.41	19.56	1.39	1.96	2.33	2.49
	$\geq$13	15.83	80.44	1.30	2.23	2.82	3.27

Note that for mom's with low S_P (the first four rows), we divide whether the child's educational attainment was low or high by whether or not he graduated from high school, while for the rest by whether he advanced beyond high school. In the third column, $\bar{S}$ denotes the average years of schooling attained in each category. All average earnings are normalized by the average earnings from 23-27 of the group with less than 12 years of schooling whose moms attained 5 years or less of schooling.

permanent level effect that persists throughout an individual's career with no evidence of increasing steepness of the age-earnings profile.

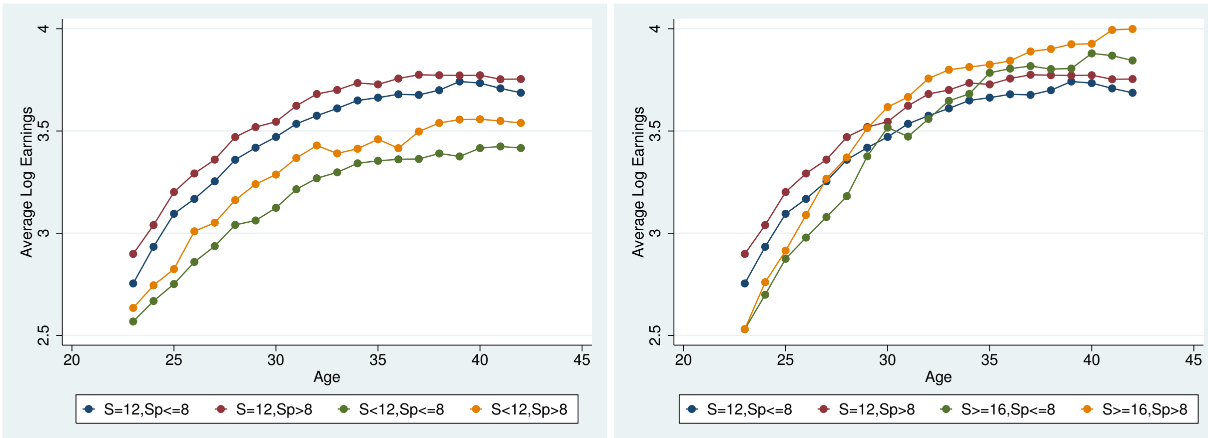
It is precisely this gap that captures parental spillovers in our model. To justify our assumption of a constant spillover, we further need to examine whether this gap remains stable across different levels of children's schooling. Furthermore, this hinges on the gap actually capturing something different from the previous Mincer regression coefficients. To see if this is true, we need to examine whether or not the earnings-schooling relationship for children is in fact log-linear, evidence of which we present next.

3.3 Evidence of Spillovers

To understand the empirical evidence in relation to the model we estimate, we drop more observations and focus on the exact data we use for the estimation. Specifically, we drop those individuals for whom we cannot construct average earnings for ages 25,30,35 and 40, which are in turn computed by simply averaging an individual's earnings from ages 23-27, 28-32, and so forth. This still leaves us with 4317 individual observations and 86340 annual earnings observations. The auxiliary parameters used in the estimation are tabulated in Table 3.

In Figure 2, we plot the average earnings profiles of individuals, now in logs, divided by their own education levels and mothers' education levels. In the left panel, we compare the profiles

Figure 2: Identifying Spillovers



(a) Children with 12 vs. less than 12 years of schooling. (b) Children with 12 vs. 16 or more years of schooling.

Earnings profiles of children of different schooling levels by mother's schooling level: 1926-1941 birth cohort. The y-axis is average log annual earnings in 2008 USD. Mothers' schooling levels are divided by 8 years or below, and more than 8 years.

of children with 12 years of schooling split by whether or not their mothers attained 8 years or schooling, against the same profiles but of children with less than 12 years of schooling. In the right, we compare the profiles of children with 12 years of schooling against those with 16 or more years of schooling (college).

There are several things to note. First, note that the average log earnings profiles of children with the same education level but different mother's schooling levels are not only nearly parallel, but more or less constant across all three categories of the children's educational attainment. This is further evidence of a constant parental spillover parameter ν that affects children's earnings with constant log-level effect. Second, the profiles of children with different education levels but the same mother's schooling levels are neither parallel nor do the gaps look similar across children's education levels. This is consistent with equation (7) in Corollary 2, since the earning differences are increasing in age (through the function C_2); and also (6) in Proposition 1, since the relationship between $\log z$ (which governs earnings) and S are non-linear through the function $F(s)$. Nonetheless, these gaps are what would intuitively identify λ, the early nature effect.

Put together, while it does seem that parental schooling is linearly related to children's earnings controlling for children's schooling, the converse is not true—in fact, it seems almost as if earnings are increasing in a concave matter in schooling, once controlling for mom's schooling. If our simple model were true, these differences are picking up the selection effects in the determination of an individual's optimal schooling decision, which is non-linear in log earnings. Then by forcing the earnings-schooling relationship to be log-linear as in a Mincer regression, parental spillovers would be soaked into the coefficient on own schooling and the coefficient on mother's schooling would instead reflect the selection, as in equation (8) in Corollary 2.

Table 4: Augmented Mincer Regressions

	(1)	(2)	(3)	(4)
S	0.076***	0.109***	0.107***	
	(74.08)	(39.42)	(36.78)	
Mom S_P	0.017***	0.017***	0.022***	0.022***
	(19.62)	(20.15)	(10.67)	(25.60)
EXP$\times S$		-0.002***	-0.002***	
		(-12.70)	(-10.97)	
EXP$\times S_P$			-0.000	
			(-2.37)	
$S < 12$				-0.375***
				(-50.71)
$S \in (12, 16)$				0.048***
				(6.42)
$S \geq 16$				0.196***
				(23.85)
R^2	0.1843	0.1860	0.1861	0.1792

OLS regressions of log earnings on different controls, t-stats shown in paranthesis. All regressions also include a linear and quadratic term for potential experience EXP, not reported in the table but all significant at the 1% level. The magnitudes of the coefficients on EXP and EXP^2 are similar to Table 2, and also statistically equal to each other for specifications (1) and (4), and (2) and (3), respectively.

Equipped with this intuition, we rerun several augmented Mincer regressions to gain a sense of the magnitudes of selection (ρ_{zh_p}) and spillovers (ν) in our model. All the regressions in Table 4 includes a linear and quadratic term for potential experience, as well as race and cohort dummies. The magnitudes of the coefficients are similar to what was found in Table 2, but with slightly larger coefficients on both S and S_P which we suspect is due to dropping noisier earnings observations at later stages of the lifecycle.

In column (1), we run the standard Mincer regression with own and mother's schooling S and S_P; according to our model, the coefficient on S_P here is explained more by nature than nurture. In columns (2)-(4), we control for S in a way that our model tells us that the coefficient on S_P would instead reveal nurture. In column (2), we include an interaction term for experience and S. Since from Figure 2, a potential reason that the profiles would not reveal nurture was their different slopes by S, controlling for this should bring us closer to recovering ν from the coefficient on S_P. For robustness, in column (3) we also include an interaction between experience and S_P, which is insignificant and hence in line with our intuition.[16]

Note that although the coefficients on S_P are all similar across different specifications, the interaction between EXP and S in (2) and (3), and the coefficients on the three education level dummies in (4) all point toward a non-linear relationship between education and earnings. Hence, according to our model, this should rather be interpreted as the implied magnitudes of nature and nurture

[16]This coefficient was also insignificant in Table 2.

being similar, with each having approximately one-fifth an effect on earnings as own schooling.

Of course, without knowing the magnitudes of the rest of the parameters in the model, and in particular α, which governs the returns to human capital accumulation, and λ, the early nature effect, we cannot recover any point estimates. Furthermore, since schooling is also a function of nature and parental human capital, a single earnings equation even if controlling for schooling does not reveal the causal effect of nature or spillovers on schooling nor earnings. As we emphasized in the introduction, we need to be able to exploit both pieces of information, which is what our model does for us. To estimate the exact values of the spillover and its causal effect, we generalize the simple model of section 2 to bring it closer to the data.

4 Estimation

The exercise above suggests a novel way of estimating intergenerational parental human capital spillovers. The model has stark implications for the impact of parental schooling on children's outcomes, and by conditioning on the schooling of the child, we are able to get around some of the selection issues commonly encountered in empirical work. The goal of the generalized model we estimate in this section is to retain the simplicity of the framework presented in section 2 and empirical intuition from section 3, and yet be able to obtain realistic predictions on the transmission of human capital and schooling across generations.

The underlying environment is one in which innate ability z is transmitted across generations. However, our framework does not require us to make specific assumptions on this transmission process per se. The magnitude of the spillover is determined only the state of the child, which consists of his or her innate ability z and the human capital level of the parent h_P. Hence, for our purposes all that is required is the estimation of ρ_{zh_P}, the correlation between z and h_P. Since an overriding objective is to fit the distribution of schooling, we will also include a taste for schooling in the objective function. Many previous studies find that psychic costs, in addition to innate ability (selection), play an important role in rationalizing schooling decisions.

We show in our counterfactual experiments that the coefficient on parental human capital in a Mincer regression captures nature, in the sense that when ρ_{zh_P} is set to zero, the counterfactual coefficient is also close to zero. We also show that the coefficient barely changes even if we set $\nu = 0$, indicating that our structural model is consistent with the difficulty of identifying nurture in the data. Nonetheless, we show that a counterfactual increase in parental human capital leads to a XX% increase in children's earnings even controlling for selection, while selection accounts for XX%. Furthermore, this happens without increasing the children's schooling. Part of this is explained by the heterogeneity in tastes for schooling. Given this, we conduct a hypothetical schooling reform counterfactual which shows that a large OLS coefficient when regressing child's schooling outcomes on parent's schooling is consistent with a near-zero IV coefficient, but has a XX% effect on earnings.

4.1 A Generalized Model of Parental Spillovers

The generalized model we estimate assumes different laws of motion during schooling and on the job. We also assume that a non-pecuniary benefit from schooling which varies across individuals. The problem faced by an individual at age 6 can be written as

$$V(6, h_0) = \max_{\{n(a), m(a), S\}} \left\{ \int_6^{6+S} e^{-r(a-6)} \left[\epsilon - m(a)\right] da + \int_{6+S}^{R} e^{-r(a-6)} h(a) \left[1 - n(a)\right] da \right\}$$

subject to

$$\begin{aligned} \dot{h}(a) &= \begin{cases} z h(a)^{\alpha_1} m(a)^{\alpha_2}, & \text{for } a \in [6, S), \\ z[n(a)h(a)]^{\alpha_W}, & \text{for } a \in [S, R), \end{cases} \\ h(6) &= h_0 = b z^{\lambda} h_P^{\nu}. \end{aligned} \tag{15}$$

The variable ϵ represents a non-pecuniary benefit from schooling (but measured in pecuniary units) that varies across individuals but stays constant throughout the life-cycle. Since ϵ only affects an individual's desire to remain (or not) in school while having no direct effect on earnings, the inclusion of taste heterogeneity allows the model to flexibly account for schooling-earnings relationships that do not solely rely on economic factors (parental human capital and learning ability). This not only helps to account for the data but also ties our hands to not label everything as nature or nurture.

The only changes we have made in addition to the taste shock is to explicitly split the schooling and working phase, during which the human capital accumulation technology differs. Schooling only involves goods inputs while OJT only involves time inputs. The technology in the schooling phase is identical to the simple model with $n(a) = 1$, while the working phase is identical to the the simple model with $\alpha_2 = 0$. We also allow α_W, the returns to human capital investments during the working phase, differ from the schooling phase. The parameter b that multiplies initial human capital captures the overall level of human capital in the model, while we have dropped the wage rate w since it is not separately identified from b in our partial equilibrium setup (i.e., it is not separately identified from units of human capital without modeling the demand for labor).

At age 6, an individual is completely characterized by (h_P, z, ϵ), and while we cannot derive closed form solutions for schooling and earnings as in the simple model,[17] the optimal schooling and resulting level of earnings can be characterized using similar methods. In Appendix B, we characterize the equations governing optimal schooling and earnings given an individual state (h_P, z, ϵ), and describe how a solution is found numerically in Appendix C.

[17] This is primarily because in general, the supply of labor, $1 - n(a)$, jumps from 0 to a strictly positive amount once an individual begins to work, unlike in the simple model where it increases continuously over time.

4.2 Indirect Inference

Our dataset contains information on parental schooling, children's schooling as well as complete earnings profiles of children. Since we only have information on the schooling of he parent, we approximate the parental human capital, or earnings, of the parent by a standard Mincerian equation relating parental schooling to earnings, $h_P = a\exp(\beta S_P)$. The only usage of β is to gain a statistical relationship over the initial conditions for the children. The coefficient a is ignored since it is not separately identified from b in the child's initial human capital (15).

We also make assumptions on the population distribution of (h_P, z, ϵ). We assume that both $(\log h_p, \log z)$ and $(\log h_P, \epsilon)$ are joint normal, but that $\log z$ and ϵ are independent conditional on h_P. Specifically, we assume that

$$\begin{bmatrix} \log h_P \\ \log z \\ \epsilon \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_h \\ \mu_z \\ \mu_\epsilon \end{bmatrix}, \begin{bmatrix} \sigma_h^2 & 0 & 0 \\ \rho_{zh_P}\sigma_h\sigma_z & \sigma_z^2 & 0 \\ \rho_{\epsilon h_P}\sigma_h\sigma_\epsilon & 0 & \sigma_\epsilon^2 \end{bmatrix} \right).$$

Given the relationship $h_P = \exp(\beta S_P)$, we have the grid points for h_P once we know S_P. To get the grid points for z and ϵ, notice that our distributional assumptions imply

$$\begin{aligned} \log z | \log h_P &\sim \mathcal{N}\left(\mu_z + \rho_{zh_P}\frac{\sigma_z}{\sigma_h}\left(\log h_P - \mu_h\right), \sigma_z^2\left(1-\rho_{zh_P}^2\right)\right) \\ \epsilon | \log h_P &\sim \mathcal{N}\left(\mu_\epsilon + \rho_{\epsilon h_P}\frac{\sigma_\epsilon}{\sigma_h}\left(\log h_P - \mu_h\right), \sigma_\epsilon^2\left(1-\rho_{\epsilon h_P}^2\right)\right). \end{aligned}$$

For each grid point of h_P, we discretize these two conditional distributions to get the grid points for z and ϵ according to Kennan (2006).

For each combination of (h_P, z, ϵ), we solve the model numerically as described in Appendix C. This induces the optimal choice of schooling and life-cycle earnings for any given initial condition. The relevant model moments, $\Psi(\Theta)$, are computed by integrating over the population distribution of this vector.

The full set of structural parameters is the vector Θ. We partition this vector into two vectors, i.e. $\Theta = [\Theta_p, \Theta_e]$, where $\Theta_p = [a, w, \beta, r, R, \mu_h, \sigma_h]$ includes parameters that are set a priori. The rest of the parameters, $\Theta_e = [b, \lambda, \nu, \alpha_1, \alpha_2, \alpha_W, \mu_z, \sigma_z, \rho_{zh_P}, \mu_\epsilon, \sigma_\epsilon, \rho_{\epsilon h_P}]$, are estimated by indirect inference.

Parameters Set a Priori As already explained above, the parameters a (the constant of from running a Mincer regression of oldest cohorts earnings) and w (the wage rate) are normalized to 1, because they are not separately identified from b (the constant multiplying initial human capital). Since human capital in our model is essentially efficiency wage units, without a demand side for human capital we cannot separate the average level of human capital from the wage.

The coefficient β is recovered from running Mincer regressions similar to (14) for the earliest cohorts in our sample, only without parental schooling, which is mostly unobserved for this small sample (513 men and 603 women). Furthermore, the β coefficient is quite stable across cohorts, ranging from approximately 0.04 to 0.06 for men and 0.05 to 0.09 for women; we fix $\beta = 0.06$ in our model. Note that this value is also in a similar range as the coefficients we recover from the pool of later cohorts in Tables 2 and 4, which includes more controls; our estimates are not sensitive to different values of β within this range.

Given $h_P = \exp(\beta S_P)$, we have $\log h_P = \beta S_P$. Thus $\mu_{h_P} = \beta \mathbb{E} S_P$ and $\sigma_{h_P} = \beta \sigma_{S_P}$. We take the mean and variance of mother's schooling, $\mathbb{E} S_P$ and σ_{S_P}, directly from their sample analogs in the data, so the only assumption we are imposing is on the correlation structure. These values are 9.26 and 3.52, respectively. We then get $\mu_{h_P} = 0.556$ and $\sigma_{h_P} = 0.211$.

The interest rate r and retirement age R are fixed at 5% and 65, respectively.

Estimated Parameters We are left with 12 parameters to be estimated. Indirect inference works by selecting a set of statistics of interest, which the model is asked to reproduce.[18] These statistics are called sample auxiliary parameters $\hat{\Psi}$ (or target moments). For an arbitrary value of θ_e, we use the structural model to compute auxiliary parameters $\Psi(\theta_e)$. The parameter estimate $\hat{\theta}_e$ is then derived by searching over the parameter space to find the parameter vector which minimizes the criterion function:

$$\hat{\Theta}_e = \arg\min_{\theta_e \in \Theta_e} \left(\hat{\Psi} - \Psi(\theta_e)\right)' W \left(\hat{\Psi} - \Psi(\theta_e)\right)$$

where W is a weighting matrix and Θ_e the estimated parameters space. This procedure generates a consistent estimate of θ_e. We use the inverse of the variance-covariance matrix of the data moments $\hat{\Psi}$ as weighting matrix. Following Hall and Horowitz (1996), it is estimated by bootstrap.

The minimization is performed using Nelder-Mead simplex algorithm. While this method does not guarantee global optima , we used more than a million different starting values to numerically search over a wide range of parameter values (for which most have natural boundaries).

The standard errors are obtained using 500 bootstrap repetitions. In each bootstrap repetition, a new set of data is produced by randomly selecting blocks of observations.[19] In the bth boostrap repetition, auxiliary parameters $\hat{\Psi}^b$ are calculated using the new set of data. An estimator $\hat{\theta}_e^b$ is found by minimizing the weighted distance between the recentered bootstrap auxiliary parameters $\left(\hat{\Psi}^b - \hat{\Psi}\right)$ and the recentered simulated auxiliary parameters $\left(\Psi(\theta_e^b) - \Psi(\hat{\theta}_e)\right)$:

$$\hat{\theta}_e^b = \arg\min_{\theta_e^b \in \Theta_e} \left(\left(\hat{\Psi}^b - \hat{\Psi}\right) - \left(\Psi(\theta_e^b) - \Psi(\hat{\theta}_e)\right)\right)' W \left(\left(\hat{\Psi}^b - \hat{\Psi}\right) - \left(\Psi(\theta_e^b) - \Psi(\hat{\theta}_e)\right)\right).$$

In all cases, the auxiliary parameters $\widehat{\Psi}$ include all moments tabulated in Table 3 years of schooling, except of course one of the child's educational attainment share conditional on mom's schooling (since they add up to 1) and the average earnings from 23-27 of the group with less than 12 years

[18]See Gourieroux et al. (1993) for a general discussion of indirect inference.

[19]See Hall and Horowitz (1996) for more details on the Block-Bootstrap. The sampling is random across households but is done in block over the time dimension.

Table 5: Parameter Estimates

Spillover/Childhood		HCapital Production	
b	2.307	α_1	0.185
λ	1.641	α_2	0.368
ν	0.964	α_W	0.422
Ability dist		Taste dist	
μ_z	-0.624	μ_ϵ	8.056
σ_z	0.049	σ_ϵ	5.439
ρ_{zh_P}	0.348	$\rho_{\epsilon h_P}$	0.730

of schooling whose moms attained 5 years or less of schooling, since it is normalized to 1. We also include five additional auxiliary parameters in $\widehat{\Psi}$: the Mincer regression coefficients on S and S_P from specification (1) of Table 4, the correlation between S and S_P, the OLS regression coefficient from regressing S on S_P, and the aggregate level of log earnings. The first four moments are included to capture the earnings and schooling gradients in the data we may miss by targeting aggregated moments, while the last moment is included to discipline b, the initial level of human capital that is identical for all individuals. In sum, we have 70 auxiliary parameters to match with 12 model parameters.

5 Results

Table 5 reports the 12 estimated parameters and standard errors (COMING SOON). The parameters that govern human capital production, $(\alpha_1, \alpha_2, \alpha_W)$ are in the lower range of estimates found in the literature that use comparable Ben-Porath human capital models, e.g. Huggett et al. (2011) use an estimate of $\alpha_W = 0.6$ for post-schooling human capital production.

There are a few things to note. First, the returns to human capital investment is slightly larger during schooling ($\alpha_1 + \alpha_2 = 0.553$) than on the job ($\alpha_W$, and since wages are normalized to 1). This means that for purposes of human capital accumulation, an individual would prefer to stay in school rather than work. Second, the magnitudes of λ and ν indicate that both abilities and spillovers are important during early childhood. As a precautionary note, that $\lambda > 1$ should not be taken as an anomaly—since learning ability units are normalized to the unobserved individual productivity of human capital production after age 6, this merely means that learning ability is relatively important before age 6. (In other words, we could have set $h_0 = zh_P^\theta$ and $\dot{h} = z^{\tilde{\lambda}} h^{\alpha_1} m^{\alpha_2}$ instead, with $\tilde{\lambda} = 1/\lambda$.)

Most importantly, the spillover parameter ν is quite large; indicating that parental human capital influences a child's initial level of human capital almost one for one. However, note that by construction of the model, this is the only way that the parent can influence the child's education (besides correlation through abilities and tastes). Since we proxy h_P by parental earnings, ν encompasses any input that influences the child's human capital that can be explained by parental earnings. For example, if the child's human capital is sensitive to early childhood investments as in Cunha and Heckman (2007), it would imply a large ν since we abstracted from dynamic investments from the parent. Relatedly, since parental inputs likely matter not only before but also after age 6, it may be that ν is capturing parental influence that in fact occurs over a longer time horizon.

A large ν may also indicate a large degree of intergenerational altruism. Although we have abstracted from individuals internalizing the spillover, in an environment where it is internalized it would incentivize individuals (of all generations) to invest more in their own human capital beyond maximizing their own life-cycle income. Then, ν would be capturing a composite of spillovers (how much parental human capital matters for children) and altruism (how much parents care about their children). For our purposes, however, we would still correctly capture the elasticity of initial human capital in response to parental human capital (which is what ν measures) as long as altruism does not vary much across the population.

One may also question that we recover a large estimate because we have assumed that parental human capital has only a level effect. However, if parental human capital also had a slope effect (i.e., by augmenting z), the level effect would have be even larger (since the level and slope effects have opposite effects on the length of schooling) while the slope effect would push the contribution of nature downward. Hence, by only assuming a level effect for parental human capital while assuming both for unobserved abilities, we are taking a conservative stance on the magnitude of parental spillovers on earnings.

Lastly, it may seem that the preference shock is an order of magnitude larger than abilities; this is not the case since we assumed that z is distributed log-normal but ϵ normal.

5.1 Interpreting the Parameters

To get a first impression of how these parameters affect the model, we elucidate the quantitative importance of each by shutting down the spillover and correlation parameters. Table 6 tabulates the counterfactual changes in the correlation between children's and parent's schooling levels, and also the Mincer coefficient β_2 on mom's schooling from running regression (1) in Table 4.

Note that shutting down ν does little to affect either moment; as expected, β_2 is mostly affected by ρ_{zh_P} than anything else. This is in line with our intuition from section 2 that β_2 captures nature more than nurture, which pervades also into our general model. Also, note that the schooling correlation is explained by countervailing forces between λ and $\rho_{\epsilon h_P}$—neither of which are of primary interest in our goal of separating spillovers from ability transmission. When $\lambda = 0$, the ability level affect disappears, inducing high ability individuals (whose parents tend to have higher levels of schooling) to increase their length of schooling, which in turn increases correlation of schooling across generations. When $\rho_{\epsilon h_P} = 0$, children of high human capital parents (who tend to have higher levels of schooling) no longer have a desire to remain in school longer, and the correlation drops to being negative. The negativity comes from the early childhood effect: children with high z and h_P, all else equal, spend less time in school in the absence of the taste

Table 6: Counterfactual Effect of Spillover and Correlation Parameters.

	$\rho(S_P, S)$	β_2
Data	0.478	0.017
Model	0.458	0.017
Spillovers		
$\nu = 0$	0.458	0.013
$\lambda = 0$	0.625	0.017
$\nu = \lambda = 0$	0.644	0.009
Correlations		
$\rho_{zh_P} = 0$	0.425	0.008
$\rho_{\epsilon h_P} = 0$	-0.121	0.014
$\rho_{zh_P} = \rho_{\epsilon h_P} = 0$	-0.121	0.004

shock. Hence, in order to recover the positive relationship observed in the data, the $\rho_{\epsilon h_P}$ must be high.

Some caution is required when interpreting ρ_{zh_P}. The innate ability of the parent affects his or her human capital stock (higher ability parents invest more in themselves), in the same why that high ability children have high human capital. This can be represented by a correlation ρ_P. They also tend to have high ability children (the nature effect), which can be represented by a correlation paramater ρ_z. Then the structural interaction between nature and nurture is captured by the the correlation parameter $\rho_{zh_P} = \rho_z \rho_P$, i.e., the only way children's ability can be correlated with parental human capital is through the parent's own ability. To the extent that Proposition 2 suggests a high ρ_P and our structural model likely does not capture everything, ρ_{zh_P} becomes a lower-bound estimate of ρ_z, the share of children's human capital explained by solely by parental innate ability.

The correlation between parental human capital and own abilities, ρ_{zh_P}, is approximately 0.348. To understand what this means, suppose we were to assume that tastes are uncorrelated with abilities except through own human capital (as in equation (11)), and that the correlation structure of (h, z, ϵ) remains stable across generations. We can then obtain a sense of the magnitude of ρ_P by imposing ρ, the correlation between (h, z) of the children in our model, on the parents. Using lifetime income as a proxy for h, our model implies a correlation of $\rho = 0.603$ between the children's h and z. The implied correlation of abilities across generations in our model is then

$$\rho_z = \frac{\rho_{zh_P}}{\rho_P} \approx \frac{\rho_{zh_P}}{\rho} = 0.578.$$

As we have posited in section 2.3, the consideration of a parental level effect not only implies large spillovers (a large ν), but also a larger value of unobserved ability correlation across generations than what has been assumed in some previous studies (e.g., Restuccia and Urrutia (2004)).

Table 7: Aggregate Effect of 1 Year Increase in Mom's Schooling.

	spillover	ability	tastes	RF
Schooling	-0.099	-0.111	0.529	0.503
(log) Earnings	0.011	0.021	-0.001	0.009

"(log) Earnings" denotes the cross section average of the present discounted value of lifetime earnings, in logarithims. The first column holds abilities and tastes constant, while the next two columns let ability or taste also vary according to their estimated correlations with h_P. The column RF is when we allow for both selection on abilities and tastes. Schooling OLS in data and estimated model is 0.458 and 0.478, respectively.

We also argued in section 2.3 that Plug and Vijverberg (2003), by proxying for abilities by parental IQ, may understate the contribution of nurture. Hence, we proceed in the opposite direction. It is helpful to restate equation (10) along with the correlations (11) as

$$S = a_0 + a_1 \left[\frac{\rho_z}{\rho_P} (\log h_P - \eta_P) + \eta_z \right] + a_2 \log h_P = a_0 + \left[\frac{a_1 \rho_z}{\rho_P} + a_2 \right] \cdot h_P + \eta, \tag{16}$$

where (η_z, η_P, η) are i.i.d. shocks. Since the identification problem comes from the correlation structure, instead of varying parental z_P as in Plug and Vijverberg (2003), we can vary parental human capital h_P to isolate the spillover. (The identification problem in the presence of the taste shock can be interpreted in a similar fashion.) This can be done by decomposing the effect of a marginal increase in h_P while holding (z, ϵ) constant, and then letting (z, ϵ) vary as would be dictated by the correlations $(\rho_{zh_P}, \rho_{\epsilon h_P})$. This not only isolates the spillover but also contrasts it against ability and taste selection.

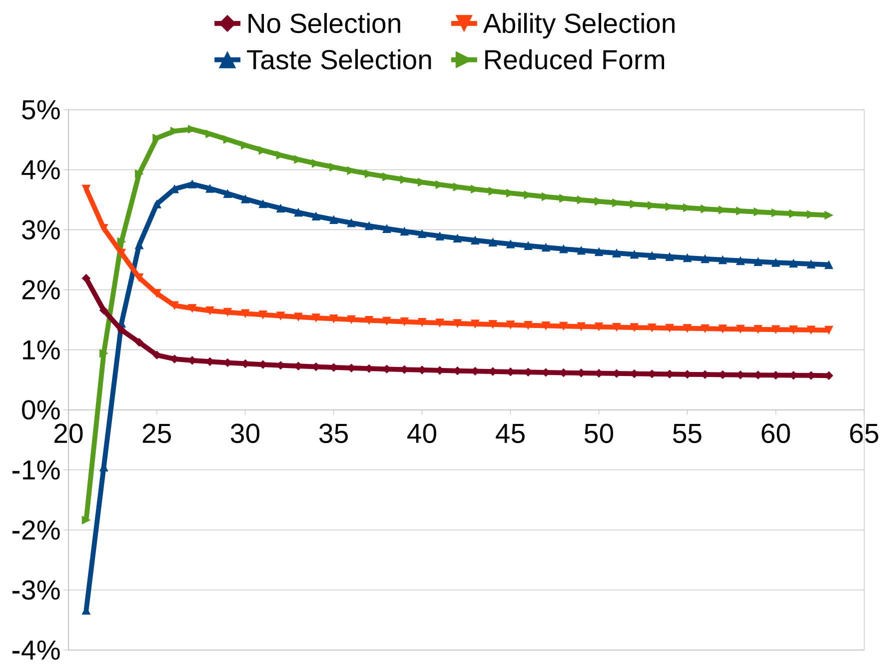
5.2 Decomposing the Parental Spillover

Given the individual state (h_P, z, ϵ) at age 6, we conduct the following experiment to isolate the schooling and earnings effects of spillovers. We increase S_P by 1 year, which translates into an increase of β units of h_P in logarithms, in four ways.[20] First, we hold constant the individual's (z, ϵ), which isolates the pure effect coming only from parental human capital. Next, we let either z or ϵ vary with h_P as dictated by the estimated correlations $(\rho_{zh_P}, \rho_{\epsilon h_P})$; this separately captures the selection effects from each. Finally, we let both z and ϵ vary together, which we label a "reduced-form" effect—i.e., this is what would be captured in an OLS regression not controlling for any selection on unobservables.

The average reponse is tabulated in Table 7. As expected, the spillover has a negative effect on schooling. What may be slightly surprising is that allowing for selection abilities has an even

[20]Since we assumed a statistical relationship between S_P and h_P in our estimation, we have also tried assuming a life-cycle structure for the parents, and forcing them to choose an additional year of schooling than would be optimal; this captures endogenous life-cycle effects since the increase in human capital from a forced additional year of schooling would be countervailed by cutting back on human capital accumulation later in the life. However, the resulting endogenous effects do not differ much from β, although there are small differences depending on the level of h_P.

Figure 3: Lifecycle Effect of 1 Year Increase in Mom's Schooling.



stronger negative effect: this means that in the cross-section, the ability level effect dominates its slope effect. As was implied from the previous counterfactual of setting $\rho_{\epsilon h_P} = 0$, only when we allow for selection on tastes do we see a positive effect of increasing parental schooling on child's schooling.

In contrast, the spillover has an approximately 1.1 percent positive effect on lifetime earnings, while abilities have a 1 percent effect. We conclude that independently of preference shocks, nature and nurture have a similar impact on lifetime earnings, with nurture playing a slightly larger role. From the table, it seems that tastes have a negative average impact: lifetime earnings drops by 1.2 percent once we allow for selection. This is because by inducing children who already earn above a certain threshold to stay in school longer, the higher earnings they earn later in life (through more human capital accumulation by staying in school) is dominated by the foregone earnings later in life. Such life-cycle effects are depicted in Figure 3.

However, this is merely an average effect. The cross-sectional impact of spillovers, abilities and tastes, in response to a one year increase in mom's schooling, is depicted in figure 4; the top panel is for schooling and the bottom panel for earnings. In both graphs, the sample is separated by levels of mom's schooling. The structural impact of spillovers is always negative on schooling and positive on earnings, taste selection always has a positive impact on schooling, and ability selection a positive impact on earnings. However, the effect of ability selection on schooling and taste selection on earnings depends on the level of mom's schooling, due to the two-dimensional heterogeneity between these two unobservables. Note that, high abilities has a negative effect on schooling for those individuals that the level effect dominates the slope effect, and high tastes for schooling may have a positive effect on earnings depending on whether the foregone earnings by increased schooling is dominated by future higher earnings, since by staying in school longer, individuals can also take advantage of the higher return human capital technology.

5.3 Counterfactual Schooling Reform

Having understood the marginal effects of increasing a parental schooling by 1 year, we now conduct a counterfactual experiment by imposing a minimum schooling requirement, which is intended to mimic compulsory schooling reforms that took place in many countries throughout the 20th century. A minimum schooling requirement has heterogeneous affects across the population since it only affects those parents who would otherwise not attain the required level of schooling, and even within this group, the additional number of years that is attained will vary. The goal of the exercise is to examine whether our intuition from above, that children's schooling responses are anemic despite increases in lifetime earnings, remains quantitatively valid, and also to gain a sense of the quantitative magnitudes such a reform may have (on life-cycle earnings).

The benchmark exercise is a simple one in which we simply impose a minimum 9 years of schooling for all parents, i.e. now the initial level of human capital in the economy is set to

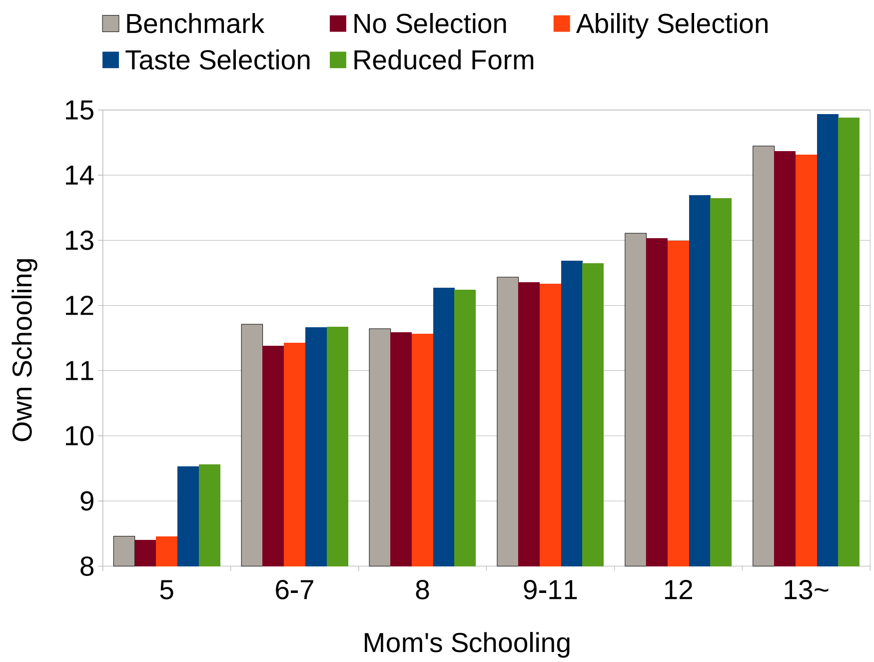
$$h_0 = bz^{\lambda} h_P^{\nu} = bz^{\lambda} \exp\left[\nu\beta \max\left\{S_P, 9\right\}\right].$$

One reason we choose 9 years as the hypothetical schooling requirement is because to make it comparable to the study by Black et al. (2005), who use the increase in the compulsory schooling requirement from 7 to 9 years in Norway in the 1960s. Of course, while the location and timing differs (the parents in our data would have been in school at turn of the 20th century), because the U.S. was a forerunner of public schooling (Goldin and Katz (2007)), the average years of schooling are quite similar in both settings (10.5 years vs 9.3 years; the medians are also similar). The reason we do not use, say, 8 years, is because in fact many states already had compulsory schooling requirements of 8 years in the U.S. at the time, so that imposing this requirement would only affect a very small number of parents in our sample.

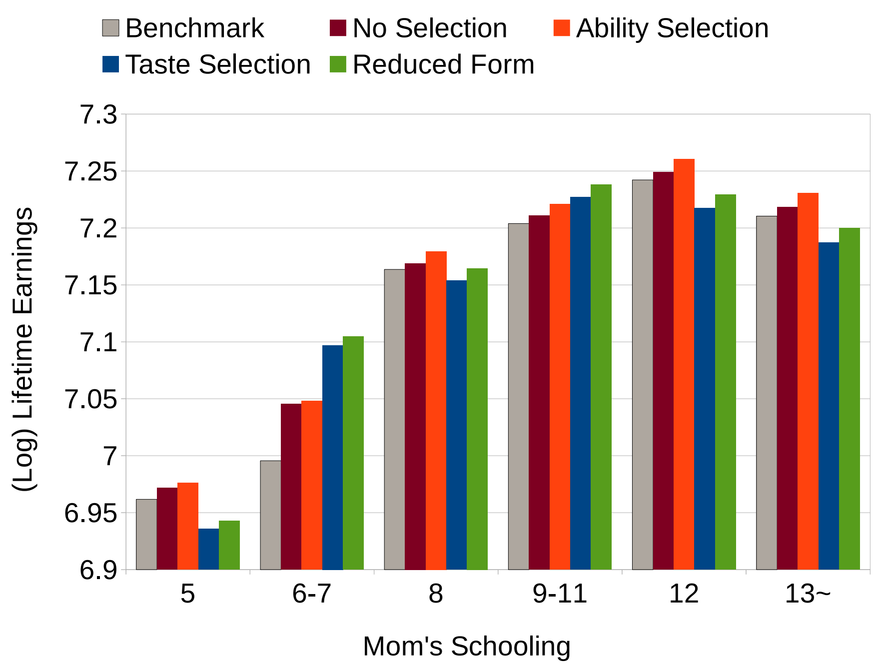
Specifically, we combine the simulated outcomes of two regimes: one without the minimum requirement (our benchmark model) and one imposing the requirement. Then we run both and OLS and IV regression on the merged simulated data, using the different regimes as an instrument. We repeat this exercise for four cases, as we did above: without controlling for any selection, and then controlling for selection on abilities and/or tastes. Table 8 shows the regression results for all cases.

Without allowing for selection on tastes, the IV coefficient is close to zero or even negative. Hence, a zero IV can be interpreted as the presence of a parental spillover, since a forceful increase in schooling is perhaps unlikely to change unobserved parental heterogeneity that is unrelated to economic outcomes, at least in the short run. However, the isolated spillover effect on lifetime earnings is 3.5%, in contrast to a 1.9–2.5% effect from ability selection. Hence, for children of low human capital parents, the spillover matters even more. Interestingly, taste selection also has a

Figure 4: Cross-section Effects of 1 Year Increase in Mom's Schooling



(a) Children's Schooling



(b) (log) Lifetime Earnings

Table 8: Counterfactual Schooling Reform

		spillover	ability	tastes	RF
Schooling	$a_{1,\text{OLS}}$	0.496	0.494	0.453	0.453
	$a_{1,\text{IV}}$	-0.090	-0.054	0.516	0.507
(log) Earnings		0.035	0.054	0.070	0.095

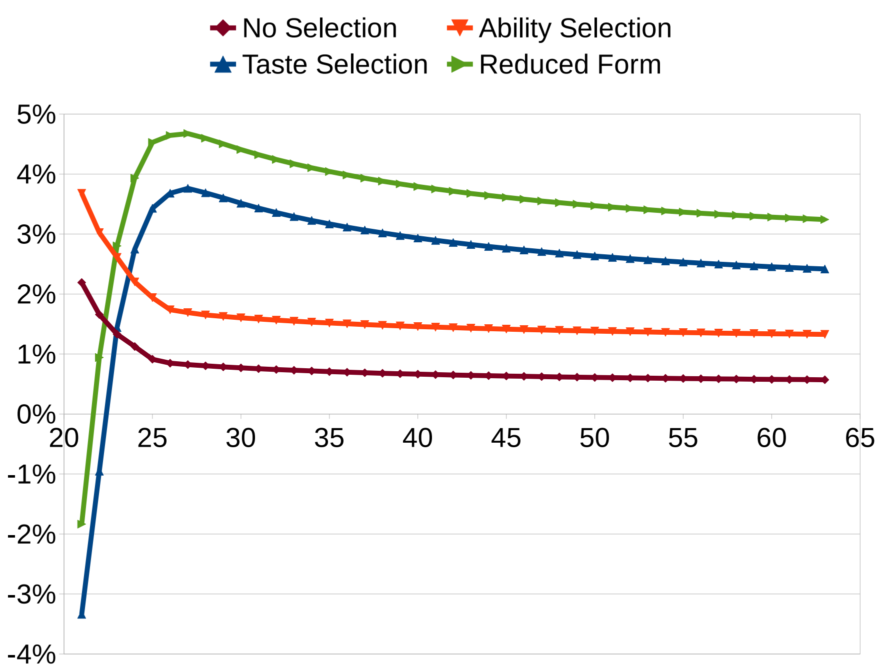
"(log) Earnings" denotes the cross section average of the present discounted value of lifetime earnings, in logarithims, for only those individuals affected by the reform. The first column holds abilities and tastes constant, while the next two columns let ability or taste also vary according to their estimated correlations with h_P. The column RF is when we allow for both selection on abilities and tastes. Schooling OLS in data and estimated model is 0.458 and 0.478, respectively. Schooling OLS in data and estimated model is 0.458 and 0.478, respectively.

large, positive effect on earnings (3.5%), in contrast to the negative effect it had on average when we increased parental schooling by one year. This also indicates that for children of low human capital parents, the future returns from additional schooling dominates the foregone returns from staying in school (which is perhaps expected, since these children have very low years of schooling to begin with.) The average life-cycle effect on earnings is depicted in figure 5.

6 Conclusion

COMING SOON.

Figure 5: Lifecycle Impact of Reform



Appendices

A Proofs to Propositions 1, 2, and Corollary 1.

The proof requires a complete characterization of the income maximization problem. While we can use standard methods to obtain the solution, we do this elsewhere and in what follows simply guess and verify the value function. For notational convenience, we drop the age argument a unless necessary. We separately characterize the solutions before and after the constraint $n \leq 1$ is binding in Lemmas 1 and 2. Then schooling time S is characterized as the solution to an optimal stopping time problem in Lemma 3. To this end, we further assume that

$$V(a,h) = q_2(a)h + C_W(a), \qquad \text{for } a \in [6+S, R),$$
$$V(a,h) = q_1(a) \cdot \frac{h^{1-\alpha_1}}{1-\alpha_1} + e^{-r(6+S-a)}C(S,h_S), \qquad \text{for } a \in [6, 6+S), \text{if } S > 0,$$

where

$$C(S,h_S) = q_2(6+S)h_S + C_W(6+S) - q_1(6+S) \cdot \frac{h_S^{1-\alpha_1}}{1-\alpha_1},$$

for which the length of schooling S and level of human capital at age $6+S$, h_S, are given, and C_W is some redundant function of age. Given the forms of $g(\cdot)$ and $f(\cdot)$, these are the appropriate guesses for the solution, and the transversality condition becomes $q(R) = 0$. Given the structure of the problem, we first characterize the working phase.

Lemma 1: Working Phase *Assume that the solution to the income maximization problem is such that $n(a) = 1$ for $a \leq 6+S$ for some $S \in [0, R-6)$. Then given $h(6+S) \equiv h_S$ and $q(R) = 0$, the solution satisfies, for $a \in [6+S, R)$,*

$$q_2(a) = \frac{w}{r} \cdot q(a) \tag{17}$$
$$m(a) = \alpha_2 \left[\kappa q(a) z\right]^{\frac{1}{1-\alpha}} \tag{18}$$
$$h(a) = h_S + \frac{r}{w} \cdot \left[\int_{6+S}^{a} q(x)^{\frac{\alpha}{1-\alpha}} dx\right] \cdot (\kappa z)^{\frac{1}{1-\alpha}} \tag{19}$$

and

$$\frac{wh(a)n(a)}{\alpha_1} = \frac{m(a)}{\alpha_2}, \tag{20}$$

where

$$
\begin{aligned}
q(a) &\equiv \left[1 - e^{-r(R-a)}\right] \\
\kappa &\equiv \frac{\alpha_1^{\alpha_1}\alpha_2^{\alpha_2}w^{1-\alpha_1}}{r}.
\end{aligned}
$$

Proof. Given that equation (2) holds at equality, dividing by (3) leads to equation (20), so once we know the optimal path of $h(a)$ and $m(a)$, $n(a)$ can be expressed explicitly. Plugging (2) and the guess for the value function into equation (4), we obtain the linear, non-homogeneous first order differential equation

$$\dot{q}_2(a) = rq_2(a) - w,$$

to which (17) is the solution. Using this result in (2)-(3) yields the solution for m, (18). Substituting (17), (18) and (20) into equation (5) trivially leads to (19). □

If $S = 0$ (which must be determined), the previous lemma gives the unique solution to the income maximization problem. If $S > 0$, what follows solves the rest of the problem, beginning with the next lemma describing the solution during the schooling period.

Lemma 2: Schooling Phase *Assume that the solution to the income maximization problem is such that $n(a) = 1$ for $a \in [6, 6+S)$ for some $S \in (0, R-6)$. Then given $h(6) = h_0$ and $q_1(6) = q_0$, the solution satisfies, for $a \in [6, 6+S)$,*

$$
\begin{aligned}
q_1(a) &= e^{r(a-6)}q_0 && (21)\\
m(a)^{1-\alpha_2} &= \alpha_2 e^{r(a-6)} \cdot q_0 z && (22)\\
h(a)^{1-\alpha_1} &= h_0^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2} \cdot \left[e^{\frac{\alpha_2 r(a-6)}{1-\alpha_2}} - 1\right] \cdot (\alpha_2 q_0)^{\frac{\alpha_2}{1-\alpha_2}} z^{\frac{1}{1-\alpha_2}}. && (23)
\end{aligned}
$$

Proof. Since $n(a) = 1$ during the schooling phase, using the guess for the value function in (4) we have

$$\dot{q}_1(a) = rq_1(a),$$

to which solution is (21). Then equation (22) follows directly from (3), and using this in (5) yields the first order ordinary differential equation

$$\dot{h}(a) = h(a)^{\alpha_1}\left[\alpha_2 q_1(a)\right]^{\frac{\alpha_2}{1-\alpha_2}} z^{\frac{1}{1-\alpha_2}},$$

to which (23) is the solution. □

The only two remaining unknowns in the problem are the age-dependent component of the value function at age 6, q_0, and human capital level at age $6+S$, h_S. This naturally pins down the length of the schooling phase, S. The solution is solved for as a standard stopping time problem.

LEMMA 3: VALUE MATCHING AND SMOOTH PASTING *Assume $S > 0$ is optimal. Then (q_0, h_S), are given by*

$$q_0 = \frac{e^{-rS}}{\alpha_2^{\alpha_2}} \cdot \left([\kappa q(6+S)]^{1-\alpha_2} z^{\alpha_1} \right)^{\frac{1}{1-\alpha}} \tag{24}$$

$$h_S = \frac{\alpha_1}{w} \cdot [\kappa q(6+S)z]^{\frac{1}{1-\alpha}}. \tag{25}$$

Proof. The value matching for this problem boils down to setting $n(6+S) = 1$ in the working phase, which yields (25).[21] The smooth pasting condition for this problem is

$$\lim_{a \uparrow 6+S} \frac{\partial V(a,h)}{\partial h} = \lim_{a \downarrow 6+S} \frac{\partial V(a,h)}{\partial h}.$$

Using the guesses for the value functions, we have

$$q_1(6+S) h_S^{-\alpha_1} = q_2(6+S) \qquad \Leftrightarrow \qquad h_S^{\alpha_1} = \frac{r}{w} \cdot \frac{e^{rS}}{q(6+S)} \cdot q_0,$$

and by replacing h_S with (25) we obtain (24). □

The fact that $h_S \propto z^{1/(1-\alpha)}$, along with the solution for $h(a)$ in the working phase in Lemma 1, proves Proposition 2. Then the solutions for $n(a)h(a)$ and $m(a)$ proves Corollary 1. We must still show Proposition 1.

Proof of Proposition 1. The length of the schooling period can be determined by plugging equations (24)-(25) into (23) evaluated at age $6+S$:

$$\begin{aligned} &\left(\frac{\alpha_1}{w} \cdot [\kappa q(6+S)z]^{\frac{1}{1-\alpha}} \right)^{1-\alpha_1} \\ &\quad \le h_0^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2^{1-\alpha_2}} \cdot \left(1 - e^{-\frac{\alpha_2 rS}{1-\alpha_2}}\right) \cdot \left([\kappa q(6+S)]^{\alpha_2} z^{1-\alpha_1} \right)^{\frac{1}{1-\alpha}}, \end{aligned}$$

with equality if $S > 0$. All this equation implies is that human capital accumulation must be positive in schooling, which is guaranteed by the law of motion for human capital. Rearranging terms,

$$h_0 \ge \frac{\alpha_1}{w} \cdot \left[1 - \frac{(1-\alpha_1)(1-\alpha_2)}{\alpha_1 \alpha_2} \cdot \frac{1 - e^{-\frac{\alpha_2 rS}{1-\alpha_2}}}{q(6+S)} \right]^{\frac{1}{1-\alpha_1}} \cdot [\kappa q(6+S)z]^{\frac{1}{1-\alpha}},$$

[21] This means that there are no jumps in the controls. When the controls may jump at age $6+S$, we need the entire value matching condition.

or now replacing $h_0 \equiv z^{\lambda} h_P^{\nu}$,

$$z^{1-\lambda(1-\alpha)} h_P^{-\nu(1-\alpha)} \leq F(S), \tag{26}$$

$$F(S)^{-1} \equiv \kappa \left(\frac{\alpha_1}{w}\right)^{1-\alpha} \cdot \left[1 - \frac{(1-\alpha_1)(1-\alpha_2)}{\alpha_1 \alpha_2} \cdot \frac{1 - e^{-\frac{\alpha_2 r S}{1-\alpha_2}}}{q(6+S)}\right]^{\frac{1-\alpha}{1-\alpha_1}} \cdot q(6+S)$$

which is the equation in the proposition. Define $\bar{S}$ as the solution to

$$\alpha_1 \alpha_2 q(6+\bar{S}) = (1-\alpha_1)(1-\alpha_2)\left(1 - e^{-\frac{\alpha_2 r \bar{S}}{1-\alpha_2}}\right),$$

i.e. the zero of the term in the square brackets. Clearly, $\bar{S} < R - 6$, $F'(S) > 0$ on $S \in [0, \bar{S})$, and $\lim_{S \to \bar{S}} F(S) = \infty$. An interior solution ($S > 0$) requires that

$$F(0) < z^{1-\lambda(1-\alpha)} h_P^{-\nu(1-\alpha)} \qquad \Leftrightarrow \qquad z^{1-\lambda(1-\alpha)} h_P^{-\nu(1-\alpha)} > \frac{r}{\alpha_1^{1-\alpha_2} (\alpha_2 w)^{\alpha_2} \cdot q(6)},$$

and S is determined by (26) at equality. The full solution is given by Lemmas 1-3 and we obtain Proposition 2 and Corollary 1. Otherwise $S = 0$ and the solution is given by Lemma 1. □

B Analytical Characterization of the Generalized Model

The solution to the schooling phase is identical to Lemma 2. In the working phase, there can potentially be a region where $n(a) = 1$ for $a \in 6 + [S, S+J)$, and $n(a) < 1$ for $a \in [6+S+J, R)$, so we can characterize the "full-time OJT" duration, J, following Appendix A. Although we normalize $w = 1$ in the estimation, we keep it here for analytical completeness.

Lemma 4: Working Phase, Generalized *Assume that the solution to the income maximization problem is such that $n(a) = 1$ for $a \in [6+S, 6+S+J)$ for some $J \in [0, R-6-S)$. Then given $h_S \equiv h(6+S)$, the value function for $a \in [6+S+J, R)$ can be written as*

$$V(a,h) = \frac{w}{r} \cdot q(a) h + D_W(a) \tag{27}$$

and the solution is characterized by

$$n(a)h(a) = \left[\frac{\alpha_W}{r} \cdot q(a) z\right]^{\frac{1}{1-\alpha_W}} \tag{28}$$

$$h(a) = h_J + \left(\frac{\alpha_W}{r}\right)^{\frac{\alpha_W}{1-\alpha_W}} \cdot \left[\int_{6+S+J}^{a} q(x)^{\frac{\alpha_W}{1-\alpha_W}} dx\right] \cdot z^{\frac{1}{1-\alpha_W}}, \tag{29}$$

where $h_J \equiv h(6+S+J)$ is the level of human capital upon ending full-time OJT. If $J = 0$, there is nothing further to consider. If $J > 0$, the value function in the full-time OJT phase, i.e. $a \in [6+S, 6+S+J)$ can *be written as*

$$V(a,h) = e^{r(a-6-S)} q_S \cdot \frac{h^{1-\alpha_W}}{1-\alpha_W} + e^{-r(6+S+J-a)} D(J,h_J) \tag{30}$$

where

$$D(J,h_J) = \frac{w}{r} \cdot q(6+S+J)h_J + D_W(6+S+J) - e^{rJ} q_S \cdot \frac{h_J^{1-\alpha_W}}{1-\alpha_W}$$

while human capital evolves as

$$h(a)^{1-\alpha_W} = h_S^{1-\alpha_W} + (1-\alpha_W)(a-6-S)z. \tag{31}$$

If $J > 0$, the age-dependent component of value function at age $6+S$, q_S, and age $6+S+J$ level of human capital, h_J, are determined by

$$q_S = we^{-rJ} \cdot \left[\frac{\alpha_W^{\alpha_W}}{r} \cdot q(6+S+J) z^{\alpha_W} \right]^{\frac{1}{1-\alpha_W}} \tag{32}$$

$$h_J = \left[\frac{\alpha_W}{r} \cdot q(6+S+J) z \right]^{\frac{1}{1-\alpha_W}}. \tag{33}$$

The previous Lemma follows from applying the proof in Appendix A. The solution for J is also obtained in a similar way we obtained S. Since human capital accumulation must be positive during the full-time OJT phase,

$$\frac{\alpha_W}{r} \cdot q(6+S+J)z \leq h_S^{1-\alpha_W} + (1-\alpha_W)Jz,$$

with equality if $J > 0$. Rearranging terms,

$$\frac{z}{h_S^{1-\alpha_W}} \leq G(J) \equiv \left[\frac{\alpha_W}{r} \cdot q(6+S+J) - (1-\alpha_W)J \right]^{-1}. \tag{34}$$

Define $\bar{J}$ as the zero to the term in the square brackets, then clearly $\bar{J} < R - S - 6$, $G'(J) > 0$ on $J \in [0, \bar{J})$, and $\lim_{J \to \bar{J}} G(J) = \infty$. Hence an interior solution $J > 0$ requires that

$$G(0) < \frac{z}{h_S^{1-\alpha_W}} \qquad \Leftrightarrow \qquad \frac{r}{\alpha_W q(6+S)} < \frac{z}{h_S^{1-\alpha_W}}, \tag{35}$$

and J is determined by (34) at equality. Otherwise $J = 0$.

We now consider the schooling phase and optimal schooling time. Since the schooling phase solution is identical to Lemma 2, and we only need consider new value matching and smooth pasting conditions.

LEMMA 5: SCHOOLING PHASE, GENERALIZED *The length of schooling, S, and level of human capital* *at age $6+S$, h_S, are determined by*

1. *if $J = 0$,*

$$\epsilon + (1-\alpha_2)\left[\frac{\alpha_2^{\alpha_2} w}{r}\cdot q(6+S) z h_S^{\alpha_1}\right]^{\frac{1}{1-\alpha_2}} = w\cdot\left(h_S + (1-\alpha_W)\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S)z\right]^{\frac{1}{1-\alpha_W}}\right) \tag{36}$$

$$h_S^{1-\alpha_1} \leq h_0^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2}\cdot\left(1-e^{-\frac{\alpha_2 r S}{1-\alpha_2}}\right)\cdot\left[\frac{\alpha_2 w}{r}\cdot q(6+S)h_S^{\alpha_1}\right]^{\frac{\alpha_2}{1-\alpha_2}} z^{\frac{1}{1-\alpha_2}} \tag{37}$$

with equality if $S > 0$. In an interior solution $S \in (0, R-6)$, the age-dependent component of the value function at age 6, q_0 is determined by

$$q_0 = \frac{we^{-rS}}{r}\cdot q(6+S)h_S^{\alpha_1}. \tag{38}$$

2. *if $J > 0$,*

$$\epsilon + (1-\alpha_2)\left(\alpha_2^{\alpha_2} w e^{-rJ}\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S+J)z\right]^{\frac{1}{1-\alpha_W}}\cdot h_S^{\alpha_1-\alpha_W}\right)^{\frac{1}{1-\alpha_2}}$$

$$= we^{-rJ}\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S+J)z\right]^{\frac{1}{1-\alpha_W}} \tag{39}$$

$$h_S^{1-\alpha_1} \leq h_0^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2}\cdot\left(1-e^{-\frac{\alpha_2 r S}{1-\alpha_2}}\right) \tag{40}$$

$$\cdot\left(\alpha_2 w e^{-rJ}\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S+J)z^{\alpha_W}\right]^{\frac{1}{1-\alpha_W}} h_S^{\alpha_1-\alpha_W}\right)^{\frac{\alpha_2}{1-\alpha_2}}\cdot z^{\frac{1}{1-\alpha_2}}$$

with equality if $S > 0$. In an interior solution $S \in (0, R-6)$, the age-dependent component of the value function at age 6, q_0 is determined by

$$q_0 = we^{-r(S+J)}\cdot\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S+J)z^{\alpha_W}\right]^{\frac{1}{1-\alpha_W}}\cdot h_S^{\alpha_1-\alpha_W}. \tag{41}$$

Proof. Suppose $S \in (0, R-6)$. The value matching and smooth pasting conditions when $J = 0$ are, respectively,

$$\epsilon - m(6+S) + e^{rS} q_0 z m(6+S)^{\alpha_2} = wh_S\left[1-n(6+S)\right] + \frac{w}{r}\cdot q(6+S)z\left[n(6+S)h_S\right]^{\alpha_W}$$

$$e^{rS} q_0 h_S^{-\alpha_1} = \frac{w}{r}\cdot q(6+S).$$

Hence (38) follows from the smooth pasting condition. Likewise, (36) follow from plugging $n(6+S)$, $m(6+S)$ from Lemmas 2 and 4 and q_0 from (38) in the value matching condition. Lastly, (37) merely states that the optimal h_S must be consistent with optimal accumulation in the schooling phase, $h(6+S)$.

The LHS of the value matching and smooth pasting conditions when $J > 0$ are identical to when $J = 0$, and only the RHS changes:

$$\epsilon - m(6+S) + e^{rS} q_0 z m(6+S)^{\alpha_2} = q_S z$$
$$e^{rS} q_0 h_S^{-\alpha_1} = q_S h_S^{-\alpha_W}.$$

Hence (41) follows from plugging q_S and h_S from (32)-(33) in the smooth pasting condition. Likewise, (39) follow from plugging $n(6+S) = 1$, $m(6+S)$ from Lemma 2, and q_0 from (41) in the value matching condition. Again, (40) requires consistency between h_S and $h(6+S)$. □

For each case where we assume $J = 0$ or $J > 0$, it must also be the case that condition (35) does not or does hold.

C Numerical Algorithm

Note that there is always a solution to (37) or (40)—i.e., we can always define a function $h_S(S)$ as a function of S. This is seen by you rearranging the equations as (bold-face for emphasis)

$$1 = \left(\frac{h_0}{\mathbf{h_S}}\right)^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2} \cdot \left(1 - e^{-\frac{\alpha_2 rS}{1-\alpha_2}}\right) \cdot \left[\frac{\alpha_2 w}{r} \cdot q(6+S)\right]^{\frac{\alpha_2}{1-\alpha_2}} z^{\frac{1}{1-\alpha_2}} \cdot \mathbf{h_S}^{-\frac{1-\alpha}{1-\alpha_2}} \tag{42}$$

$$1 = \left(\frac{h_0}{\mathbf{h_S}}\right)^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2} \cdot \left(1 - e^{-\frac{\alpha_2 rS}{1-\alpha_2}}\right) \cdot \left(\alpha_2 w e^{-rJ} \left[\frac{\alpha_W^{\alpha_W}}{r} \cdot q(6+S+J) z^{\alpha_W}\right]^{\frac{1}{1-\alpha_W}}\right)^{\frac{\alpha_2}{1-\alpha_2}} \cdot z^{\frac{1}{1-\alpha_2}} \cdot \mathbf{h_S}^{-\frac{1-\alpha+\alpha_2\alpha_W}{1-\alpha_2}}, \tag{43}$$

respectively. Hence, for any given value of S, both RHS's begin at or above 1 at $h_S = h_0$, goes to 0 as $h_S \to \infty$, and is strictly decreasing in h_S. The solution $h_S(S)$ to both (42) and (43) are such that

1. $h_S = h_0$ when $S = 0$ or $S + J = R - 6$
2. $h_S(S)$ is hump-shaped in S (i.e., there $\exists S$ s.t. h_S reaches a maximum).

Given a solution, we can solve equations (36) or (39), to which a solution may or may not exist depending on the parameters. Since we want to search over a large parameter region, we employ the following algorithm to handle any combination of possible solutions.

C.1 $S \in (0, R-6), J = 0$

1. Set $J = 0$. Get $h_S(S)$ from (42), this is well defined.

2. Now rearrange (36) as

$$\begin{aligned}\epsilon &= w \cdot \left(h_S(S) + (1-\alpha_W) \left[\frac{\alpha_W^{\alpha_W}}{r} \cdot q(6+S)z \right]^{\frac{1}{1-\alpha_W}} \right) - (1-\alpha_2) \left[\frac{\alpha_2^{\alpha_2} w}{r} \cdot q(6+S) z h_S(S)^{\alpha_1} \right]^{\frac{1}{1-\alpha_2}} \\ &\equiv T_0(S)\end{aligned}$$

If $\epsilon < T_0(0)$, it must be that $S = 0$. Otherwise, find the smallest solution to $\epsilon = T_0(S)$.

At this point, we may or may not have a candidate solution (including $S = 0$). If there is a candidate solution, check condition (35); if it is consistent with $J = 0$, save the solution as S_1.

C.2 $S \in (0, R-6), J > 0$

1. For any value of $J > 0$, $h_S(S; J)$ is well defined from (43).

2. This has to be consistent with (34):

$$h_S(S; J) = h_{SJ}(S, J) \equiv \left(z \cdot \left[\frac{\alpha_W}{r} \cdot q(6+S+J) - (1-\alpha_W) J \right] \right)^{\frac{1}{1-\alpha_W}}. \tag{44}$$

For given S, we can always find (h_S, J) from (34) if $h_{SJ}(S, 0) > h_0$. In fact, if $h_{SJ}(0,0) < h_0$, we can stop since J cannot be interior. Otherwise, find S s.t. $h_{SJ}(S, 0) = h_0$.

3. Now we can just repeat step C.1 above, but only up to the point that an interior solution for J exists. Rearrange (39) as

$$\begin{aligned}\epsilon = \quad & w e^{-rJ(S)} \left[\frac{\alpha_W^{\alpha_W}}{r} \cdot q(6+S+J(S))z \right]^{\frac{1}{1-\alpha_W}} \\ & - (1-\alpha_2) \left(\alpha_2^{\alpha_2} w e^{-rJ(S)} \left[\frac{\alpha_W^{\alpha_W}}{r} \cdot q(6+S+J(S))z \right]^{\frac{1}{1-\alpha_W}} \cdot h_S(S; J(S))^{\alpha_1 - \alpha_W} \right)^{\frac{1}{1-\alpha_2}} \\ \equiv T_1(S)\end{aligned}$$

If $\epsilon < T_1(0)$, it must be that $S = 0$. Otherwise, find the smallest solution to $\epsilon = T_0(S)$.

Again, we may or may not have a candidate solution. If there is a candidate solution, check condition (35); if it is consistent with $J > 0$, save the solution as S_2.

C.3 Final Solution

At this point, we are in one of three cases:

1. If there is only one solution, it is the solution.

2. No solution from either section: $S = R - 6$ is the solution.

3. Both have solutions: compare the two value functions at age 6 given the solutions $(S_1, J_1 = 0)$ and $(S_2, J_2 > 0)$ from steps C.1-C.2 using the fact that the function D_W in (27) can be written

$$\begin{aligned}&D_W(6+S+J)\\ =&w\left(\frac{\alpha_W}{r}\right)^{\frac{\alpha_W}{1-\alpha_W}}\left\{\int_{6+S+J}^{R} e^{-r(a-6-S-J)}\left[\int_{6+S+J}^{a} q(x)^{\frac{\alpha_W}{1-\alpha_W}}dx-\frac{\alpha_W}{r}\cdot q(a)^{\frac{1}{1-\alpha_W}}\right]da\right\}\cdot z^{\frac{1}{1-\alpha_W}}\end{aligned}$$

and

$$\begin{aligned}V(6,h_0) &= \int_6^{6+S} e^{-r(a-6)}\left[\epsilon - m(a)\right]da + e^{-rS}V(6+S,h_S)\\ &= \frac{1-e^{-rS}}{r}\cdot\epsilon - \frac{1-\alpha_2}{r\alpha_2}\cdot(\alpha_2 z q_0)^{\frac{1}{1-\alpha_2}}\left(e^{\frac{r\alpha_2 S}{1-\alpha_2}}-1\right)+e^{-rS}V(6+S,h_S).\end{aligned}$$

The candidate solution that yields the larger value is the solution.

D Tables

Table 9: Summary Statistics by Lifetime Earnings Decile

	Lifetime Earnings		Years of Schooling			Earnings growth	
	Mean	Median	Respondent	Mother	Father	Mean	Median
Lowest	409.45	439.14	10.36	8.19	7.62	3.99	1.01
Second	720.76	725.79	10.35	8.36	7.93	3.11	1.65
Third	888.49	889.85	11.19	8.62	7.93	3.57	2.11
Fourth	1013.96	1015.04	12.14	9.25	8.65	3.00	2.26
Middle	1127.57	1123.96	12.46	9.40	8.84	3.75	2.16
Sixth	1289.67	1288.40	12.58	9.59	9.24	4.06	2.22
Seventh	1490.83	1493.82	12.65	9.73	9.34	5.17	2.10
Eighth	1701.22	1697.35	12.72	10.05	9.39	5.33	2.21
Ninth	1934.17	1932.52	13.33	9.91	9.48	6.13	2.41
Highest	2562.82	2330.27	14.36	10.96	10.67	8.29	3.02

-Lifetime earnings is calculated as the sum of earnings from age 25 to age 55, where
-earnings in each age is adjusted to 2008 dollars and measured in \$1000.
-Earnings growth is measured as the ratio of average earnings between age 49 and
-age 51 to the average earnings between age 24 and age 26.

References

Becker, G. S. and N. Tomes (1986, July). Human capital and the rise and fall of families. *Journal of Labor Economics 4*(3), S1–39.

Behrman, J. R. and M. R. Rosenzweig (2002, March). Does increasing women's schooling raise the schooling of the next generation? *American Economic Review 92*(1), 323–334.

Behrman, J. R. and M. R. Rosenzweig (2005, December). Does increasing women's schooling raise the schooling of the next generation? reply. *American Economic Review 95*(5), 1745–1751.

Behrman, J. R. and P. Taubman (1989, December). Is Schooling "Mostly in the Genes"? Nature-N urture Decomposition Using Data on Relatives. *Journal of Political Economy 97*(6), 1425–46.

Behrman, J. R., P. Taubman, and T. J. Wales (1977). Controlling for and measuring the effects of genetics and family environment in equations for schooling and labor market success. In P. Taubman (Ed.), *Kinometrics: Determinants of Socioeconomic Success within and between Families*. North-Holland.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy 75*, 352.

Björklund, A., M. Lindahl, and E. Plug (2006). The origins of intergenerational associations: Lessons from swedish adoption data. *The Quarterly Journal of Economics 121*(3), 999–1028.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *The American Economic Review 95*(1), pp. 437–449.

Card, D. (1999). The causal effect of education on earnings. Volume 3A of *Handbook of Labor Economics*. Amsterdam: Elsevier.

Cunha, F. and J. J. Heckman (2007, May). The technology of skill formation. *American Economic Review 97*(2), 31–47.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, 05). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Del Boca, D., C. Monfardini, and C. Nicoletti (2012, March). Children's and parents' time-use choices and cognitive development during adolescence. Working Papers 2012-006, Human Capital and Economic Opportunity Working Group.

Goldin, C. and L. F. Katz (2007, March). The race between education and technology: The evolution of u.s. educational wage differentials, 1890 to 2005. NBER Working Papers 12984, National Bureau of Economic Research, Inc.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics 8*(S1), S85–S118.

Hall, P. and J. L. Horowitz (1996, July). Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators. *Econometrica 64*(4), 891–916.

Huggett, M., G. Ventura, and A. Yaron (2011, December). Sources of lifetime inequality. *American Economic Review 101*(7), 2923–54.

Kennan, J. (2006). A note on discrete approximations of continuous distributions.

Oreopoulos, P. and M. E. Page (2006, October). The Intergenerational Effects of Compulsory Schooling. *Journal of Labor Economics 24*(4), 729–760.

Plug, E. and W. Vijverberg (2003, June). Schooling, family background, and adoption: Is it nature or is it nurture? *Journal of Political Economy 111*(3), 611–641.

Restuccia, D. and C. Urrutia (2004). Intergenerational persistence of earnings: The role of early and college education. *American Economic Review 94*, 1354–1378.

Sacerdote, B. (2002). The nature and nurture of economic outcomes. *American Economic Review 92*(2), 344–348.

Sacerdote, B. (2011). Nature and nurture effects on children's outcomes: What have we learned from studies of twins and adoptees? Volume 1 of *Handbook of Social Economics*, pp. 1 – 30. North-Holland.

Scholz, J. K., A. Seshadri, and S. Khitatrakun (2006). Are americans saving "optimally" for retirement? *Journal of Political Economy 114*, 607–643.