The Causal Effect of Parents' Education on Children's Earnings*

Sang Yoon (Tim) Lee[1]

Queen Mary University of London, TSE, and CEPR, United Kingdom

Nicolas Roys

Royal Holloway, University of London, United Kingdom

Ananth Seshadri

University of Wisconsin–Madison and NBER, United States

Abstract

We develop and estimate a model of endogenous schooling and earnings to isolate the causal effect of parents' education on children's outcomes. Identification uses earnings differences among children with equal schooling whose parents have different schooling. Health and Retirement Study data indicate that the observed intergenerational schooling correlation is largely driven by correlated unobserved preferences for schooling. This is partly offset by a slightly negative structural relationship between parents' and children's schooling choices. Nevertheless, an exogenous one-year increase in parents' schooling raises children's lifetime earnings by 2 percent on average.

JEL Codes: E24, I24, I26, J24, J62

Keywords: parental spillovers; human capital; lifetime earnings; schooling choice

Running Head: Moms' Education and Sons' Earnings

*Manuscript submitted May 19, 2024; final revision submitted June 11, 2026.

[1]We thank Antonio Ciccone, Mariacristina De Nardi, Steven Durlauf, Eric French and Chris Taber for very helpful conversations. The paper also benefited from conference and seminar participants at the 2015 ASSA and EEA meetings, University of Cyprus, Essex, and TSE. Junjie Guo and Wei Song provided outstanding research assistance. Corresponding author. Email: sylee.tim@qmul.ac.uk.

1 Introduction

Parents have a large influence on their children's outcomes. Does this merely reflect selection—correlation in unobserved heterogeneity across generations? Or does it also partly reflect human capital spillovers from parent to child? In the former case, government subsidies aimed at improving education would only impact one generation. But in the presence of intergenerational spillovers, the returns from such public investments are reaped by all succeeding members of a dynasty, resulting in long-lasting effects.

To address these questions, we posit a standard life-cycle model of human capital accumulation à la Ben-Porath (1967, henceforth BP), which encompasses both schooling and learning on-the-job. An individual's length of schooling and earnings profile is endogenously determined by his initial level of human capital at age 6, and his learning ability, which governs the amount of human capital he can accumulate in a unit of time and remains constant throughout life. We extend this canonical model by assuming that an individual's initial level of human capital is itself a function of his learning ability and his parent's human capital. In addition, we assume that learning abilities are persistent over generations, and since parents' human capital is a result of parents' abilities, this results in children's learning abilities and their parents' human capital being correlated. Thus the effect of parents' human capital on children's outcomes from altering children's initial level of human capital is causal, while any effect through its correlation with children's learning abilities reflects selection.

The key result from our model is that an individual's schooling is a function of his learning ability and his parent's human capital, while earnings are a function of his learning ability and schooling. That is, parents' human capital will affect how long their children stay in school, but conditional on the length of schooling, have no direct effect on their children's earnings. This allows us to separately identify parental spillovers (the causal effect of parents' human capital on children's earnings) from selection by jointly using information on children's schooling and earnings, in addition to parents' schooling.

We argue that our model can rationalize findings from the previous literature that find a positive OLS coefficient when regressing children's years of schooling on their parents', but no evidence of a causal effect (e.g. Behrman and Rosenzweig 2002; Black et al. 2005). High human capital parents tend to have higher learning abilities, which are passed on to their children. Such children would then stay longer in school to reap the benefits. This generates a positive selection effect. But in BP, all else equal, individuals with higher levels of initial human capital spend less time in school since they can reach higher levels of human capital more quickly. So in our model, as long as higher human capital parents have children with higher levels of initial human capital, such children will spend less time in school—a somewhat surprising result. We refer to this as the "level effect": better initial conditions substitute the need for longer investments.

According to a simplified version of our model, the observed positive intergenerational schooling relationship indicates that selection effects are stronger than the level effect. The negative level effect can also help explain why IV estimates of intergenerational schooling find little to no effect. Note that the level effect is not a feature we artificially assume, but inherent in most life-cycle models of human capital accumulation that feature an initial level of human capital (e.g. Heckman et al. 1998; Huggett et al. 2011).

Importantly, this does not mean that the causal effect of parents' human capital on children's *earnings* (i.e., the spillover effect) is negative. According to BP, earnings are only a function of schooling and own abilities, which are unobserved. And our model implies that schooling is determined by abilities and parents' human capital. So among children who attain the same years of schooling, abilities have a one-to-one relationship with parents' human capital. This relationship allows us to identify the magnitude of the spillover from how parents' human capital affects the earnings of *children who attain the same level of schooling*.

Reduced-form evidence from the Health and Retirement Study (HRS) data reveals that children with the same years of schooling have parallel earnings profiles with constant gaps, which are increasing in their moms' years of schooling. The gaps indicate that moms affect how much human capital a child accumulates before labor market entry, compared to another child with the same years of schooling. The fact that these earnings profiles are parallel suggests that spillovers no longer operate once children enter the labor market. And, the size of the gaps is similar whether we look at children who do or don't finish high school, or children who do or don't enroll in college. All such evidence is consistent with our model assumptions.

We then estimate our general model by GMM. Given our identifying assumptions, we use moms' schooling as a proxy for parents' human capital, and allow children's human capital production technology to differ depending on whether he is in school or is working. We also allow for heterogeneous, non-pecuniary preferences for schooling.Including preference heterogeneity not only helps replicate key moments in the data, but it also prevents the overestimation of the effects of parents' human capital and children's learning abilities on children's schooling and earnings outcomes.

Preferences are identified by an exclusion restriction: abilities affect both schooling and earnings outcomes, whereas preferences exclusively affect schooling choices. Empirically, the reduced-form effect of moms' schooling on children's earnings is too small for the positive intergenerational schooling relationship to be driven solely by selection on abilities. Instead, our estimates suggest that a positive correlation between moms' schooling and preferences for sons' schooling is the main determinant of the positive intergenerational schooling relationship. But the level effect, while negative, is quantitatively small, so that it is not the case that moms with more years of schooling want to significantly reduce their children's schooling.

Finally, we find that forcefully increasing all moms' schooling by a year can raise the lifetime earnings of children by 2 percent, while only having limited effect on their schooling. This effect is heterogeneous across individuals and also over the life-cycle, and similar to our reduced-form estimate of 1.7 percent. We argue that this constitutes a lower-bound for parental spillovers, as the model only allows parents' human capital to affect children's initial level of human capital and suppresses other channels that may also be operational, such as parents' human capital having a causal impact on children's learning abilities. Thus our structural estimates can also be used in more stylized quantitative models with intergenerational links, such as in Adhami et al. (2023).[1]

Related literature A broad literature has studied the causal effect of parents on children's outcomes (Black and Devereux 2011; Holmlund et al. 2011; Sacerdote 2011).The common challenge for all these studies is to separately identify the unobserved correlation between parents' and their children's endowments (abilities and preferences independent of parents' observables) from the unobserved, causal impact of parental spillovers. Unfortunately, different approaches have led to

[1]While closely related, our analysis does not directly speak to the sources of intergenerational persistence. That said, our finding that parents' education is significant for children's earnings is consistent with Gayle et al. (2022), who decompose the sources of the observed intergenerational persistence of earnings.

a wide range of estimates (Black and Devereux 2011). We contribute to this literature by incorporating insights from a human capital model of education and earnings.[2]

Focusing only on the intergenerational transmission of education, there is some consensus that parents' schooling has little causal impact on children's schooling, with estimates often small, insignificant, or even negative (Holmlund et al. 2011). Some studies focus on teasing out the selection effect, such as Behrman and Taubman (1989) and Plug and Vijverberg (2003), who find strong, inherited genetic effects on children's schooling outcomes, and only limited room for causal effects. A larger literature has focused on measuring causal effects directly via means of an instrumental variable. Some solutions have been to employ special data on twin parents (Behrman and Rosenzweig 2002), adopted vs. biological children (Plug 2004) or to take advantage of compulsory schooling reforms during the parents' generation as a natural experiment (Black et al. 2005). Our structural model is explicit about the unobserved factors of learning abilities and preferences for schooling, and can rationalize the wide range of estimates found in the empirical literature.

Most of the existing literature attempts to control for selection on abilities, yet parental influences may also operate through non-pecuniary channels. Several studies have shown that pecuniary motives alone fall short of explaining education choices (Heckman et al. 2008) and non-pecuniary motives are estimated to be quite large in life-cycle models with schooling choice (Heckman et al. 1998). In our model, non-pecuniary motives are captured through heterogeneous preferences for schooling. While some of these preferences may be genetically determined, others are shaped by the environment in which the child grows up. This intergenerational transmission of preferences complicates the estimation of causal effects, but it also provides some discipline for interpreting results from special datasets.

Behrman and Rosenzweig (2002) find a negative to zero causal effect of moms' schooling on children's schooling, and conclude that the observed positive intergenerational schooling relationship must be driven by correlation between moms' schooling and unobserved factors.[3] Black et al.

[2]Our work differs from studies of early childhood human capital formation and its aggregate consequences, such as Lee and Seshadri (2019); Daruich (2026). These papers find that better early environments lead to higher human capital. While we find that higher human capital parents result in higher human capital children, our focus is on whether better early environments may also substitute for lengthier schooling.

[3]They also show some evidence that this is related to higher-educated moms spending more time in the labor market, as is also implied by our model results.

(2005) find that the causal schooling relationship is not significantly different from zero. Through the lenses of our model, twin moms who have different levels of schooling (as in Behrman and Rosenzweig 2002) likely differ less in abilities and preferences than adjacent cohorts before and after a schooling reform (as in Black et al. 2005). In particular, even though a reform is unlikely to cause a jump in abilities, it could cause a change in preferences. So in the case of twin moms, the child of the mom with more schooling may himself attain less (or at least not more) schooling, due to the level effect. But following a schooling reform, latter cohorts may develop a stronger preference for schooling that is passed on to their children, which can countervail the level effect.

More generally, while our model posits that the relationship between parents' human capital and preferences for children's schooling at least partly reflects selection, how much we can control for this endogeneity might be context-dependent. Specifically, an instrumental variable itself may directly affect preferences, or there could be an omitted variable that affects both the instrument and preferences. Thus, our findings can rationalize different findings in the literature by the degree of how any exogenous variation they exploit is related to preferences.[4]

These preferences may capture psychological or non-cognitive factors that induce a child to attain more or less education (Oreopoulos et al. 2008; Rege et al. 2011), and/or the fact that children from less advantaged families are more likely to be misinformed about education returns (Betts 1996; Avery and Turner 2012; Hoxby and Avery 2013). Recent research differentiates how cognitive and non-cognitive skills formed early in life can explain various measures of well-being in adulthood (Cunha et al. 2010), and the childhood environment has long been suspected as what may explain the large estimates for non-pecuniary motives found in structural models of earnings (Bowles et al. 2001; Heckman et al. 2006). In our context, we simply lump such factors together into heterogeneous preferences for schooling.

As another example, while adoptee studies find a causal effect of parents' schooling that is lower than the observed level of intergenerational schooling persistence, they still remain slightly positive (Sacerdote 2002) or are close to zero (Plug 2004). Children adopted into different families are likely to develop non-cognitive skills or acquire information that shapes their preference for

[4]IV estimates may also reflect differences in parents' selection into treatment. However, our quantitative exercises only consider a forceful 1-year increase for all parents, or a compulsory schooling reform that mandates a minimum level of schooling for all parents. We thank an anonymous referee for pointing this out.

schooling. If these unobservables are positively correlated with parental education, one would expect the estimated effect of parents' schooling on children's schooling to be larger in adoptee studies than in twin studies. Similarly, the schooling of moms and dads may also have different effects on the schooling of daughters and sons due to differences in how preferences are transmitted.[5] Regardless, the main message from our model is that the intergenerational effects on children's schooling and earnings need not be the same.

Like schooling, earnings also exhibit strong intergenerational persistence (Solon 1999; Chetty et al. 2014). But in contrast to the empirical literature on schooling, most studies find that non-genetic factors do have a positive causal effect on children's earnings (Björklund et al. 2006; Sacerdote 2007). The fact that parental causal effects are small or non-existent for children's schooling but not for their earnings is sometimes viewed as an inconsistency (Black and Devereux 2011). The vast majority of the literature exclusively focuses on only the schooling-to-schooling, or earnings-to-earnings (or income) relationship. Some do discuss whether the effect of parents' earnings on children's schooling is similar to its effect on children's earnings (e.g. Chang et al. 2023), but little is known on the effect of *parents' schooling on children's earnings*.[6]

Our data do not allow us to differentiate between parents' education and earnings, so we proxy their human capital and abilities using data from the earlier HRS AHEAD cohort. We then focus on differentiating how they affect *their children's education and earnings* in a structural model in which both are jointly determined, rather than focusing on single outcomes one-by-one.[7] By considering two sets of outcome variables, we are able to isolate the size of intergenerational spillovers. In our model, parents can have opposing effects on their children's schooling and earnings outcomes. Even though the level effect is negative, it is estimated to be quantitatively small, and the positive relationship between parents' and children's schooling is overall driven by

[5] Our model is only based on the mom-son relationship. Fernández et al. (2004) and subsequent studies find evidence that supports a causal transmission of attitudes across generations between mothers and sons, but none focus on exogenous changes in parents' education.

[6] The only paper we know that investigates the effect of *parents' education on children's earnings or incomes* is Connolly et al. (2022). They find a positively significant effect, but do not control for endogeneity or selection and do not make any claims of causality.

[7] In this sense, our study is also related to Bowles and Gintis (2002), which estimates how much of the intergenerational persistence in earnings can be explained by the correlation between dads' earnings and other variables, such as children's education.

preferences. But higher initial human capital still implies higher earnings, leading us to conclude that parents have a positive causal effect on children's earnings.

The rest of the paper is organized as follows. Section 2 posits a model of human capital accumulation with spillovers, of which we solve a simpler version to derive analytical results and make empirical predictions. In Section 3 we describe the HRS data that we use and compare reduced-form evidence against the simple model's predictions. Section 4 estimates the full structural model to the HRS. We show, quantitatively, that the estimated model inherits properties of the simpler version. Section 5 examines the main result of increasing moms' schooling by 1 year, and also the counterfactual result of a hypothetical compulsory schooling reform. Section 6 concludes.

2 Schooling and Earnings with Parental Spillovers

A parent of age B has a child at time 0. At this point, she (the parent) can be described by a state vector $(y_i(0); h_{Pi}, z_{Pi})$, representing her initial wealth, human capital, and an ability that is stochastically transmitted to her child.[8] While the parent's wealth will evolve over time based on her savings and child-investment decisions, both her human capital and ability remains fixed throughout her lifetime.[9]

We impose an intergenerational borrowing constraint: parents cannot borrow against their children's future earnings. However, the parent only faces the natural borrowing limit based on their own lifetime earnings. Consequently, y represents the present discounted value of the parent's total potential future earnings plus initial assets.[10] Let θ describe the parent's altruism for the child, and individual utility is CRRA with parameter χ, i.e. $u(c) = c^{1-\chi}/(1-\chi)$. When the

[8] In our empirical analysis, we assume that the parent is the mom and the child is a son, for reasons explained below.

[9] This assumption about human capital reflects the constraint that while the mom may continue working during child-rearing, she lacks the time to invest in her own human capital accumulation. For an analysis that incorporates parental human capital accumulation alongside child investment decisions, see Lee and Seshadri (2019). Since this is not a major tenet of our analysis here, we maintain the constant human capital assumption for analytical tractability and computational feasibility.

[10] Although this abstracts from period-to-period credit constraints, a large literature emphasizes the role of long-term family factors over short-term borrowing constraints in determining children's educational outcomes (e.g., Cameron and Heckman 1998, 2001; Cameron and Taber 2004), especially in the pre-2000s period. For a more detailed discussion, see footnote 26 in Section 4.1.

child is attached to the parent, the pair's instantaneous utility is given by:

$$U(c_{Pi}, c_i) = (1-\theta)u(c_{Pi}) + \theta u(c_i) = qu(C_i), \quad q \equiv \left[(1-\theta)^{1/\chi} + \theta^{1/\chi}\right]^{\chi}, \quad C_i \equiv c_{Pi} + c_i,$$

where (c_{Pi}, c_i) is the consumption of the parent and child and q is interpreted as an adult consumption equivalent scale.

The child's own learning ability, z_i, is initially unknown and only revealed at age 6. The realization of z_i is i.i.d. conditional on the parent's ability z_{Pi}.At child age 0, the parent solves the following optimization problem (dropping family i subscripts),

$$\text{(1a)} \quad W_0(y(0); h_P, z_P) = \max_{\{C(a)\}} \left\{ \int_0^6 e^{-\rho a} qu(C(a)) + e^{-6\rho} \mathbb{E} W_6(y(6); h_P, z) \right\}$$

where ρ is the subjective discount rate, $C(a)$ is the pair's consumption when the child is age a, and the expectation is taken over the child's abilities z, subject to

$$\text{(1b)} \quad \dot{y}(a) = ry(a) - C(a)$$

where r is the interest rate. We assume the parent dedicates all her time to childcare during the child's early years, earning no labor income.

Starting at age 6, the parent makes educational decisions for the child, including how long he remains in school before becoming independent. Assuming that opportunities to end schooling occur at discrete intervals Δ, she faces a dynamic discrete choice problem:

$$\text{(2)} \quad W_a(y(a), h(a); h_P, z) = \max\{J_a(y(a), h(a); h_P, z), I_a(y(a), h(a); z)\}$$

where $h(a)$ is the child's human capital at age a, and (J, I) are the values of continuing and stopping school, respectively. The value of continuing at age a_0 is

$$\begin{aligned} & J_{a_0}(y(a_0), h(a_0); h_P, z) \\ \text{(3a)} \quad & = \max_{\{C(a_0), n(a_0), m(a_0)\}} \left\{ \int_{a_0}^{a_1} e^{-\rho(a-a_0)} qu(C(a)) da + e^{-\rho\Delta} \mathbb{E} W_{a_1}(y(a_1), h(a_1); h_P, z) \right\} \end{aligned}$$

where $a_1 \equiv a_0 + \Delta$ and the expectation is taken over a preference shock that is only realized in the future, subject to

$$\text{(3b)} \quad \dot{y}(a) = ry(a) - w_P h_P n_P(a) - m(a) - C(a),$$

$$\text{(3c)} \quad \dot{h}(a) = zh(a)^{\alpha_1} m(a)^{\alpha_2(a)} [n_P(a) h_P]^{\alpha_3(a)}, \qquad n_P(a) \in [0,1], \qquad m(a) > 0$$

$$\text{(3d)} \quad h(6) = \phi z_i^{\lambda} h_{Pi}^{\nu}.$$

where (3b) describes the evolution of parental wealth, while (3c) governs the accumulation of the child's human capital. Human capital is produced using parental time and goods inputs, (n_P, m). Because $y(0)$ includes the parent's potential lifetime earnings, the cost of time devoted to children is her forgone labor income, $w_P h_P n_P(a)$. Both this opportunity cost and the cost of goods reduce parental wealth in (3b). Finally, (3d) specifies the child's initial stock of human capital at age six. The parameter ϕ is common across families and scales parents' human capital into units of the children's initial human capital.

The exponents $(\alpha_1, \alpha_2, \alpha_3)$ govern the child's self-productivity and the returns to goods (m) and parental time investments (n_P), respectively. While the coefficients on parental time n_P and parental human capital h_P are constant, note that we allow (α_2, α_3) to vary with the child's age, enabling both the relative importance of inputs and their complementarity to change over time.

The causal effect of parents' human capital on children's earnings comes from (3d): the child's initial level of human capital at age 6 depends on h_P. This relationship captures how parents' human capital affects the quality of time they spend with young children, with the magnitude of this effect determined by the parameter ν. Initial human capital may also depend on the child's own learning ability z, the extent of which is measured by λ.

The value of attaining S years of schooling and finishing is given by:

$$\text{(4)} \qquad I_{S'}(y(S'), h(S'); h_P, z) = \max_{b \geq 0} \left\{ (1-\theta) M_{S'}(y(S') - b) + \theta V_{S'}(b, h(S'); z) \right\} + \xi(S; h_P, z)$$

where $S' \equiv S + 6$ is the child's age upon independence, M is the parent's value after the child's independence, V is the child's value, and b is an intergenerational transfer. The intergenerational constraint $b \geq 0$ rules out the parent borrowing against any of the child's future earnings. Preferences for S years of schooling, ξ, is allowed to vary with (h_P, z). Since the parent is altruistic, these preferences may originate from the parent, the child, or both, reflecting only their desire for the child to continue (or discontinue) schooling without directly affecting the child's future earnings.

A parent whose child becomes independent at age S', left with assets y_0, solves a consumption-savings problem until death at age T:

$$M_{S'}(y_0) = \max_{\{c_P(a)\}} \int_0^{T-S'-B} e^{-\rho a} u(c_P(a)) da,$$

where $S' + B$ is the parent's age when the child becomes independent (upon finishing schooling), subject to $y(0) = y_0$, $y(T - S' - B) \geq 0$, and

$$\dot{y}(a) = ry(a) - c_P(a).$$

Once independent, the child solves a standard Ben-Porath model of human capital accumulation given assets b and human capital h_0, retiring at age R and dying age T:

$$\text{(5a)} \quad V_{S'}(b, h_0; z) = \max_{\{c(a), n(a), m(a)\}} \int_0^{T-S'} e^{-\rho a} u(c(a)) da$$

subject to $h(0) = h_0$, $y(0) = b$, $y(T - S') \geq 0$ and

$$\dot{y}(a) = ry(a) + wh(a)(1 - n(a)) - m(a) - c(a),$$

$$\text{(5b)} \quad \dot{h}(a) = \begin{cases} z[n(a)h(a)]^{\alpha_{1W}} m(a)^{\alpha_{2W}}, & n(a) \in [0, 1],\ m(a) > 0, \quad \text{for } a \in [0, R - S'), \\ 0 & \text{otherwise} \end{cases}$$

where w is the wage earned by the child's generation and the exponents $(\alpha_{1W}, \alpha_{2W})$ represents the on-the-job human capital returns to time and good investments $[n(a), m(a)]$, which the child now decides independently. The coefficients on time n and human capital h are constrained to be equal, consistent with the empirical evidence surveyed in Browning et al. (1999).[11]

The structure of the model is depicted in Figure 1, which shows how the parental state vari-

[11] For instance, Heckman et al. (1998) estimate this technology using NLSY79 data and cannot reject this restriction across persons of different ability and education groups.

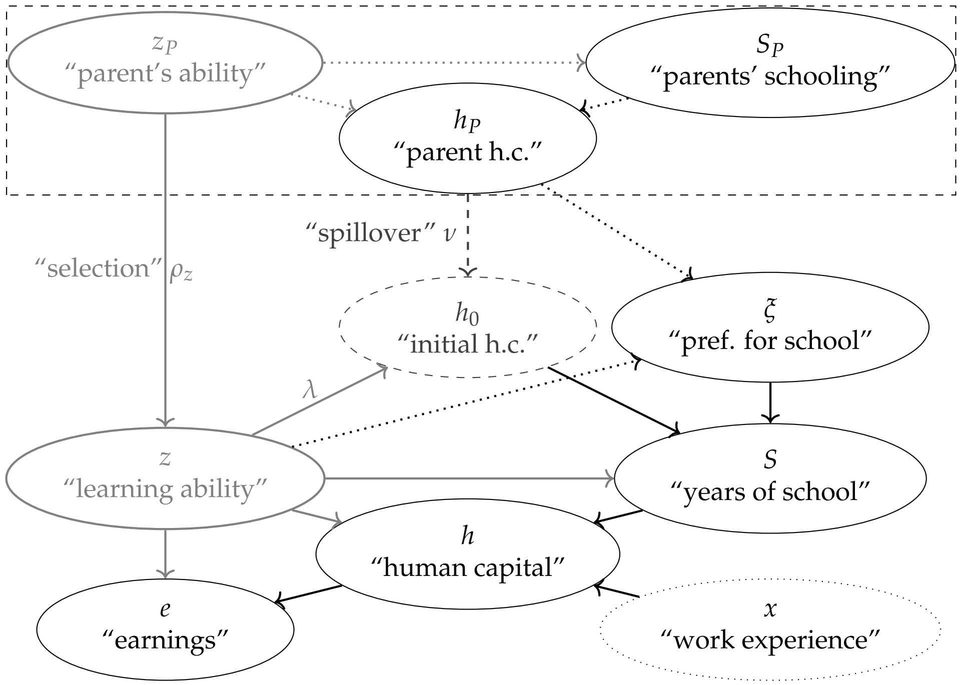


Figure 1: Structure of the Model.
The parent's state is summarized inside the dashed rectangle. Selection effects are depicted in solid gray, and spillovers in dashed dark gray. Initial human capital h_0 affects earnings directly, but also through its effect on schooling. In contrast, preferences for schooling ξ only affects schooling and has no direct effect on earnings.

ables in the upper panel affect children's schooling and earnings outcomes in the lower panel. The arrows depict the direction of the model's mechanisms.

2.1 Initial Conditions

Unlike the canonical BP model, we assume that an altruistic parent makes decisions for the child during childhood, taking into account both his later adulthood value and her own old-age. Our model incorporates several additional departures: We allow for non-pecuniary preferences for schooling (equation 4), and the returns to human capital investments can vary with the child's age and differ between schooling and working phases (equations 3c and 5b). Most importantly for our analysis, we assume that individuals' initial human capital depends on both their own ability and their parents' human capital (equation 3d).

The main objective of this paper is to understand how exogenously increasing parental schooling S_P affects children's schooling and earnings outcomes. To this end, we set the parent's human capital state h_P as follows. Ignoring assets, (3) implies that the child's human capital upon completing school, h_S, is a function of $(S, z; h_P)$. Abstracting from grandparents, the parent's human capital h_P must therefore be a function of (S_P, z_P). Furthermore, (2) implies that S_P itself is also a function of z_P. These relationships are depicted in the upper panel of Figure 1.

Since S_P is a function of z_P and z depends only on z_P, we can derive the best linear predictors:

$$\log z_P = \log \tilde{z}_P + \beta_{z_P S_P} S_P$$

$$\text{(6)} \qquad \log z = \log \tilde{z} + \beta_z \log z_P = \underbrace{\log \tilde{z} + \beta_z \log \tilde{z}_P}_{\equiv Z} + \underbrace{\beta_z \beta_{z_P S_P}}_{\equiv \beta_{z S_P}} S_P,$$

where $(\log \tilde{z}, Z)$ are uncorrelated with S_P. We can also project $\log h_P$ on S_P and Z (ignoring its mean) as

$$\text{(7)} \qquad \log h_P \approx \beta_{h_P S_P} S_P + \beta_{h_P Z} Z.$$

so that $\beta_{h_P Z}$ captures only the direct effect of z_P on $\log h_P$, excluding the indirect effect operating through S_P. The effect of S_P on h_P that can be exogenously affected by policy is captured by $\beta_{h_P S_P}$, i.e., the policy-relevant effect. To be clear, parental ability z_P and its transmission to z are both assumed to be purely exogenous and unaffected by policy interventions.

We can now derive the theoretical OLS coefficient of projecting $\log h_0$ on S_P as:

$$\text{(8)} \qquad \frac{\text{Cov}(\log h_0, S_P)}{\text{Var}(S_P)} = \frac{\text{Cov}(\nu \log h_P + \lambda \log z, S_P)}{\text{Var}(S_P)} = \nu \cdot \beta_{h_P S_P} + \lambda \cdot \beta_{z S_P}.$$

Thus the causal effect of parents' schooling on children's human capital is captured by $\beta_{h_P S_P}$ through the parameter ν, and we define the spillover as *the increase in earnings caused by this transmission*.[12] In contrast, the selection term $\beta_{z S_P}$ is only non-zero because S_P depends on the parent's learning ability z_P (equation 7), which is exogenously transmitted to her child as z, capturing selection—that is, a change in the joint distribution of (z_P, S_P). This is depicted as a gray dashed line in Figure 1. The parameter λ governs the strength of this selection effect on early human capital formation.[13]

[12] An increase in children's earnings would be caused not only by $\beta_{h_P S_P}$ but also by higher h_P- parents investing more time and money in their children.

[13] Our setup assumes away any causal effect that parental human capital h_P might have on z, but such an effect

In the estimation, we take (h_P, z_P) by moms' schooling S_P directly from the data, so $\beta_{h_P S_P}$ is observed. However, since abilities z are unobserved, we face the challenge of separately identifying (ν, λ) from β_{zS_P}, i.e., how much of the correlation between a child's schooling or earnings and S_P stems from a causal effect on h_0 versus indirect effects through selection. A primary objective of our structural estimation is to sort these out.

Preferences and Selection Similarly, we need to identify the relationship between (h_P, z) and preferences for schooling $\xi(S; h_P, z)$. Unlike abilities, we do not assume that such preferences are transmitted across generations but rather that they may originate from the parent, the child, or both, since the decision-making parent can influence the child's choices.

We do not explicitly model how such preferences are formed. In reality, an exogenous shift in parents' schooling may or may not affect preferences for children's schooling. If preferences are correlated with parental human capital but fixed across the population, policies that alter parents' schooling may have no effect on children's schooling choices. Conversely, if the correlation reflects causation, such policies would have long-lasting intergenerational effects on educational attainment. Since the reality likely lies somewhere between these extremes, in our counterfactual experiments we consider both cases. Importantly, in either scenario, we highlight that the intergenerational schooling effect of parental human capital is distinct from the intergenerational earnings effect.

In the remainder of this section, we demonstrate that a simplified model without preferences for schooling provides insights into the precise mechanisms at work, and shows how the spillover from ν can be separately identified from (λ, β_{zS_P}) using a panel of individuals' schooling and earnings, and their parents' schooling. In the quantitative analysis, beginning in Section 4, we reintroduce preferences for schooling. These preferences generate effects on children's schooling similar to selection on abilities but exert opposite effects on earnings.

cannot be separately identified from β_{select}. Thus our estimate would provide a lower bound of the spillover, which nevertheless turns out to be significantly positive. We discuss this further in Sections 2.3 and 4.3.

2.2 Simple Model Without Preferences

To derive simple analytical expressions and compare them against the data, we first solve a simplified version of the model. To this end, we abstract from any preferences for schooling and the parent's own welfare concerns by imposing $\theta = 1$ and $\xi(S; h_P, z) = \alpha_3 = 0$. Since any borrowing against the child's future earnings would only be invested in the child, we also ignore the intergenerational borrowing constraint. When the child reaches age 6, the parent's problem reduces to maximizing the child's lifetime income, after which the child can perfectly smooth consumption. Finally, we impose $\alpha_1 = \alpha_{1W}$ and $\alpha_2(a) = \alpha_{2W}$ so that the human capital production technology is identical in the schooling and working periods, making the model equivalent to BP, with the sole addition that initial human capital is a function of z and h_P:

$$\text{(9a)}\quad V(h_P, z) = \max_{\{n(a), m(a)\}} \left\{ \int_6^R e^{-r(a-6)} \left[wh(a) \left[1 - n(a)\right] - m(a) \right] da \right\}$$

subject to (3d) and

$$\text{(9b)}\quad \dot{h}(a) = z[n(a)h(a)]^{\alpha_1} m(a)^{\alpha_2}, \qquad n(a) \in [0, 1], \qquad m(a) > 0.$$

This is an optimal stopping time problem where $n = 1$ until $6 + S$. Since the objective function is linear, the constraint set strictly convex, and the law of motion strictly positive and concave (as long as $\alpha_1 + \alpha_2 < 1$), the optimization problem is well-defined and the solution is unique (Léonard and Van Long 1992). The Hamilton-Jacobi-Bellman (HJB) equation is

$$rV(a, h) - \frac{\partial V(a, h)}{\partial a} = \max_{n, m} \left\{ wh(1 - n) - m + \frac{\partial V(a, h)}{\partial h} \cdot z(nh)^{\alpha_1} m^{\alpha_2} \right\}.$$

where $V(a, h)$ denotes the present discounted sum of net income at age a. As usual, the HJB equation can be interpreted as a no-arbitrage condition. The left-hand side is the instantaneous cost of having a human capital level of h at age a, while the right-hand side is the instantaneous return. The first order conditions for the controls are

(10a)

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

(10b)

$$m = \alpha_2 z(nh)^{\alpha_1} m^{\alpha_2} \cdot V_h,$$

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

$$\text{(11)} \quad r \cdot V_h - V_{ah} = w(1-n) + \frac{\alpha_1 z(nh)^{\alpha_1} m^{\alpha_2}}{h} \cdot V_h + z(nh)^{\alpha_1} m^{\alpha_2} \cdot V_{hh},$$

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 equal the marginal return. Equations (10)-(11) along with the law of motion (9b), initial condition (3d) and terminal condition $V_h = 0$—the appropriate transversality condition for a fixed terminal time problem—characterize the complete solution. Here we only present the important results, all proofs are relegated to Online Appendix A.1. To save on notation, it is useful to define $\alpha \equiv \alpha_1 + \alpha_2$ and

$$q(a) \equiv 1 - e^{-r(R-a)}, \qquad \kappa \equiv \frac{\alpha_1^{\alpha_1} \alpha_2^{\alpha_2} w^{1-\alpha_1}}{r}.$$

PROPOSITION 1: OPTIMAL SCHOOLING CHOICE *Define the function $F(s)$ as*

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

which is only a function of prices (w, r) and the parameters (α_1, α_2). The optimal choice of schooling S is uniquely determined by

$$\text{(12)} \quad F'(S) > 0, \qquad F(S) \geq z^{1-\lambda(1-\alpha)} h_P^{-\nu(1-\alpha)} \quad \textit{with equality if } S > 0,$$

so F implicitly defines S as a function of (h_P, z).

Proof. See Online Appendix A.1. □

The higher the learning ability z of an individual, the higher his optimal choice of schooling (as long as $\lambda(1-\alpha) < 1$). Intuitively, conditional on an initial level of human capital, higher z individuals benefit more from schooling. But the causal effect from the parent's human capital h_P on schooling is *negative* (as long as $\nu(1-\alpha) > 0$). In our model, this level effect—substitution between the initial level of human capital and length of schooling—arises as long as lifetimes are finite and the returns to human capital investments are decreasing ($\alpha < 1$): Individuals with high initial values face low returns to schooling earlier, reducing their incentive to stay in school and instead enter the labor market earlier.

PROPOSITION 2: POST-SCHOOLING HUMAN CAPITAL AND EARNINGS PROFILES *For an individual who attains S years of schooling, human capital at the end of schooling is*

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

Earnings at experience $x \in [0, R-6-S)$ *is*

$$(13) \quad e(x; S, z) = wh(x)\left[1 - n(x)\right] = C(x; S) \cdot z^{\frac{1}{1-\alpha}},$$

where

$$(14) \quad C(x; s) = r\kappa^{\frac{1}{1-\alpha}} \left\{ \int_0^x q(6+s+\hat{x})^{\frac{\alpha}{1-\alpha}} \left[1 + \frac{\alpha_1 e^{-r(R-6-s-\hat{x})}}{1-\alpha}\right] d\hat{x} \right\}.$$

Proof. See Online Appendix A.1. □

The above proposition tells us that, conditional on a child's length of schooling, the human capital level of a child during his working years is determined only by his learning ability z. His initial stock of human capital, h_0, has no effect on the *amount* of human capital accumulated (quality) except through the *length* of schooling (quantity), S. So both the causal effect of ν and early childhood learning λ are subsumed in the length of schooling.[14]

[14] All else equal, raising h_0 (quality) would reduce S (quantity), diluting its effect on earnings (the level effect). The reason why h_0 has *exactly* zero effect on earnings is because we are imposing $(\alpha_1, \alpha_2) = (\alpha_{1W}, \alpha_{2W})$, which is the standard BP model (see Lemma 3 in the Appendix). The richer model we estimate allows the production function to change before and after S, so earnings do depend on h_0, but we verify that this does not affect the qualitative implications derived from Proposition 2.

Equation (13) can be interpreted as a type of Mincer equation that relates earnings to schooling. Conditional on ability z, the slope of individuals' log-earnings experience profiles vary by length of schooling S according to the function C. Conversely, for individuals with the same S, experience profiles must be parallel across different z's, which is standard in models that use BP. But while Proposition 2 does rely on the fact that S subsumes the effect of initial human capital, it does not exploit the structural relationship that S is a function of (h_P, z) in Proposition 1. In the next subsection we show that this is key: Both propositions must be utilized for us to isolate the spillover effect.

2.3 Identification of Spillovers

We will now demonstrate how to identify spillovers given a data sample of children's schooling and earnings outcomes, and the schooling levels of their parents. In addition, we will also argue that a positive intergenerational schooling relationship, a robust finding in empirical studies, implies that the spillover effect is dominated by other factors such as selection on abilities. However, quantifying its exact magnitude is less straightforward, which motivates the structural model in Section 4. There, we also find that preferences rather than selection on abilities is the dominating force, while the identification scheme for spillovers carries over. Nonetheless, we also provide a reduced-form approximation of ability selection in Online Appendix A.2. The online appendix also proposes a reduced-form identification scheme for λ.

To isolate the spillover effect, we use Proposition 1 to relate children's schooling to the parents'. Using the projections in (6)-(7) and applying them to (12) at equality (ignoring constants),

(15a)

$$\begin{aligned}\log F(S_i) &= [1-\lambda(1-\alpha)]\log z_i - (1-\alpha)\nu \log h_{Pi}\\ &\approx [1-\lambda(1-\alpha)]\left(\beta_{zS_P} S_{Pi} + Z_i\right) - (1-\alpha)(\nu\beta_{h_P S_P} S_{Pi} + \nu\beta_{h_P Z} Z_i)\end{aligned}$$

(15b)

$$\frac{\log F(S_i)}{(1-\alpha)[1-\lambda(1-\alpha)]} \approx \Big[\underbrace{\frac{\beta_{zS_P}}{1-\alpha}}_{\equiv b_{\text{select}}} - \underbrace{\frac{\nu\beta_{h_P S_P}}{1-\lambda(1-\alpha)}}_{\equiv b_{\text{spill}}}\Big]\cdot S_{Pi} + \Big[\frac{1}{1-\alpha} - \frac{\nu\beta_{h_P Z}}{1-\lambda(1-\alpha)}\Big]\cdot Z_i.$$

where the function F is an increasing function of S (Proposition 1) and Z_i is uncorrelated with S_{Pi}.

This shows that the intergenerational schooling coefficient is positively related to selection β_{zS_P} but negatively to the spillover parameter ν under reasonable parameter values of (λ, α). Therefore a positive coefficient would indicate that the selection effect outweighs the spillover effect.

Next, we jointly exploit the structural relationships in (12) and (13) to separately identify spillover effects. In what follows, suppose we have a random sample of individuals whose schooling levels of their parents and themselves, (S_{Pi}, S_i), and potential experience $x_i \equiv a_i - 6 - S_i$, are observed without error. But the earnings of an individual i with x years of experience, $e_{i,x}$, is only observed with error, so that $\log e^*_{i,x} = \log e_{i,x} + u_{i,x}$ where $u_{i,x}$ is i.i.d. across all (i, x).

Corollary 1: Identifying Spillovers. *Suppose that $(S_i, x) > 0$ for all i. If we regress*

$$(16) \quad \log e^*_{i,x} = [\varphi_{S_i,x} + b_{S_i} S_{Pi}] + \tilde{u}_{i,x}.$$

where $\{\varphi_{s,x}\}$ is a full set of schooling$\times$experience effects, and $\tilde{u}_{i,x}$ regression error, the estimates

$$\hat{b}_s = \bar{b} = b_{spill} = \beta_{h_P S_P} \nu / \left[1 - \lambda(1-\alpha)\right].$$

Proof. Proposition 1 implies that, among children with $S_i = \hat{S} > 0$, it must hold that

$$z_i = F(\hat{S})^{\frac{1}{1-\lambda(1-\alpha)}} \cdot h_{Pi}^{\frac{\nu(1-\alpha)}{1-\lambda(1-\alpha)}} \approx F(\hat{S})^{\frac{1}{1-\lambda(1-\alpha)}} \cdot \exp(\beta_{h_P S_P} S_{Pi} + \beta_{h_P Z} Z_{S_i = \hat{S}})^{\frac{\nu(1-\alpha)}{1-\lambda(1-\alpha)}}$$

where the approximation follows from (7). Applying this and $S_i = \hat{S}$ in (13) yields

$$(17) \quad \log e_{i,x}(\hat{S}) = \underbrace{\log C(x;\hat{S}) + \frac{\log F(\hat{S})}{(1-\alpha)[1-\lambda(1-\alpha)]} + \frac{\nu \beta_{h_P Z} Z_{S_i = \hat{S}}}{1-\lambda(1-\alpha)}}_{\hat{\varphi}_{\hat{S},x}} + \frac{\nu \beta_{h_P S_P}}{1-\lambda(1-\alpha)} \cdot S_{Pi}$$

among individuals with $S_i = \hat{S}$, since Z_i's are uncorrelated with S_{Pi}. Since the term in brackets in (16) apply only to individuals with the same level of schooling, the coefficients b_s are estimated as claimed, and the regression error only captures measurement error. $\square$

Corollary 1 shows that holding own schooling levels fixed, the spillover is identified by the within-group variation in log-earnings and parents' schooling. Moreover, it implies that among children with the same length of schooling, differences in parents' schooling manifests itself as constant gaps across log-earnings profiles, which reveals the spillover.[15]

Why does the coefficient on parents' schooling in (16) reveal spillovers? According to Proposition 2, conditional on S and potential experience, children's earnings are only a function of z. While S itself is endogenous to z (equation 15a), holding S fixed controls for the selection effect, which is exactly what is being done in (16). Then the regression explicitly reveals the structural relationship between (h_P, z) *within* each schooling subgroup given in (15a), rather than the population (h_P, z) relationship. Our model also implies that within-subgroup relationships should be the same across groups, which we empirically verify in the next section.

We stress that while the propositions and corollaries of course depend on our model, the identification scheme is more general than that. The two identifying assumptions are that $S = S(S_P, z)$: Children's schooling is a function of their abilities and parents' schooling; and $e = e(x; S, z) = e[x; S(S_P, z), z]$: earnings are a function of their abilities and only their *own* schooling. That is, parents directly influence children's schooling but not their earnings. The earnings effect is only indirect through how children's earnings are affected by children's schooling.

In the next section, we run the proposed regressions in Corollary 1 using HRS data. We will show that earnings profiles are close to parallel with gaps determined by own and parents' schooling, and present some raw evidence on the magnitude of b_{spill}. Then in Section 4, we estimate a generalized model that includes unobserved heterogeneity in preferences for schooling.

3 Data Analysis

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. It is a national panel study with an initial sample (in 1992) of 12,652 persons in 7,702 households that 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 sample that we use for our 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 were conducted every two years

[15] Corollaries 1-2 in Online Appendix A.2 show how b_{select} and λ can be identified in reduced-form.

since 1992. As the HRS matured, new cohorts have been added.

The HRS is usually used to study elderly Americans close to or in retirement, but there are several features that make it suitable for our purposes (compared to NLSY79, for example). First, these older individuals and their parents were less affected by compulsory schooling regulations and other government interventions, and more than half never advanced to college. This makes the sample suitable for our model in which large schooling variations are important for identifying causal effects.[16] In particular, we can directly exploit information on the length of schooling as implied by our model, rather than imputing educational attainment based on whether individuals gain certain qualifications (such as a college degree). Second, the education premium was quite stable prior to the 1980s, so it is unlikely for these cohorts to have been surprised by an unexpected rise in education returns. It is also less likely that the effect of education on earnings outcomes or the reverse effect of expected earnings on education choices changes much from cohort to cohort. Third, the HRS contains information on own schooling, schooling of both parents, and can be augmented with restricted Social Security earnings data through which we observe entire life-cycle earnings histories.

3.1 Descriptive Statistics

For the purposes of our study, we keep 5,760 male respondents born between 1924 and 1941 from the 1992 sample.[17] We further drop 646 individuals with missing information on their own education or mom's years of schooling. This leaves us with 5,114 individuals. Table 1 describes this sample by level of education. Average schooling and mom's schooling are about 12.3 and 9.2 years, respectively, both with a standard deviation of approximately 3.5 years, although both are top-coded at 17 years in the data.

A large fraction of HRS respondents allowed researchers restricted access to their Social Security earnings records. Combined with self-reported earnings in the HRS, these earnings records

[16] The HRS displays much more schooling variation than other datasets. Many papers studying intergenerational schooling relationships also focus on earlier periods, but lack life-cycle earnings information (except for Black et al. (2005), which uses Norwegian administrative data, although they do not use that information in their main analysis).

[17] We focus on men as most women from this sample have only very short earnings histories. The initial HRS sample selected individuals born in 1931-1941, but spouses born in different years were also included in the sample.

	HSD	HSG	SMC	CLG	Total
Schooling	7.97	12.00	13.88	16.54	12.33
	(2.67)		(0.68)	(0.50)	(3.41)
Mom's Schooling	7.00	9.26	10.11	11.07	9.24
	(3.70)	(2.98)	(3.11)	(3.27)	(3.60)
Dad's Schooling	6.41	8.72	9.88	11.02	8.91
	(3.70)	(3.36)	(3.59)	(3.77)	(3.96)
% White	73.31	84.70	85.00	89.30	82.81
% Black	21.94	13.24	12.23	6.62	13.82
% Hispanic	20.36	4.98	5.86	2.72	8.67
Earnings 23-27	16.69	23.39	22.20	18.34	20.24
	(12.23)	(13.23)	(12.77)	(12.37)	(13.00)
Earnings 28-32	24.96	34.78	34.13	33.42	31.76
	(16.36)	(17.92)	(18.60)	(19.56)	(18.50)
Earnings 33-37	30.90	41.46	41.68	44.15	39.33
	(18.78)	(21.04)	(21.72)	(23.64)	(21.85)
Earnings 37-42	34.38	44.27	45.92	52.17	43.79
	(21.21)	(24.16)	(28.16)	(31.88)	(26.96)
# Obs	1349	1647	940	1178	5114

*HSD<12, HSG=12, 12<SMC<16, CLG=16+ years of schooling

**Years of schooling top-coded at 17.

***Standard deviations in parentheses.

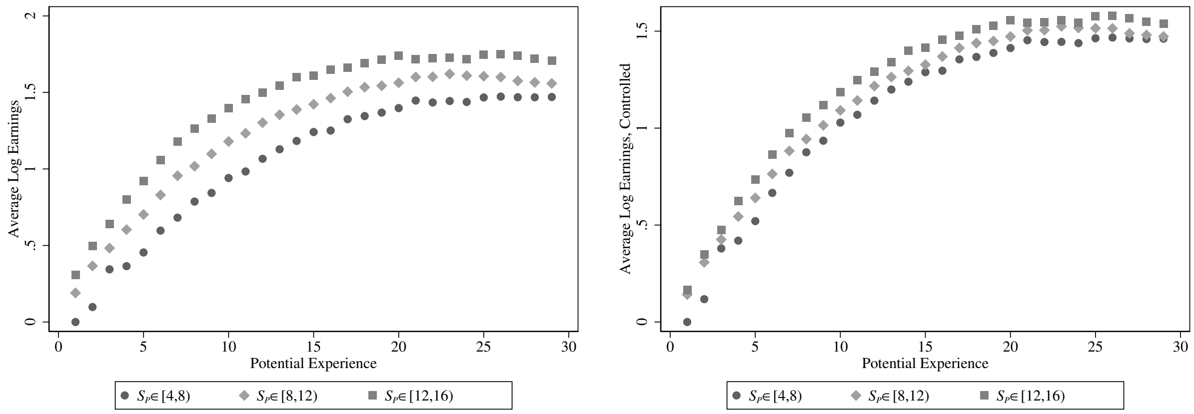
****Earnings inflated to 2008, measured in $1000.

Table 1: Summary Statistics by Education

provide almost the entire history of earnings for most of the HRS respondents. Social security earnings records exceeding the maximum level subject to social security taxes were top-coded in the years 1951 through 1977, so we impute them assuming

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

where $e^*_{i,t}$ 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, as well as an unanticipated white noise component $u_{i,t}$. We employed random-effect assumptions with homoskedastic errors to estimate the above model separately for men



(a) Children of moms with [4,8), [8,12), and [12,16) years of schooling.

(b) Children of moms with [4,8), [8,12), and [12,16) years of schooling, controlling for own schooling-specific experience effects.

Figure 2: Experience-Earnings Profiles by Mothers' Schooling

Earnings profiles of children by different levels of moms' schooling: 1924-1941 birth cohort. *y*-axis: average log annual earnings in 2008 USD, zero and missing earnings are dropped. Mothers' schooling levels are divided into 4 to 7 years, 8 to 11 years, and 12 or more years. In each subfigure, the lowest level of average log earnings is normalized to 0.

with and without a college degree. Scholz et al. (2006) gives details of the above earnings model, the procedure used to impute top-coded earnings, and the resulting coefficient estimates.

3.2 Visualizing Spillovers

According to our model, the spillover is subsumed in the choice of schooling (Proposition 1). For individuals with the same level of schooling, earnings profiles should be parallel with a constant gap. Spillovers are identified from the within group variation in parents' schooling among individuals with identical levels of schooling (Corollary 1).

In Figure 2(a), we divide individuals into 3 subsamples, depending on whether their moms attained $S_P \in [4,8)$, $[8,12)$, or $[12,16)$ years of schooling. These correspond approximately to primary, secondary, and tertiary education. Average earnings profiles are close to parallel by mom's schooling, suggesting that potential experience effects do not vary much across schooling levels.[18,19] In Figure 2(b), we control for own schooling-specific experience effects, corresponding to the estimates for $\hat{\varphi}_{s,x}$ in Corollary 1. Since schooling is correlated across generations, the gaps narrow but remain close to parallel: In the model, the slope of earnings profiles is determined by $\log C(x; S)$, so controlling for S, the slopes should be identical. Online Appendix Figure A.2, which plots the gaps between the profiles, shows that parallelism holds from 5 to 25 years of potential experience, only failing at very low and high experiences (especially for low S_P's).

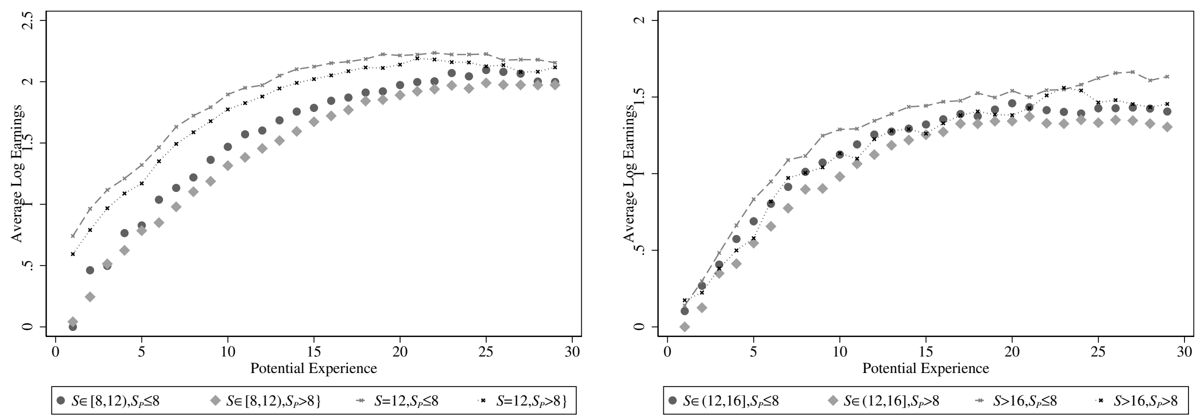
The corollary also tells us that the gap between these controlled profiles is the spillover effect *if* it is constant across subsamples of children with the same levels of schooling. So in Figure 3, we compare earnings profiles across moms' schooling levels by groups with similar levels of own schooling. Individuals are categorized into 4 subsamples depending on whether they have $S \in [8, 12)$, exactly 12, $(12, 16]$, or more than 16 years of schooling. Each group is further divided according to whether the mom has more than 8 years of schooling—corresponding to the end of junior high in most states in 1900 U.S.—and is about half of the moms in our sample. Figure 3(a) depicts the average log-earnings profiles of individuals with $[8, 12)$ and those with exactly 12 years of schooling, and Figure 3(b) individuals with $(12, 16]$ and those with 17 or more years of schooling, respectively.

At all four education levels, the average log-earnings profiles of children with the same schooling but different mom's schooling are nearly parallel with a constant gap, as predicted by Corollary 1.[20] This points to a permanent parental effect that persists throughout an individual's career, which is precisely what the spillover intends to capture. Moreover, the magnitude of the gap is similar across the four categories of children's education, as implied by Corollary 1: To be precise, the gaps should capture $b_s = \bar{b} = b_{\text{spill}} = \beta_{h_P S_P} \cdot \nu / [1 - \lambda(1 - \alpha)]$ in the corollary.

[18] For robustness, we have tried dividing children and their moms according to different levels of education. Experience-earnings profiles are near-parallel except when we split mom's schooling into very fine categories with few observations. We also confirmed that this evidence is present in available data from the NLSY79 and PSID. Controlling for cohort and time effects has little effect as well: time effects in particular absorb some of the experience effects, but do not change the relative position of the profiles. Details available upon request.

[19] It could also mean that own schooling levels are evenly distributed across moms' schooling levels, but this is not the case in the data as shown in Table 5.

[20] Online Appendix Figure A.3 plots the gaps between the profiles in Figure 3. Just like the average profiles, parallelism fails at low and high experiences for schooling-specific profiles, but holds from 5 to 25 years of potential experience.



(a) Children with [8,12) vs. 12 years of schooling.

(b) Children with (12,16] vs. 17+ years of schooling.

Figure 3: Identifying Spillovers

Earnings profiles of children of different schooling levels by moms' schooling: 1924-1941 birth cohort. y-axis: average log annual earnings in 2008 USD, zero and missing earnings are dropped. Mothers' schooling levels are divided into 8 years or below, and more than 8 years. In each subfigure, the lowest level of average log earnings is normalized to 0.

Given this visual evidence, we next obtain the reduced-form magnitudes of the spillover effect b_{spill} presented in Figures 2-3. To capture the spillover, we run different versions of a Mincerian regression:

$$(18) \quad \log e_{i,x} = f(S_i, x) + bS_{Pi} + \epsilon_{i,x}$$

where $f(\cdot)$ is a function of both schooling and potential experience which we specify in various different ways below, and $\epsilon_{i,x}$ an error term. We estimate different versions of (18) for earnings data from ages 23 to 42, and tabulate the results in Table 2.[21,22] In Table 3, we also allow the estimates of b to vary by own-schooling, S_i.

We consider four measures for S_{Pi}: the mom's and dad's years of schooling, respectively, their sum, and also including both as separate controls. Our theory does not speak to which of these

[21] Neither the estimates nor earlier figures change much depending on whether we use all available earnings records or restrict individuals' age. We restrict ourselves to this age interval because this is what we estimate the model to in the next section. We begin at age 23 because many earnings records are missing prior to this age, and end at age 42 because it is close to the peak of the earnings profiles for most individuals and our model has nothing to say about the irregular labor supply or retirement behavior at older ages.

[22] The estimates are also robust to the inclusion of race or cohort dummies. In our structural estimation to follow, such effects should be absorbed in the unobserved heterogeneity in preferences for schooling, so controlling for them here would make the reduced-form estimates incomparable with the structural estimates.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)*	(8)*
Mom S_P		0.017	0.017			0.015	0.017	0.017
		(22.71)	(21.82)			(14.98)	(21.87)	(21.89)
Dad S_P				0.011		0.003		
				(15.03)		(3.53)		
Mom + Dad					0.008			
					(20.05)			
S	0.090	0.082					0.190	0.228
	(99.73)	(84.21)					(17.19)	(6.29)
x effects	✓	✓	✓	✓	✓	✓		
S effects			✓	✓	✓	✓		
(S, x) effects			✓	✓	✓	✓		
3rd-order							✓	
4th-order								✓
R^2	0.190	0.194	0.204	0.203	0.205	0.205	0.201	0.203

Table 2: Mincer Regressions

OLS regressions of (log) earnings on own and parents' years of schooling. HRS initial cohort, 5,114 individuals males born 1924-1941, ages 23-42. Columns (1)-(6) include a full set of potential experience (age-6-S) effects, and columns (3)-(6) also include a full set of own schooling$\times$experience effects. Columns (7)-(8) instead include 3rd and 4th order polynomials of (x, S). t-stats shown in parentheses.
$*$The coefficients shown for S is conditional on $S = 12$ years of schooling.

measures is more appropriate. However, many studies argue that moms have a larger influence on children (Behrman and Rosenzweig 2002; Del Boca et al. 2017), which is also true in our sample as shown in Table 2.

Column (1) is a standard Mincer regression with a linear coefficient for own schooling, and a full set of potential experience effects (but not interacted). The return to schooling is estimated to be 9 percent. 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 in the last century (Goldin and Katz 2008). The returns slightly decrease to 8.2 percent when we include moms' years of schooling in column (2). The coefficient on moms' education is 1.7 percent and statistically significant. This suggests that being born to a mom with five additional years of schooling has about the same effect on earnings as would an additional year of own schooling.

Column (3) corresponds exactly to the regression in Corollary 1, which includes a full set of own schooling-specific experience effects. When we replace moms' schooling with dads' schooling in (4), the parent's coefficient is attenuated to 1.1 percent but still statistically significant. The parents' coefficient drops further when we let S_{Pi} equal the sum of the schooling of both parents

	$S < 8$	$S \in [8, 12)$	$S = 12$	$S \in (12, 16]$	$S > 16$	W.Avg
Mom S_P	0.017	0.021	0.022	0.020	0.016	0.020
	(6.40)	(12.21)	(16.40)	(13.93)	(7.09)	
x effects	o	o	o	o	o	
R^2	0.076	0.121	0.151	0.182	0.256	
Sample	6,116	16,961	29,727	26,256	10,833	

Table 3: Mincer Regressions by Own Schooling

OLS regressions of (log) earnings on moms' years of schooling. HRS initial cohort, males born 1924-1941, ages 23-42. All columns include a full set of potential experience (age-6-S) effects. t-stats shown in parentheses. The sample size is the number of earnings observations. The last column is the population weighted average of the coefficients.

in (5). If we include both separately as in (6), the coefficient on moms is slightly reduced from 1.7 percent to 1.5 percent while the coefficient on dads drops significantly from 1.1 percent to 0.3 percent. This implies that moms' schooling has dominant explanatory power, which we take as the proxy for parents' human capital in what follows. In columns (7)-(8), we drop all schooling and experience effects and instead control for schooling and experience using 3rd- and 4th-order polynomials.[23] The coefficient on S_P remains virtually the same regardless of how we control for experience.

The striking feature is that no matter how we control for own schooling and experience, the coefficients on mom's S_P are virtually identical and highly significant across all specifications. The coefficient is stable even when we re-ran all these regressions controlling for race and cohort dummies, and also for white males separately.

To summarize, moms' schooling has a stronger relationship with sons' earnings than dads', and the estimated effect $\hat{b}$ is about 1.7 percent, which is a measure of the gaps between the three lines in Figure 2. This is in the range of previous empirical work (Card 1999), and according to our model (Corollaries 1) captures a causal spillover effect to the extent they are identical across subgroups of children with similar levels of schooling.

In Table 3, we group individuals into schooling intervals rather than exact years (S_i less than 8, $\in [8, 12)$, exactly 12, $\in (12, 16]$, or more than 16), and for each group, regress earnings only on moms' schooling, controlling for experience effects.[24] The coefficient is significant across all groups at approximately 2 percent, although slightly lower for children with either very high or

[23]The exact specification can be found in Online Appendix A.2, equation (13).

[24]The estimated S_P coefficient barely changes when we use a polynomial for experience rather than dummies.

very low education. We also formally test for equality of the coefficients using an F-test, resulting in a test statistic equal to 1.53 with an associated p-value of 0.19, so we cannot reject that the coefficients are equal. This justifies our assumption of a constant ν that applies equally for all education groups (Corollary 1).

Motivation for a Structural Model In our sample, the intergenerational schooling regression coefficient is 0.46. Thus, (15b) implies that the selection effect must be larger than the spillover effect. In fact, in Online Appendix A.2 we find that $b_{\text{select}} \approx 0.111$, almost 6 times larger than b_{spill}. However, this analysis ignores that children's schooling and earnings are jointly determined endogenous outcomes, and simulating the model leads to a counterfactually high correlation between parental schooling and children's earnings. Instead, in the next section we will find that the ability selection effect is muted, and that it is preferences that predominantly explains intergenerational schooling persistence. Moreover, to compare our model against the empirical literature on intergenerational causal effects, we must perform counterfactuals, which necessitates consistent estimates of all model parameters (α, φ, etc.) This is what we turn to next.

4 Structural Estimation

Since schooling and earnings are solely functions of abilities (z) and parents' schooling (S_P), the previous regressions may overestimate the role of ability selection and spillovers if there are other dimensions of unobserved heterogeneity, in particular non-pecuniary preferences for schooling. On the one hand, families with highly educated moms may have stronger preferences for their children to pursue schooling, independent of any financial gain. Then to explain a positive intergenerational schooling relationship in the context of (15b), the effect of ability selection would need to be smaller, the spillover effect larger, or some combination of the two. On the other hand, given that a mother's schooling is also positively correlated with her child's earnings, the effect of ability selection and/or spillovers must be greater if the child's educational attainment was driven mainly by non-pecuniary rather than pecuniary motives.

To quantitatively assess the relative importance of each, we now turn to estimating our general model in the opening of Section 2. We make the following assumptions:

1. Moms have children at age 24, and their human capital remains constant until the children complete their education.

2. The choice of schooling is restricted to a discrete set of years, $\{8, 10, 12, 14, 16, 18\}$. Accordingly, we solve a discrete-time version of the model, where each time period corresponds to two years.

3. Non-pecuniary benefits from schooling are modeled as logit shocks.

4. Following Del Boca et al. (2014), total returns to human capital investment during the schooling phase is a linear function of age a, $\zeta(a) \equiv \alpha_2(a) + \alpha_3(a)$. This return evolves linearly over the child's age range from 6 to 24, from ζ_0 to ζ_1. The parent's contribution in human capital investment, $\zeta_P(a) \equiv \alpha_3(a)/\zeta(a)$, also changes linearly with age, from ζ_{P0} to ζ_{P1}.

5. On-the-job training uses only time inputs and differs from the schooling phase: $\alpha_{1W} = \alpha_W \neq \alpha_1$ and $\alpha_{2W} = 0$.

Although we cannot derive closed form solutions for schooling and earnings as in Section 2, Corollary 1 still disciplines our choice of moments. Accordingly, we will demonstrate that the underlying intuition for identification in the simple model carries over to the general model. Then in Section 5, we will use these estimates to perform counterfactuals and to interpret the empirical results from the IV literature.

4.1 Population Distribution Assumptions

Parents Our HRS sample reports parents' schooling, but does not contain any information on parents' earnings. So we instead rely on the HRS AHEAD cohorts between 1890-1923, approximately one generation older than the original HRS cohort, to set parental states by years of schooling. Specifically, we first run an individual fixed-effects regression between the ages of 23-42 also with a full set of experience-by-schooling effects, separately for men and women (AHEAD cohorts have 0-17 years of schooling). Then in line with (13), we set the estimated fixed effects as the parents' z_P, and each individual's predicted PDV of lifetime earnings as y_P.

Applying Propositions 1-2 and ignoring grandparents, S_P is solely a function of z_P, while h_{S_P} is a function of both (S_P, z_P). Thus, similar to (7), we regress the estimated $\log z_P$'s on a complete set of years-of-schooling effects, and store the estimated effects as $\log h_P$. We are imposing a one-to-one relationship between parents' S_P and h_P, ignoring the direct effect of Z on h_P in (7). This choice is guided by the fact that the identification of $(\nu, \lambda, b_{\text{select}})$ does not depend on Z (Corollaries 1-2), and our main objective is to estimate the effect of exogenously increasing parents' schooling while controlling for ability selection. The resulting correlation in the AHEAD data between $(\log h_P, \log z_P)$ is $\rho_{z_P S_P} = 0.48$, and regressing $\log h_P$ on S_P yields a coefficient of $\beta_{h_P S_P} = 0.1$.

Initial wealth $y(0)$ is intended to capture household resources rather than the mom's resources alone.[25] To construct the distribution of $y(0)$, we combine information on lifetime earnings from AHEAD with observed assortative matching patterns from the HRS as follows. For a mom with S_M years of schooling, we draw the mother's component of household resources, denoted y_{S_M}, from the distribution of the present discounted value (PDV) of lifetime earnings in AHEAD for women with S_M years of schooling. To account for the dads' resources, we first estimate the distribution of dads' schooling conditional on moms' schooling from the HRS, denoted $G(S_D \mid S_M)$. We then assign a dad to each mom according to this empirical matching distribution, and for each paternal schooling level S_D, compute the average male PDV of lifetime earnings in AHEAD, denoted $\bar{y}_{S_D}$.

Initial household wealth is then defined as the average of maternal and paternal lifetime resources, $y(0) = y_{S_M} + \bar{y}_{S_D}$. Thus, moms with the same level of schooling may face substantially different initial resource levels because they are matched to dads with different schooling and earnings capacities according to the observed assortative matching patterns in the data. This procedure generates a rich and empirically grounded distribution of initial household wealth while maintaining the tractability of the model.[26]

[25]Not only do dads usually earn more than moms, but also we assume moms devote all of their time to childcare without supplying any market labor during the child's first six years of life, as well as spend time with children during their schooling years.

[26]Loading both parents' lifetime income into $y(0)$ effectively abstracts from period-to-period credit constraints. We do so mainly to avoid making arbitrary assumptions about the joint distribution of moms' and dads' earnings processes since HRS only reports parents' schooling and not earnings, and instead focus on the joint determination of children's schooling and earnings outcomes. If the model featured such credit constraints, we suspect the estimated value of ν would be larger to capture the effect of h_P on children's earnings, but do not expect this would affect our counterfactual exercises given that our results attribute only a minor role to assets.

In summary, the h_P of HRS individuals whose mom and dad have (S_M, S_D) years of schooling is assumed to be a scalar that is determined by their moms' $S_P = S_M$, their z_P is assumed to be drawn from a distribution that varies by S_M, and $y(0)$ from a distribution that varies by both (S_M, S_D), as observed in AHEAD. Since we will calibrate w_P so that the units of h_P will not matter, and also since $y(0)$ only captures parents' lifetime earnings between ages 23-42, we normalize the means of h_P and $y(0)$ to 1.[27] The resulting correlation in AHEAD between $(\log y(0), \log z_P)$ is 0.32.

Abilities and Preferences Moms' years of schooling in HRS have a mean and standard deviation of 9.24 and 3.6 years, respectively as shown in Table 1.

We approximate the distribution of $(y(0), z_P)$ for each S_P with a multivariate lognormal distribution with the same mean and covariance structure as estimated above. Assuming that z has a mean of μ_z and is distributed lognormal with variance σ_z^2, we discretize (z_P, z) so that each have the same number of grid-points and construct a transition matrix assuming a correlation of ρ_z. Then β_z in (6) can be expressed as $\beta_z = \rho_z \cdot \sigma_z / \sigma_{z_P}$, so that β_{zS_P} in (15b) can be computed as

$$(19) \quad \beta_{zS_P} = \beta_z \cdot \beta_{z_P S_P} = \rho_z \cdot \underbrace{\rho_{z_P S_P}}_{0.48} \cdot \sigma_z / \underbrace{\sigma_{S_P}}_{3.6} \approx 0.133 \cdot \rho_z \sigma_z.$$

Preferences for schooling depend on parent's human capital and ability according to

$$(20) \quad \xi_i(S_i; h_P, z) \equiv \bar{\gamma}_{h_P}(S_i)(\log h_{Pi} - \mu_{\log h_P}) + \bar{\gamma}_z(S_i)(\log z_i - \mu_{\log z}) + \delta_{S_i} + \tilde{\xi}_{S_i}.$$

where $\tilde{\xi}_i \equiv [\tilde{\xi}_{S_i}]_{S_i \in \{10,12,14,16,18\}}$ is i.i.d. logistic with scale parameter σ_{ξ}. We assume that $\bar{\gamma}_{h_P}(S)$ linearly increases with years of schooling S from 0 at $S = 8$ to γ_{h_P} at $S = 18$, while $\bar{\gamma}_z(S)$ linearly increases every 4 years so is 0 for $S \in \{8, 10\}$, $\gamma_z/2$ for $S \in \{12, 14\}$ and γ_z for $S \in \{16, 18\}$. The constants γ_{h_p} and γ_z capture the correlations between ξ_i and (h_{Pi}, z_i), respectively. All details can be found in Online Appendix A.3.

To solve the dynamic discrete choice problem, we set a grid over (y, h) that evolves over time and solve the model backwards as described in Online Appendix A.3. This induces the optimal choice

[27] Since earnings beyond age 43 would be correlated with younger age earnings and heavily discounted, restricting age as we do should not have a large effect on the distribution of $y(0)$.

of S and resulting life-cycle earnings for any given initial condition $(y(0), h_P, z_P)$.

4.2 Generalized Method of Moments

Given the initial distribution of $(y(0), h_P, z_P)$, there are 29 parameters in the model, which we partition into

$$\theta_0 = [\rho, r; \chi, \theta; R, T; \delta_8]$$
$$\theta_1 = \left[(\nu, \lambda, \rho_z); (\alpha_W, \phi, \alpha_1, \zeta_0, \zeta_1, \zeta_{P0}, \zeta_{P1}); (\mu_z, \sigma_z); w_P; (\delta_{10}, \delta_{12}, \delta_{14}, \delta_{16}, \delta_{18}, \gamma_{h_P}, \gamma_z, \sigma_\xi); w\right].$$

The first partition, θ_0, are parameters that are set *a priori*. The rest are estimated by indirect inference to fit schooling and earnings moments.

Parameters Set a Priori We fix the interest rate at 5 percent, which is in the range of the after-tax rate of return on capital reported in Poterba (1998) and used in Heckman et al. (1998).Retirement age is fixed at 65 and death at 78.

The main concern in our model is children's lifetime outcomes which trades off against parents' old-age outcomes, rather than short-term fluctuations. So we assume log-utility, which is also symmetric with (20). And because we estimate the relative wages of both the parent and child generation, θ is not identified and instead fixed at 0.5.[28] Since preferences for schooling are relative, we normalize $\delta_8 = 0$ and estimate $\delta_S, S \in \{10, 12, 14, 16, 18\}$ to fit schooling shares. The entire list of exogenously fixed parameters are summarized in Table 4.

Estimated Parameters The remaining 22 parameters in θ_1 are estimated by indirect inference. We use 72 empirical moments. The first 62 come from the HRS: schooling and earnings outcomes by level of moms' schooling in the HRS, tabulated in the last 5 columns of Table 5, as well as the aggregate share of individuals who attain $S \in \{8, 10, 12, 14, 16, 18\}$ years of schooling, tabulated

[28] Typically, θ is identified from data on intergenerational transfers. However, because θ is constrained to the interval $(0, 1)$, it limits the range of intergenerational transfers that the model can admit. By contrast, the relative level of w to w_P allows the model to generate a much wider range of intergenerational transfers for any fixed θ.

Parameter	Value	Description
$(y_0, h_P, z_P; S_P)$		Initial assets, mom's human capital, and ability by mom's years of schooling; distribution approximated from HRS AHEAD cohorts
(ρ, r)	5%	Interest rate; after-tax rate of return on capital in Poterba (1998)
R	65	Retirement age, exogenous
T	78	Age of death, exogenous
χ	1	Log utility
θ	0.5	Not separately identified from (w, w_P), normalization
δ_8	0	Preference for 8 years of schooling, normalization

* See text for more details.

Table 4: Parameters Set a Priori

in the top panel of Table 6.[29]

For each of these 12 groups, we 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. For each level of moms' schooling, 1 of the 2 educational attainment shares of the children are redundant (since they add up to 1). All average earnings are normalized by the lowest level of average earnings, i.e. the age 25 average earnings of children with less than 12 years of schooling whose moms had 5 or less years of schooling, which is also dropped.

Four additional moments are added in the bottom panel of Table 5: the correlation between S and S_P, the OLS coefficient from regressing S on S_P, and the Mincer regression coefficients on S and S_P from specification (2) of Table 2. These are included to capture earnings and schooling gradients in the data we may miss by targeting average moments.

Nine additional moments come from AHEAD: moms' (relative) unemployment at ages 30-48 from AHEAD, in 2 years intervals, shown in the bottom panel of Table 6. Unemployment is defined as observations with zero earnings, and several women in the AHEAD cohort have long periods (more than 7 years) of unemployment. Our model assumes that moms stay home from age 24 to 29 (for the first 6 years of the child's life), and implicitly assumes they all work from age 48 onward, while in AHEAD we do not know the reason for women earning zero nor their family composition. Thus instead of the absolute moments, we target the difference between

[29]The empirical share of each of the 6 mom's schooling groups, as well as the average years of schooling for each subgroup with higher or lower education, are tabulated in Online Appendix Table A.3.

Mom's S_P	Child's S	Fraction (%)	Average Earnings at Age 23-27	28-32	33-37	38-42
≤5	≤11	59.82	1.00	1.43	1.75	1.93
	≥12	40.18	1.26	1.91	2.34	2.63
6-7	≤11	42.61	1.10	1.67	2.01	2.22
	≥12	57.39	1.27	1.93	2.42	2.75
8	≤12	67.30	1.36	1.99	2.45	2.73
	≥13	32.70	1.28	2.11	2.74	3.09
9-11	≤12	63.35	1.40	2.03	2.43	2.67
	≥13	36.65	1.29	2.12	2.66	2.99
12	≤12	45.57	1.55	2.27	2.67	2.87
	≥13	54.43	1.37	2.32	2.92	3.42
≥13	≤12	22.61	1.42	2.00	2.41	2.66
	≥13	77.39	1.30	2.21	2.84	3.32
(S, S_P) correlation and OLS:			0.47	0.44		
Mincer coefficients on (S, S_P):			0.08	0.02		

Table 5: Targeted Empirical Moments

For moms with low S_P (the first four rows), we divide low/high educational attainment of children by whether or not he graduated from high school, while for the rest we divide by whether or not he advanced beyond high school. The share of each S_P-group, and the average years of schooling for each S-group are summarized in Online Appendix Table A.2. 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. (S, S_P) OLS denotes the coefficient from regressing S on S_P, and the Mincer coefficients are from specification (2) in Table 2.

unemployment at women's ages $a_P = 30, \ldots, 46$ and age 48. Last, we also include the average share of inheritances to lifetime earnings, which Hendricks (2007) reports to be approximately 3 percent from a survey of literature and various data sources. In sum, we target 72 empirical moments with 22 model parameters.

Denote these target moments $\hat{\Psi}$. For an arbitrary value of θ_1, we compute the implied model moments $\Psi(\theta_1)$ as described in Online Appendix A.3. The parameter estimate $\hat{\theta}_1$ is found by

$$\hat{\theta}_1 = \arg\min_{\theta_1 \in \Theta_1} \left(\hat{\Psi} - \Psi(\theta_1)\right)' W \left(\hat{\Psi} - \Psi(\theta_1)\right)$$

where W is a weighting matrix. This procedure generates a consistent estimate of θ_1. We use a diagonal weighting matrix $W = diag(V^{-1})$, where V is the variance-covariance matrix of $\hat{\Psi}$. This weighting scheme allows for heteroskedasticity and can have better finite sample properties than the optimal weighting matrix (Altonji and Segal 1996) in practice. Additionally, since we draw

Child's S					8	10	12	14	16	18
Fraction (%)					13.39	8.78	36.41	15.12	13.79	12.41
Mom's Age	30	32	34	36	38	40	42	44	46	48
UE (%)	45.95	46.66	45.43	43.34	41.37	40.49	39.70	38.27	37.11	36.31
Target	9.64	10.35	9.12	7.03	5.06	4.18	3.39	1.96	0.80	-

Table 6: Targeted Empirical Moments, Continued

The top panel shows the fraction of men in the HRS who attain $(S-2, S]$ years of schooling, except for the first entry which is the fraction who achieve 8 or less years of schooling. The bottom panel shows the fraction of unemployed women in AHEAD, defined as those who report zero earnings, by age. The last row shows the difference between the unemployment rate at age x and age 48, which we target in the estimation.

moments from 3 separate sources (62 moments from HRS, 9 from AHEAD and 1 intergenerational transfer moment), we normalize the weights so that the sum of weights from each source sums to 1. Minimization is performed using a MADS (Mesh Adaptive Direct Search) pattern search algorithm, with several hundred different starting values to numerically search over a wide range of the parameter space (most of which have naturally defined boundaries). Asymptotic standard errors for the parameter estimates are obtained from

$$\sqrt{N}(\hat{\theta}_1 - \theta_1^*) \to \left(G'WG\right)^{-1} G'W\hat{V}WG\left(G'WG\right)^{-1}$$

as N approaches ∞, where $\hat{\theta}_1$ is the estimate, θ_1^* is the unknown, true parameter, and N the sample size. The matrix G is the $M \times P$ Jacobian of $\Psi(\theta_1)$ with respect to θ_1, where $(M = 72, P = 22)$ are the number of moments and parameters, respectively, and is computed numerically. The estimate of V, $\hat{V}$, is estimated via 2000 bootstrap draws from our HRS and AHEAD samples.

Identification As is usual with this class of models, identification is hard to prove formally. While all moments influence all parameters, our choice of moments is guided by intuition from the simpler model of how certain moments should have more influence on certain parameters. Here we provide a brief description with a more detailed explanation of key parameters in the following subsection.

According to Corollary 1, ν can be identified by the coefficient on mom's schooling in an earnings regression with a complete set of schooling$\times$experience effects, while Table 2 showed that the coefficient remains stable across different specifications. Although this relationship is affected by the inclusion of preferences for schooling, the intuition should still hold conditional on such preferences. Thus ν should still be identified from earnings differences among children with similar levels of own schooling but different levels of moms' schooling, as tabulated in Table 5. Similarly, λ is identified from how earnings vary by own schooling levels among children whose moms have similar schooling levels (Corollary 2). Conditional on preferences and (ν, λ), the degree to which sons' earnings ultimately depend on moms' schooling is determined by ρ_z (Corollary 1), so the three parameters are jointly identified by the Mincer coefficient on mom's schooling. As we explained in Section 2.3, the intuition for identification carries over from the simple model since our identifying assumption is that schooling S is determined by (h_P, z) and earnings are determined only by (S, z).

The human capital production technology parameters are identified by life-cycle earnings moments. Since α_W governs the speed of human capital growth in the working phase, it is identified by the average slope of experience-earnings profiles. On the other hand, the initial productivity of human capital production ϕ and schooling parameters $(\alpha_1, \zeta(a))$ are identified by initial earnings and the subsequent slope of earnings profiles, and how they vary by own and moms' schooling levels. When ϕ is high, the slope of the earnings profile becomes lower because early age human capital is higher. Similarly, for high values of $\alpha(a) \equiv \alpha_1 + \zeta(a)$, individuals with high S will have higher early age earnings and thus flatter earnings profiles. For high values of α_1, individuals with high S_P will benefit more from higher age-6 human capital and thus have higher initial earnings, leading to flatter earnings profiles. The mean of abilities μ_z and standard deviation of log abilities σ_z are identified by the overall mean and variation of earnings.[30]

Moms' wages w_P and the parameters $\zeta_P(a)$ which govern the importance of moms' time input into children's human capital production, are identified by women's relative unemployment in AHEAD. When w_P is high, spending time with children becomes costly, while a high $\zeta_P(a)$ increases the returns to time investments. Women's relative unemployment by age, in Table 6 identifies how $[\zeta(a), \zeta_P(a)]$ evolves linearly with age from (ζ_0, ζ_{P0}) to (ζ_1, ζ_{P1}).

Since δ_8 is normalized to 0, the 5 preference parameters δ_S, $S \in \{10, 12, 14, 16, 18\}$ can account for the shares of the 6 schooling levels. Then given all the other parameters, (γ_{h_P}, γ_z), determines how educational attainment varies by moms' schooling and individual earnings. The inclusion of γ_{h_P} is particularly important: It determines the intergenerational persistence of the schooling

[30] Since we normalized the earnings of the first cell in Table 5 to 1, μ_z is in normalized units.

Spillovers		HC Production				Abilities	
ν	0.541	α_1	0.296	α_W	0.204	ρ_z	0.460
	(0.002)		(0.004)		(0.002)		(0.001)
λ	0.052	ζ_0	0.482	ζ_1	0.625	μ_z	0.479
	(0.003)		(0.003)		(0.002)		(0.006)
ϕ	1.554	ζ_{P0}	0.772	ζ_{P1}	0.561	σ_z	0.641
	(0.033)		(0.006)		(0.000)		(0.001)
		w_P	0.022	w	0.012		
			(0.000)		(-)		
Preferences							
γ_{h_P}	4.792	δ_{10}	0.010	δ_{14}	0.790	δ_{18}	1.004
	(0.091)		(0.001)		(0.009)		(0.010)
γ_z	1.604	δ_{12}	1.014	δ_{16}	0.847	σ_ξ	0.335
	(0.026)		(0.015)		(0.009)		(0.006)

*Standard errors in parentheses.

Table 7: Parameter Estimates

(both the regression coefficient and correlation), and without it, moms' effect on how long children stay in school would be purely pecuniary. The shape parameter ξ governs how much the next generations' education outcomes depend on the other parameters, so it is identified by how connected own schooling and earnings outcomes are at each level of moms' schooling.

Given that the altruism parameter $\theta = 0.5$ is fixed, parents transfer a smaller (larger) amount of their assets to children when the next generation earns a higher (lower) wage. Thus, the share of inheritance to lifetime earnings identifies the child generation's wage w given the parent generation's wage w_P. But since this is an aggregate moment and not identified from our sample, we cannot construct standard errors for w.

4.3 Interpreting the Parameters

Table 7 presents the 22 estimated parameters and their asymptotic standard errors, with model-implied moments reported in Online Appendix Tables A.3-A.4. The model closely matches educational attainment shares and generally reproduces earnings moments. More importantly, the four regression moments (Table 5, lower panel; Table 9, first panel) are also well matched, though the own-schooling coefficient on earnings is somewhat lower.

Human Capital Production The estimated values of $(\zeta_0, \zeta_1, \zeta_{P0}, \zeta_{P1})$ imply that between ages 6 and 22, the returns to monetary investments, $\alpha_2(a)$, increase from 0.110 to 0.274. By contrast, the returns to maternal time inputs remain relatively stable, declining only slightly from 0.372 to 0.351. These patterns are qualitatively consistent with Lee and Seshadri (2019), who show that parental time investments play a much larger role than monetary investments, with their importance diminishing only moderately as children age.

The child's return to human capital production during the schooling phase, $\alpha_1 + \alpha_2(a)$, falls within the range of estimates, reported in the literature that use comparable BP technologies. Estimates surveyed by Browning et al. (1999) lie in the range 0.5 to 0.9. By contrast, the working phase estimate α_W is lower, implying that, the mother would prefer to keep her son in school longer for human capital accumulation purposes. This pattern arises from the introduction of preference shocks for schooling in the structural model: high-S_P mothers keep their children in school longer, and because these children tend to have higher abilities (z), they must obtain relatively higher returns in the schooling phase rather than in the working phase for the model to generate a positive relationship between schooling and earnings (conditional on experience). If returns were higher in the working phase, those effects would get absorbed into experience and the relationship between schooling and earnings would become weaker.

Parental Spillover and Early Childhood The magnitude of $\nu = 0.541$ seemingly implies large spillovers. Since h_P increases on average by 10 percent with an additional year of mom's schooling, a one-year increase in schooling would raise her child's initial human capital by approximately $\beta_{h_P S_P} \nu \approx 5\%$, controlling for abilities and preferences. As shown in Section 5, however, the causal effect on lifetime earnings is more modest—about 2 percent, which is close to the 1.7 percent estimate reported in Section 3.2, Table 2. The reason is that higher human capital early in life leads to less schooling and less human capital accumulation.

The astute reader might wonder if our large estimate reflects the assumption that parents only have a level effect, i.e., that h_P only has a causal effect on h_0 but not on z, nor on the speed of human capital accumulation. Unfortunately, such a slope effect is not separately identified from ρ_z in the model of Section 2. To examine this issue, we ran several numerical simulations using the structural model and we found that incorporating a slope spillover has little influence on parameter estimates other than ρ_z. This suggests that adding a slope spillover would mainly reduce the estimated ability selection effects. Thus, rather than overstating the spillover, our estimate can be interpreted as a conservative lower bound.

Selection on Learning Abilities The estimate for ρ_z implies that on average, moms with one standard deviation more schooling than the population mean have children with learning abilities $\rho_{zS_P} \equiv \rho_{z_P S_P} \cdot \rho_z = 0.22$ standard deviations higher. Using the empirical estimate of $\sigma_{S_P} \approx 3.6$ together with the model estimate of σ_z, this corresponds to an increase of approximately $\rho_{zS_P}\sigma_z/\sigma_{S_P} \approx 4\%$ in ability for each additional year of maternal schooling. Combined with $\lambda = 0.052$, this indicates that selection on abilities has only a minimal effect on initial human capital.

Instead, higher abilities sustain through life and raise both schooling and earnings. According to Proposition 2, the effect on earnings at all ages is approximately $\Delta \log z/(1-\alpha_W) \approx 5\%$ for a one-year difference in moms' schooling. According to the model, this implies that earnings differences become more pronounced at older ages once schooling and on-the-job training are completed, as we soon show in Section 5.

This also explains why the estimated value of λ is small, suggesting that moms' education S_P plays a crucial role in early human capital formation independently of learning abilities. If this were not the case, given that S_P is correlated with z_P, a higher λ would imply a counterfactually large effect of moms' schooling on lifetime earnings, given the estimated value of ρ_z.

Following similar calculations, Proposition 2 implies that a standard deviation of 0.647 for z translates into a $\sigma_z/(1-\alpha_W) \approx 80.4\%$ standard deviation in sons' log-earnings, once we control for schooling and experience. This result is consistent with the data: The residual variance in both specifications (2) and (3) in Table 2 is 71 percent.

Preferences for Schooling As expected, the estimates for δ_S are increasing in S except for δ_{12}, which takes on a large value for the model to match the K-12 system in the United States. Given the log utility for consumption and preferences for schooling, the estimated scale parameter σ_ξ translates into a standard deviation of approximately 61 percent in utility.

To facilitate the interpretation of the preference parameters, Table 8 presents the pecuniary and non-pecuniary benefits from schooling for different groups of individuals. The top panel

Schooling	$S = 10$	$S = 12$	$S = 14$	$S = 16$	$S = 18$
PDV Lifetime Earnings gain by S_P (log-difference)					
$S_P < 8$	-0.111	-0.224	-0.341	-0.465	-0.220
$S_P \in [8, 12)$	-0.110	-0.222	-0.338	-0.459	-0.205
$S_P = 12$	-0.110	-0.221	-0.337	-0.456	-0.244
$S_P > 12$	-0.110	-0.221	-0.336	-0.456	-0.262
Utility gain from h_P by S_P					
$S_P < 8$	-0.622	-1.119	-1.509	-1.807	-2.028
$S_P \in [8, 12)$	-0.280	-0.503	-0.678	-0.812	-0.911
$S_P = 12$	0.118	0.211	0.285	0.341	0.383
$S_P > 12$	0.435	0.782	1.054	1.262	1.416
PDV Lifetime Earnings gain by z (log-difference)					
Q1	-0.110	-0.223	-0.338	-0.455	-0.057
Q2	-0.111	-0.223	-0.338	-0.455	-0.055
Q3	-0.110	-0.222	-0.338	-0.457	-0.166
Q4	-0.109	-0.222	-0.338	-0.457	-0.234
Utility gain from z					
Q1	-	-1.007	-0.953	-1.547	-1.421
Q2	-	-0.935	-0.848	-1.525	-1.381
Q3	-	-0.439	-0.485	-0.337	-0.445
Q4	-	0.037	0.026	0.133	0.086

Table 8: Pecuniary and Non-Pecuniary Benefits of Schooling

All values discounted to age 24, and relative to $S = 8$.

reports the PDV lifetime earnings gain from additional schooling relative to $S = 8$ by moms' years of schooling. While not shown in the table, the lifetime earnings of individuals with the most educated moms are approximately 30 percent to 50 percent larger than those with the least educated moms, across all schooling levels. Nonetheless, the relative gains from extra schooling do not vary much by moms' S_P, except at $S = 18$.[31] While individuals with higher S benefit from the superior human capital production technology in the schooling phase (relative to the working phase), the forgone earnings losses are significant. This loss attenuates at $S = 18$, but remains large for individuals with higher S_P as they also tend to have higher z (through ρ_z), and thus

[31] Because δ_S induces all individuals to prefer S years of schooling, a monotonic relationship in relative gains should not be expected.

larger forgone earnings.

Despite these losses, individuals with higher S_P remain in school longer due to the utility gains from γ_P. As shown in the lower part of the top panel, individuals whose moms have low S_P dislike longer schooling, whereas the opposite is true for those with higher S_P.

Similarly, individuals in the highest quartile have the highest pecuniary benefits across all schooling levels–ranging from 1.5 to 2 times the lifetime earnings of those in the lowest quartile)—while the relative gains from additional schooling remain fairly stable across quartiles, except at $S = 18$. Because individuals with higher z are more likely to have parents with higher S_P, they tend to remain in school longer, thereby losing more pecuniary value at higher schooling levels. Nonetheless, the utility gains from additional schooling are generally larger for individuals with higher z, although the variation in magnitude is more limited since γ_z is much smaller than γ_{h_P}.

That γ_{h_P} is substantially larger than γ_z suggests that non-pecuniary motives for staying in school are more strongly influenced by moms' schooling than by children's learning abilities. This aligns with the idea that highly educated moms often provide a family environment that actively encourages their own children to pursue further education. This correlation between preferences for children's schooling (non-pecuniary benefits) and parents' schooling plays an important role in the following analyses.

The estimated value of γ_z is low because abilities have a limited influence on the intergenerational schooling relationship. Similar to how the point estimate for λ was small given the estimate of ρ_z, the value of γ_z must also be small given σ_z. Otherwise, the effect of own schooling on observed earnings would be too large.

4.4 Quantitative Analysis

Shutting down parameters To understand the effect of spillovers and selection in light of our empirical analysis from Section 3, it is useful to see how the gradient moments change when we shut down the key parameters. The second panel of Table 9 shows the changes in these moments when we set the key parameters ν, ρ_z, or both, to zero.[32] The third panel shows the changes in the same moments when we don't allow preferences for schooling to vary by h_P or z by setting γ_{h_P}, γ_z

[32]Similar exercises with λ had minimal effects, as its estimated value is already close to zero.

	Corr_S	OLS_S	S coeff.	S_P coeff.
Data	0.469	0.444	0.082	0.017
Model	0.484	0.416	0.045	0.017
Pecuniary				
$\nu = 0$	0.491	0.423	0.037	-0.005
$\rho_z = 0$	0.433	0.366	0.046	0.000
$\rho_z = \nu = 0$	0.440	0.373	0.039	-0.022
Non-pecuniary				
$\gamma_{h_P} = 0$	0.070	0.054	0.044	0.043
$\gamma_z = 0$	0.452	0.361	-0.010	0.044
$\gamma_z = \gamma_{h_P} = 0$	-0.015	-0.011	0.001	0.044

Table 9: Effect of Spillover and Correlation Parameters.
(μ_{h_P}, μ_z) means that we shut down the correlation between preferences and (h_P, z). Corr_S and OLS_S denote, respectively, the correlation between S and S_P, and the OLS coefficient from regressing S on S_P. (S, S_P) coefficients correspond to the coefficients from the Mincer regression in column (2) of Table 2.

or both, to zero. These exercises show that the intuition for the identification of spillovers based on the simple model in Section 2 carries over to the estimated model. They also show how we can separately identify selection on abilities from non-pecuniary effects.

When $\nu = 0$, the level effect disappears, inducing high ability individuals (whose moms tend to have more schooling) to increase their length of schooling. This, in turn, slightly raises the correlation of schooling across generations, but the quantitative effect is small. Ability selection also has a modest effect: When both ρ_z and ν are set to zero, schooling persistence drops only slightly compared to the benchmark. This contrasts with the simple model's prediction, where according to (15b), it should be zero. This is partly because the estimated returns to schooling are much larger than those from on-the-job training, but more because the persistence of schooling across generations is primarily driven by the intergenerational transmission of heterogeneous preferences for education.

Although turning off either ν or ρ_z has a limited effect on the own schooling S coefficient, it causes the moms' schooling S_P coefficient to drop to zero. The intuition for this can be found in the propositions in Section 2.2. According to Proposition 1, the initial human capital effect of both is absorbed into the choice for schooling, while the relationship between earnings and schooling in Proposition 2 remains the same; this implied earnings become more or less independent of moms' schooling. In fact, when both are zero, the S_P coefficient is negative because of the correlation of preferences for schooling with S_P: Sons of highly educated moms remain in school longer for non-pecuniary reasons at the cost of lesser earnings.

The third panel confirms the crucial role of preferences in explaining the persistence of schooling across generations. When preferences do not vary with moms' schooling (row $\gamma_{h_P} = 0$), children of high human capital parents (who tend to have higher levels of schooling) no longer have a preference for longer schooling. Consequently, both Corr_S and OLS_S become negligible. In this scenario, selection based on abilities alone is not enough to generate a high persistence. However, the Mincer coefficient on moms' schooling increases by almost 2.5-fold: When moms' schooling does not affect non-pecuniary benefits, its impact on pecuniary earnings becomes larger. Thus if we were to force the model to explain the intergenerational schooling relationship solely through selection on abilities, the Mincer coefficient on moms' schooling would be too high compared to the data. Conversely, the observed coefficient is too low to explain the intergenerational persistence of schooling through selection on abilities alone.

Similarly, the S_P-coefficient increases by a similar magnitude when we don't allow preferences to vary with abilities (4th column of row $\gamma_z = 0$). Moms with high S_P still tend to have higher z children. If a higher z no longer provides any non-pecuniary benefits for schooling, the earnings effect is magnified, as earnings are explained by z, controlling for schooling and experience (Proposition 2). However, the own-schooling coefficient also drops to nearly zero. While the high estimate for γ_{h_P} drives the intergenerational persistence of schooling, a positive value of γ_z ensures that among children with high h_P, those with a higher z still achieve more schooling. Nonetheless, setting $\gamma_z = 0$ has little effect on schooling persistence, as it is primarily influenced by γ_{h_P}. In fact, when preferences do not vary along either dimension, schooling persistence is only slightly smaller than when it varies solely with abilities (last row), while the dependence of earnings on parents (S_P-coefficient) remains high.

Reduced-form comparison Given the large point estimate of γ_{h_P}, children of higher educated moms like to stay in school longer. We can now evaluate the degree to which overlooking preferences in the basic model biases our estimates of selection on abilities and spillover effects. Compare the model-predicted values of $(b_{\text{spill}}, b_{\text{select}})$ in equation (15b) to their reduced-form estimates, which were 0.017 and 0.111, respectively, as shown in Table 2 and Online Appendix Table A.1. According to our modeling assumptions, $\beta_{h_P S_P} = 0.1$, and using (19) and our estimates for (ρ_z, σ_z), we can calculate the value of $\beta_{z S_P}$:

$$\beta_{z S_P} = 0.133 \cdot \rho_z \cdot \sigma_z = 0.133 \cdot 0.46 \cdot 0.641 = 0.039$$

Now, using our estimates for (ν, λ, α_W) and the range of $[\alpha_1(a), \zeta(a)]$, noting that $\zeta(a)$ ranges from ζ_0 to ζ_1, we can determine the values of b_{spill} and b_{select}.

$$\begin{aligned} b_{\text{spill}} =& \nu \beta_{h_P S_P} / \left[1 - \lambda(1-\alpha)\right] \begin{cases} = 0.056 & \text{if } \alpha = \alpha_W, \\ \in [0.054, 0.055] & \text{if } \alpha = \alpha_1 + \zeta(a), \end{cases} \\ b_{\text{select}} =& \beta_{z S_P} / (1-\alpha) \begin{cases} = 0.049 & \text{if } \alpha = \alpha_W, \\ \in [0.176, 0.491] & \text{if } \alpha = \alpha_1 + \zeta(a). \end{cases} \end{aligned}$$

When $\alpha = \alpha_W$, the implied ability selection coefficient is an order of magnitude smaller than its reduced-form estimate (Corollary 1 in Online Appendix A.2), while the implied spillover coefficient is slightly larger. This discrepancy arises because our structural model allows for non-pecuniary benefits whereas the simple model (and thus Corollary 1) does not. Choices based on non-pecuniary benefits reduce lifetime earnings but raise schooling, while in the model, a higher value of ν leads to less time in school but results in higher lifetime earnings. Therefore, to explain the reduced-form S_P coefficient of around 2 percent, our model requires higher values for both ν and the returns to schooling $\zeta(a)$. Consequently, when $\alpha = \alpha_1(a) + \zeta(a)$, the simple-model-implied selection coefficient is also larger, although it remains similar in magnitude to its reduced-form estimate.

In summary, while the intuition for identification carries over, by attributing the intergenerational schooling relationship to preferences, the structural model admits a larger estimate for ν and thus a greater spillover effect.[33]

[33]While the intergenerational borrowing constraint does play a large role in explaining the aggregate level of the next generation's earnings (e.g. eliminating the constraint raises children's earnings by an average of 10 percentage points), it has minimal quantitative impact on schooling and regression moments.

5 Counterfactual Experiments

We conduct two main experiments using the model estimates. First, we decompose the effects of a one-year increase in moms' schooling. Second, we implement a counterfactual compulsory schooling reform. Although the spillover parameter ν only affects children's initial human capital, we find that a uniform one-year increase in moms' schooling leads to an average 2 percent increase in children's lifetime earnings when controlling for selection on abilities and preferences for schooling. At the same time, the effect on children's years of schooling is modest. This pattern is explained by parents with higher human capital having a slightly negative effect on children's schooling (Proposition 1), which is offset by the stronger preferences for schooling among children of higher human capital parents.

5.1 Decomposing Spillovers from Selection

Given a mom-son pair's state $(y(0), h_P, z_P)$ when the child is born, we compute the change in the child's schooling and earnings outcomes in response to a uniform one-year increase in moms' schooling, which translates into an increase of $\beta_{h_P S_P}$ units of $\log h_P$.

We perform several experiments to control for spillovers, preferences for schooling, assets, and ability selection. In Table 10, column (1), we hold constant (y_0, z_P, ξ) and also the schooling choice, which isolates the pure quality effect coming from higher S_P. To be precise, we hold (2) constant, not the values (J, I) nor the outcome variable S itself: Thus, since ours is a dynamic discrete choice model, S may still change because $(y(a), h(a))$ will differ as the mom-son pair ages. In column (2), we still hold (y_0, z_P, ξ) constant, but allow the individuals to re-optimize their choice of S. Since higher initial human capital substitutes for schooling, earnings should further increase while schooling decreases. The combined effect of the pure quality increase and schooling choice adjustment constitutes the spillover effect. In column (3), we allow $y(0)$ to also rise with h_P, as dictated by the distributional assumptions in Section 4.1. This captures the additional effects of higher-educated moms having higher PDV lifetime earnings (whether their own, or from a higher-earning husband through assortative mating), in addition to the spillover effect. Column (4) allows ξ to rise according to $\gamma_{h_P} \log h_P$, and finally in column (5), we also allow $\log z_P$ to rise with $\log h_P$, which we label a "reduced-form" effect—i.e., this simply compares the average

	(1) fixed S	(2) spillover	(3) +assets	(4) pref.	(5) RF
Schooling diff (years)	0.003	-0.011	-0.002	0.451	0.501
Avg earnings diff (%)	0.021	0.022	0.024	0.006	0.026
Age 30 earnings diff(%)	0.024	0.024	0.028	0.016	0.033
Age 50 earnings diff(%)	0.011	0.011	0.013	0.008	0.037

Table 10: Aggregate Effect of a One-Year Increase in Moms' Schooling
Units in % are approximated by log-point differences. The second row denotes the change in the cross-sectional average of the present discounted sum of lifetime earnings, in logarithms. Column (1) holds children's schooling constant, while column (2) holds assets, preferences, and abilities constant but allows schooling to vary according to the model. Columns (3)-(5) progressively allow assets, preferences and abilities to vary according to their estimated correlations with h_P. The coefficients from an intergenerational schooling OLS regression in the data and model are 0.469 and 0.444, respectively.

outcomes of children with S_P years of moms' schooling, to the average outcomes of those with $S_P + 1$ years of moms' schooling.

As discussed in the Introduction and Section 2.1, the policy-relevant effect from raising parents' schooling should lie between columns (3) and (4): Ability persistence is assumed to be an exogenous process, but we do not model how preferences for children's schooling are formed.

The first and second rows of Table 10 list the average effects of a one-year increase in S_P on schooling S and the present-discounted value of lifetime earnings, respectively. The third and fourth rows show the average change in earnings at ages 30 and 50. The change in S is measured in years and the change in earnings in log-point differences.

As expected, the spillover effect on schooling is slightly negative: increasing all moms' schooling by a year would lead to an average 0.011 year decline in the next generation's schooling. Also as expected, allowing for $y(0)$ to rise increases both schooling and earnings, but the effect is small. This indicates that, at least through the lens of our model, dads with higher schooling (who tend to have more assets) would have a positive effect on schooling and earnings, but the effect is modest.

When preferences also rise with moms' schooling, children's schooling increases by 0.451 years. Comparing columns (3) and (4), the effect of preferences alone is a slightly larger 0.453 years. Allowing abilities to also rise with S_P increases this by only another 0.05 years (columns 4 vs 5). Higher abilities do not substantially affect intergenerational schooling relationships, whether through pecuniary or non-pecuniary motives, as we saw in Table 9. Preferences correlated with higher S_P induce individuals to stay in school longer than what maximizes lifetime earnings, and since human capital accumulation faces diminishing returns, higher abilities do not further increase schooling substantially.[34] In sum, moms' S_P alone generates an intergenerational schooling relationship that is close to its empirical counterpart, while its relationship with abilities plays only a minimal role.[35]

The spillover raises lifetime earnings by roughly 2.1 percent when schooling choices are held constant, increasing to 2.2 and 2.4 percent once we account for dynamic discrete choice adjustments and richer parents, respectively. Overall, higher parental assets play only a limited role in children's earnings. However, when we also consider that sons with higher S_P moms may exhibit stronger preferences for schooling, the average impact falls by 1.8 percentage points (columns 3 vs. 4). This occurs because these stronger schooling preferences lead individuals to remain in school longer, deviating further from lifetime earnings maximization. Yet, the resulting negative effect does not fully offset the positive spillover effect. Selection on abilities raises earnings by an additional 2 percentage points (columns 4 vs 5).

We conclude that, independently of children's preferences for schooling, the causal effect of moms' education on earnings is roughly similar to the selection effect on abilities— that is, higher-ability moms having higher-ability children.

From which stage of the life-cycle do the gains in lifetime earnings arise? In Figure 4 we plot earnings effects over an individual's entire life-cycle. Each line plots the change in average earnings following a one-year increase in mom's schooling, for each age, compared to the benchmark estimates.

In "Fixed S", "spillover" and "+assets," we find that the lifetime earnings spillover effect arises almost entirely from early labor market entry, with only minimal gains occurring after age 24 (the latest entry age we allow in the model). This also means that on average, children of moms with less schooling catch up with those of moms with more schooling by staying in school longer, so

[34]Since $h_0 = z^{\lambda} h_P^{\nu}$, some of the incentive to increase schooling (since children can learn more in a fixed amount of time) is counterbalanced by the higher h_0 (since there is less need to learn when human capital is already high). Yet, this effect is minimal given the low estimated value of $\lambda = 0.05$.

[35]Online Appendix Figure A.4 shows that raising moms' S_P overall reduces their time investment in children across all scenarios (consistent with Behrman and Rosenzweig 2002), although some of the reduction at later ages is compensated by slightly more time spent with children before they turn 10 years old. This implies that raising dads' schooling would not have the same effect as raising moms', as long as men spend less time with children.

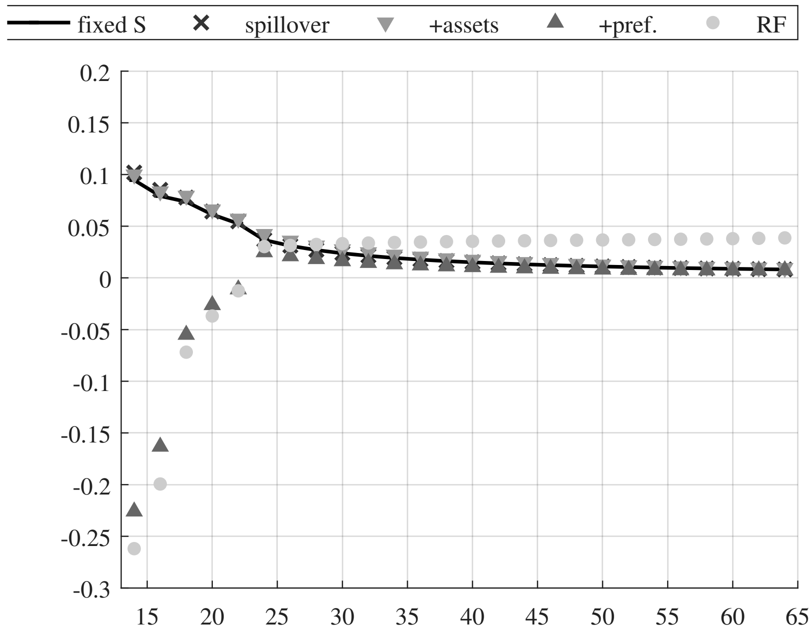


Figure 4: Life-Cycle Effect of a One-Year Increase in Mom's Schooling

that their earnings do not differ substantially later in life. Longer schooling induced by preferences increases lifetime earnings later in life slightly (through more human capital accumulated in school), but this is dominated by the foregone earnings earlier in life. Allowing for higher abilities (RF) has a fixed positive effect throughout the working phase (as expected from Section 2) which compensates for the foregone earnings earlier in life.

5.2 Counterfactual Compulsory Schooling Reform

We next impose a minimum schooling requirement that mimics compulsory schooling reforms. Such reforms took place in many countries throughout the 20th century and only affect parents who would otherwise not have attained the required level of schooling. Our exercise will show that it is possible to simultaneously estimate a large schooling OLS coefficient and a smaller or possibly a negative IV coefficient, indicating that a reform in the parents' generation may have a limited causal effect on children's schooling. Nonetheless, it can still have a positive causal effect on children's earnings.

We impose a minimum of 8 years of schooling for all parents and set the initial level of human

	(1) fixed S	(2) spillover	(3) +assets	(4) +pref.	(5) RF
Schooling OLS	0.468	0.467	0.467	0.442	0.439
Schooling IV	0.015	-0.009	-0.003	0.305	0.350
Avg earnings diff (%)	0.040	0.042	0.043	0.004	0.036
Age 30 earnings diff(%)	0.042	0.044	0.452	0.014	0.043
Age 50 earnings diff(%)	0.020	0.021	0.215	0.007	0.053

Table 11: Counterfactual Schooling Reform

Units in % are approximated by log-point differences. The second row denotes the change in the cross-sectional average of the present discounted sum of lifetime earnings, in logarithms, for only those individuals affected by the reform. Column (1) holds children's schooling constant, while column (2) holds assets, preferences, and abilities constant but allows schooling to vary according to the model. Columns (3)-(5) progressively allow assets, preferences and abilities to vary according to their estimated correlations with h_P. The coefficients from an intergenerational schooling OLS regression in the data and model are 0.469 and 0.444, respectively.

capital of all children to

$$h_0 = bz^{\lambda} h_P^{\nu} = bz^{\lambda} \exp\left[\nu \beta_{h_P S_P} \cdot \max\left\{S_P, 8\right\}\right].$$

We choose 8 years as it was in fact the compulsory schooling requirement in many U.S. states at the time, or soon after.[36] In our data, 25 percent of moms have fewer than 8 years of schooling, and the reform would raise the schooling of these moms by an average of 3.5 years.

For the schooling OLS and IV regressions, we merge simulated samples from two regimes: one without the minimum requirement (our benchmark model) and one with the requirement imposed. These represent mom-child pairs pre- and post-reform, respectively. We then run OLS and IV regressions on the merged data, using a dummy variable for the different regimes as an instrument. We repeat this exercise for the five cases discussed in the previous subsection: holding the next generation's schooling S fixed; allowing S to vary while controlling for $(z, \tilde{\xi})$; and then allowing z and/or $\tilde{\xi}$ to rise with h_P according to their correlations.

Table 11 reports the results. The first row shows the regression estimates for all cases. The second row reports the change in the present discounted sum of lifetime earnings, while the third and fourth rows show the average change in earnings at ages 30 and 50, respectively.

[36] Black et al. (2005) study the case of Norway, whose compulsory schooling requirement increased from 7 to 9 years in the 1960s. While the location and timing differ, our moms' average years of schooling is only 1 year lower (10.5 vs 9.3 years). However, the percentage of moms affected is almost twofold (12.4 percent vs 25 percent).

Our experiment generates a range of possible *model* IV estimates. Which of these should be compared with *empirical* IV estimates depends on the context of the instrument used in a particular study. As noted in the introduction, most instruments are chosen with an emphasis to capture exogenous variation in parents' schooling independent of their *abilities*. We suspect that some instruments may also be correlated with parents' preferences, which are passed on to their children. In that case, the relevant estimates likely fall somewhere between columns (3) and (4).

The OLS coefficients are somewhat harder to interpret because a quarter of the population in one regime receives varying levels of treatment (some moms are farther from 8 years than others). Even so, the OLS estimates rise in all four cases, reaching slightly above the model's benchmark value of 0.444. The IV regressions measure how children's schooling change when their moms obtain one more year of schooling, but only among those moms affected by the reform. The controlled effect is negative but smaller in magnitude (columns 2 and 3). The IV coefficient increases when we allow preferences for schooling to rise (columns 3 vs 4), and rises again slightly when we also allow moms' abilities to increase (columns 4 vs 5). The reduced-form coefficient is smaller than in Table 10, when we imposed a one-year increase in schooling for all moms. The difference arises because according to our modeling assumptions in (3d) and (20), both spillovers and preferences for schooling have weaker effects at lower levels of S_P, which is precisely the group affected by the reform.

We conjecture that this helps explain the puzzling finding in many studies— those using data sets on twins, adoptees, or compulsory schooling reforms— that IV estimates of the effect of parents' schooling children's schooling are often smaller than OLS estimates, and sometimes even zero or negative. First, the structural effect may in fact be negligible, as we have argued throughout this paper. More importantly, while the instruments used in these studies are likely to account for ability transmission, it is less clear whether and how they also account for the correlation of parents' schooling with preferences. The IV coefficient may fall close to zero because reforms at the national level, or adoption into a new family, may partly encourage children to develop stronger preferences for schooling.

We can think of this as moms who are *compelled* to attain more schooling having some influence on the family environment or on their children's non-cognitive abilities — factors that help their children remain in school longer. But this transmission of preferences is likely weaker than

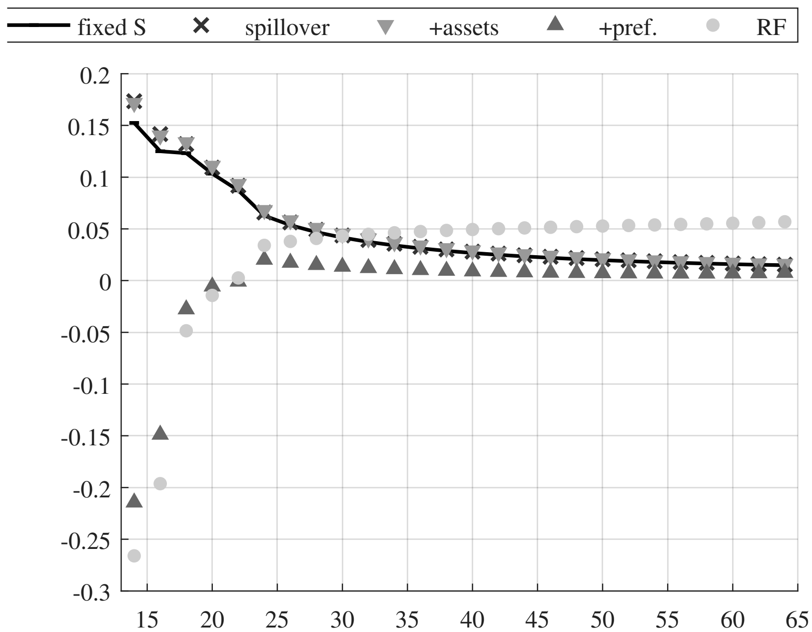


Figure 5: Life-Cycle Impact of Reform

for moms who *choose* to become highly educated. In our experiment, this implies that raising the schooling of moms who would otherwise attain very low levels could increase their children's schooling by anywhere from -0.01 to 0.3 years.[37] Similarly, adoptees may adjust to the schooling environment of their adoptive families, though this adaptation is likely weaker than that of biological children.

Also note that the controlled effect on lifetime earnings effect is much larger here than in Table 10, where we uniformly increased all moms' schooling by one year. As shown in Figure 5, the positive impact on earnings from earlier labor market entry is stronger and persists throughout the life-cycle. This contrasts with Figure 4, where the effect dissipated after age 24. For children of moms with low S_P, the gains in earnings from raising moms' human capital can be substantial.

At the same time, the negative effect of preferences remains. If a reform targeting low-educated moms also raises their children's preferences for schooling, then the effect on schooling may be small and the effect on lifetime earnings close to zero. In column (4), when mom-son pairs also adjust their preferences in response to the reform, lifetime earnings rise by only 0.4 percentage points. Since moms' schooling increases by 3.5 years on average, this implies that an additional

[37]This is consistent with the range of estimates found in Black et al. (2005) across different specifications.

year of moms' schooling raises sons' lifetime earnings by just 0.1 percentage points—considerably smaller than the 0.6 percent increase we saw in column (4) of Table 10. For sons of low educated moms, weaker initial preferences for schooling mean that raising such preferences can offset, and even outweigh, the positive spillover effects.

In summary, a positive IV coefficient on schooling may actually correspond to a smaller effect on earnings, and vice versa. This cautions against assuming a positive relationship between the two intergenerational coefficients simply because own schooling and own earnings are positively related.

6 Conclusion

We developed a model of human capital which features endogenous schooling and earnings to isolate the causal effect of parents' education on children's education and earnings outcomes. The model has several important ingredients—preferences for schooling that are correlated across generations, persistence in ability across generations, and finally, a parent's human capital is an input into the production of the child's human capital. We label the last feature a parental spillover. Despite the positive relationship between the child's own schooling and their earnings, the model shows that the causal effect of parent's schooling on children's schooling can be negligible, even when the causal effect on children's earnings is positive. Because children of higher-human capital parents begin life with higher human capital themselves, they may spend less time in school yet still achieve equal or higher earnings.

Our model is consistent with several features of the joint distribution of parents' schooling, children's schooling and children's earnings over the life-cycle. It is also consistent with a positive OLS correlation between parents' and children's schooling. Our estimates attribute this pattern primarily to the intergenerational transmission of preferences for schooling, rather than to selection on ability or parental spillovers.

Our results are also consistent with prior empirical work showing that IV estimates of the effect of parental schooling on children's schooling are typically smaller than OLS estimates and often negligible. The model highlights a mechanism for this: increasing parental human capital can reduce children's schooling due to diminishing returns to human capital accumulation, even if higher-human-capital parents also transmit stronger preferences for schooling. Nevertheless, the causal effect of mom's schooling on children's lifetime earnings is estimated to be up to 2 percent. While not directly comparable, our finding that the causal effect on earnings is of a similar magnitude to the ability selection effect is in line with the "nature-nurture" literature, which generally concludes that nurture effects are at most similar to, or smaller than, nature effects (Plug and Vijverberg 2003; Björklund et al. 2006).

Our estimates are based on U.S. cohorts born in the 1930s, whose parents were born in the 1900s to 1910s. Although education systems in advanced countries have changed markedly since, we believe our structural framework contributes to the literature that is also often based on data from older cohorts. In addition, our results underscore two insights applicable to more recent cohorts, for whom high school and college completion matter more than years of schooling alone: i) When considering education, it is important to consider the heterogeneous gains across individuals who attain the same education levels (e.g., the *quality* of degrees), and thus ii) causal intergenerational effects on quantity-based indicators of education (e.g., whether or not an individual attains a college degree) cannot be found without controlling for such differences in quality.

Data Availability Statement

We use confidential Social Security Administration earnings records linked to Health and Retirement Study respondents, which the authors are not authorized to disclose or redistribute. Qualified researchers may apply to the Health and Retirement Study for access to current restricted HRS-linked Social Security earnings products through the secure virtual data enclave of the Michigan Center for the Demography of Aging (MiCDA). Available products and vintages may differ from the historical files we used. Our openICPSR archive (Lee et al. 2026) provides disclosure-safe empirical inputs and code to reproduce the quantitative model results: `https://doi.org/10.3886/E251556V2`.

References

ADHAMI, M., M. BILS, C. I. JONES AND P. J. KLENOW, "Population and Welfare: Measuring Growth When Life Is Worth Living," Working Paper 31999, National Bureau of Economic Research, December 2023.

ALTONJI, J. G. AND L. M. SEGAL, "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics* 14 (July 1996), 353–66.

AVERY, C. AND S. TURNER, "Student Loans: Do College Students Borrow Too Much–Or Not Enough?," *Journal of Economic Perspectives* 26 (Winter 2012), 165–92.

BEHRMAN, J. R. AND M. R. ROSENZWEIG, "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?," *American Economic Review* 92 (March 2002), 323–334.

BEHRMAN, J. R. AND P. TAUBMAN, "Is Schooling "Mostly in the Genes"? Nature-Nurture Decomposition Using Data on Relatives," *Journal of Political Economy* 97 (December 1989), 1425–46.

BEN-PORATH, Y., "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy* 75 (1967), 352–365.

BETTS, J. R., "What Do Students Know about Wages? Evidence from a Survey of Undergraduates," *The Journal of Human Resources* 31 (1996), 27–56.

BJÖRKLUND, A., M. LINDAHL AND E. PLUG, "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *The Quarterly Journal of Economics* 121 (2006), 999–1028.

BLACK, S. E. AND P. J. DEVEREUX, "Recent Developments in Intergenerational Mobility," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, volume 4B (Amsterdam: Elsevier, 2011), 1487–1541.

BLACK, S. E., P. J. DEVEREUX AND K. G. SALVANES, "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," *The American Economic Review* 95 (2005), 437–449.

BOWLES, S. AND H. GINTIS, "The Inheritance of Inequality," *Journal of Economic Perspectives* 16 (Summer 2002), 3–30.

BOWLES, S., H. GINTIS AND M. OSBORNE, "The Determinants of Earnings: A Behavioral Approach," *Journal of Economic Literature* 39 (2001), 1137–1176.

BROWNING, M., L. P. HANSEN AND J. J. HECKMAN, "Micro Data and General Equilibrium Models," in J. B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics*, volume 1A (Amsterdam: Elsevier, 1999), 543–633.

CAMERON, S. V. AND J. J. HECKMAN, "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy* 106 (1998), 262–333.

———, "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy* 109 (2001), 455–499.

CAMERON, S. V. AND C. TABER, "Estimation of Educational Borrowing Constraints Using Returns to Schooling," *Journal of Political Economy* 112 (2004), 132–182.

CARD, D., "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, volume 3A (Amsterdam: Elsevier, 1999), 1801–1863.

CARROLL, C. D., "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters* 91 (2006), 312–320.

CHANG, Y., S. N. DURLAUF, S. LEE AND J. Y. PARK, "A Trajectories-Based Approach to Measuring Intergenerational Mobility," NBER Working Paper 31020, National Bureau of Economic Research, March 2023.

CHETTY, R., N. HENDREN, P. KLINE, E. SAEZ AND N. TURNER, "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," *American Economic Review* 104 (May 2014), 141–47.

CONNOLLY, M., C. HAECK AND J.-W. P. LALIBERTÉ, "Parental Education and the Rising Transmission of Income between Generations," in R. Chetty, J. N. Friedman, J. C. Gornick, B. Johnson and A. Kennickell, eds., *Measuring Distribution and Mobility of Income and Wealth* (Chicago: University of Chicago Press, 2022), 289–315.

CUNHA, F., J. J. HECKMAN AND S. M. SCHENNACH, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica* 78 (2010), 883–931.

DARUICH, D., "The Macroeconomic Consequences of Early Childhood Development Policies," *Journal of Political Economy* 134 (2026).

DEL BOCA, D., C. FLINN AND M. WISWALL, "Household Choices and Child Development," *Review of Economic Studies* 81 (2014), 137–185.

DEL BOCA, D., C. MONFARDINI AND C. NICOLETTI, "Parental and Child Time Investments and the Cognitive Development of Adolescents," *Journal of Labor Economics* 35 (2017), 565–608.

DRUEDAHL, J. AND T. H. JØRGENSEN, "A General Endogenous Grid Method for Multi-Dimensional Models with Non-Convexities and Constraints," *Journal of Economic Dynamics and Control* 74 (2017), 87–107.

FERNÁNDEZ, R., A. FOGLI AND C. OLIVETTI, "Mothers and Sons: Preference Formation and Female Labor Force Dynamics," *The Quarterly Journal of Economics* 119 (11 2004), 1249–1299.

GAYLE, G.-L., L. GOLAN AND M. A. SOYTAS, "What is the source of the intergenerational correlation in earnings?," *Journal of Monetary Economics* 129 (2022), 24–45.

GOLDIN, C. D. AND L. F. KATZ, *The race between education and technology* (Cambridge, MA: Harvard University Press, 2008).

HECKMAN, J., L. LOCHNER AND C. TABER, "Explaining Rising Wage Inequality: Explanations With A Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents," *Review of Economic Dynamics* 1 (January 1998), 1–58.

HECKMAN, J. J., L. J. LOCHNER AND P. E. TODD, "Earnings Functions and Rates of Return," *Journal of Human Capital* 2 (2008), 1–31.

HECKMAN, J. J., J. STIXRUD AND S. URZUA, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics* 24 (July 2006), 411–482.

HENDRICKS, L., "Retirement Wealth and Lifetime Earnings," *International Economic Review* 48 (2007), 421–456.

HOLMLUND, H., M. LINDAHL AND E. PLUG, "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods," *Journal of Economic Literature* 49 (September 2011), 615–51.

HOXBY, C. AND C. AVERY, "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students," *Brookings Papers on Economic Activity* 2013 (2013), 1–65.

HUGGETT, M., G. VENTURA AND A. YARON, "Sources of Lifetime Inequality," *American Economic Review* 101 (December 2011), 2923–54.

LEE, S. Y. T. AND A. SESHADRI, "On the Intergenerational Transmission of Economic Status," *Journal of Political Economy* 127 (2019), 855–921.

LEE, S. Y. T., A. SESHADRI AND N. ROYS, "Replication Files for "The Causal Effect of Parents' Education on Children's Earnings"," (Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2026-08-10), `https://doi.org/10.3886/E251556V2`.

LÉONARD, D. AND N. VAN LONG, *Optimal Control Theory and Static Optimization in Economics* (Cambridge: Cambridge University Press, 1992).

MCFADDEN, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed., *Frontiers in Econometrics* (New York: Academic Press, 1973), 105–142.

OREOPOULOS, P., M. PAGE AND A. H. STEVENS, "The Intergenerational Effects of Worker Displacement," *Journal of Labor Economics* 26 (2008), 455–483.

PLUG, E., "Estimating the Effect of Mother's Schooling on Children's Schooling Using a Sample of Adoptees," *The American Economic Review* 94 (2004), 358–368.

PLUG, E. AND W. VIJVERBERG, "Schooling, Family Background, and Adoption: Is It Nature or Is It Nurture?," *Journal of Political Economy* 111 (June 2003), 611–641.

POTERBA, J. M., "The rate of return to corporate capital and factor shares: new estimates using revised national income accounts and capital stock data," *Carnegie-Rochester Conference Series on Public Policy* 48 (1998), 211–246.

REGE, M., K. TELLE AND M. VOTRUBA, "Parental Job Loss and Children's School Performance," *The Review of Economic Studies* 78 (2011), 1462–1489.

SACERDOTE, B., "The Nature and Nurture of Economic Outcomes," *American Economic Review* 92 (2002), 344–348.

———, "How Large Are the Effects from Changes in Family Environment? A Study of Korean American Adoptees," *The Quarterly Journal of Economics* 122 (February 2007), 119–157.

———, "Nature and Nurture Effects On Children's Outcomes: What Have We Learned From Studies of Twins And Adoptees?," in J. Benhabib, A. Bisin and M. O. Jackson, eds., *Handbook of Social Economics*, volume 1A (Amsterdam: North-Holland, 2011), 1–30.

SCHOLZ, J. K., A. SESHADRI AND S. KHITATRAKUN, "Are Americans Saving "Optimally" for Retirement?," *Journal of Political Economy* 114 (2006), 607–643.

SOLON, G., "Intergenerational Mobility in the Labor Market," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, volume 3A (Amsterdam: Elsevier, 1999), 1761–1800.