

Online Appendix to “The Causal Effect of Parents’ Education on Children’s Earnings”

Sang Yoon (Tim) Lee

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

A Appendix

A.1 Proofs to Propositions 1 and 2

The proof requires a complete characterization of the income maximization problem. While we can use standard methods to derive the solution, here we simply guess and verify the value function. For notational convenience, we drop the age argument a unless necessary. Although we normalize $w = 1$ in the estimation, we keep it here for analytical completeness.

We separately characterize the solutions during the schooling and working phases in Lemmas 1 and 2. Schooling choice S is characterized as the solution to an optimal stopping time problem in Lemma 3. To this end, we guess that

$$\begin{aligned} V(a, h) &= q_2(a)h + v_W(a), & \text{for } a \in [6 + S, R), \\ V(a, h) &= q_1(a) \cdot \frac{h^{1-\alpha_1}}{1-\alpha_1} + e^{-r(6+S-a)} \cdot v(S, h_S), & \text{for } a \in [6, 6 + S), \text{ if } S > 0, \end{aligned}$$

where

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

for which the length of schooling S and level of human capital at age $6 + S$, h_S , are given, and v_W is some redundant function of age. These are the appropriate guesses for the solution, and the transversality condition becomes $q(R) = 0$. Now we characterize the working phase:

LEMMA 1: WORKING PHASE *Suppose optimal schooling satisfies $S \in [0, R - 6]$. Then given $h(6 + S) \equiv h_S$ and $q(R) = 0$, the solution satisfies, for $a \in [6 + S, R]$,*

$$(1) \quad q_2(a) = \frac{w}{r} \cdot q(a)$$

$$(2) \quad m(a) = \alpha_2 [\kappa q(a)z]^{\frac{1}{1-\alpha}}$$

$$(3) \quad h(a) = h_S + \frac{r}{w} \cdot \left[\int_{6+S}^a q(x)^{\frac{\alpha}{1-\alpha}} dx \right] \cdot (\kappa z)^{\frac{1}{1-\alpha}}$$

and

$$(4) \quad \frac{wh(a)n(a)}{\alpha_1} = \frac{m(a)}{\alpha_2},$$

where

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

Proof. Given that equation (10a) holds at equality, dividing by (10b) leads to equation (4), so once we know the optimal path of $h(a)$ and $m(a)$, $n(a)$ can be expressed explicitly. Plugging (10a) and the guess for the value function into equation (11), we obtain the linear, non-homogeneous first order differential equation

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

to which (1) is the solution. Using this result in (10a)-(10b) yields the solution for m , (2). Substituting (1), (2) and (4) into equation (5b) trivially leads to (3). $\square$

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

LEMMA 2: SCHOOLING PHASE *Suppose optimal schooling satisfies $S \in (0, R - 6)$. Then given $h(6) = h_0$ and $q_1(6) = q_0$, the solution satisfies, for $a \in [6, 6 + S)$,*

$$(5) \quad q_1(a) = e^{r(a-6)} q_0$$

$$(6) \quad m(a)^{1-\alpha_2} = \alpha_2 e^{r(a-6)} \cdot q_0 z$$

$$(7) \quad h(a)^{1-\alpha_1} = h_0^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2} \cdot \left[e^{\frac{\alpha_2 r(a-6)}{1-\alpha_2}} - 1 \right] \cdot (\alpha_2 q_0)^{\frac{\alpha_2}{1-\alpha_2}} z^{\frac{1}{1-\alpha_2}}.$$

Proof. Using the guess for the value function in (11) we have

$$\dot{q}_1(a) = r q_1(a),$$

to which solution is (5). Then equation (6) follows directly from (10b), and using this in (5b) yields the first order ordinary differential equation

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

to which (7) is the solution. □

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

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

$$(8) \quad q_0 = \frac{e^{-rS}}{\alpha_2^{\alpha_2}} \cdot \left([\kappa q(6+S)]^{1-\alpha_2} z^{\alpha_1} \right)^{\frac{1}{1-\alpha}}$$

$$(9) \quad h_S = \frac{\alpha_1}{w} \cdot [\kappa q(6+S) z]^{\frac{1}{1-\alpha}}.$$

Proof. The value matching for this problem boils down to setting $n(6 + S) = 1$ in the working phase, which yields (9).¹ The smooth pasting condition for this problem is

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

Using the guesses for the value functions, we have

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

and by replacing h_S with (9) we obtain (8). □

This, and the solutions for $[n(a)h(a), m(a)]$ during the working phase in Lemma 1 proves Proposition 2:

$$\begin{aligned} wh(a)(1 - n(a)) &= wh(a) - wh(a)n(a) \\ &= (\kappa z)^{\frac{1}{1-\alpha}} \left\{ \alpha_1 \left[q(6 + S)^{\frac{1}{1-\alpha}} - q(a)^{\frac{1}{1-\alpha}} \right] + r \cdot \int_{6+S}^a q(x)^{\frac{\alpha}{1-\alpha}} dx \right\} \\ &= (\kappa z)^{\frac{1}{1-\alpha}} \left\{ -\frac{\alpha_1}{1-\alpha} \cdot \int_{6+S}^a q(x)^{\frac{\alpha}{1-\alpha}} q'(x) dx + r \cdot \int_{6+S}^a q(x)^{\frac{\alpha}{1-\alpha}} dx \right\} \\ &= (\kappa z)^{\frac{1}{1-\alpha}} \left\{ \int_{6+S}^a q(x)^{\frac{\alpha}{1-\alpha}} \left[r - \frac{\alpha_1}{1-\alpha} q'(x) \right] dx \right\} \end{aligned}$$

after some simplification and inserting $q'(x) = -re^{-r(R-x)}$. We must still show Proposition 1.

Proof of Proposition 1. The length of the schooling period can be determined by plugging equations (8)-(9) into (7) evaluated at age $6 + S$:

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

with equality if $S > 0$. This implies that human capital accumulation must be positive in school-

¹This means that there are no jumps in the controls. When the controls may jump at age $6 + S$, we need the entire value matching condition.

ing, which is guaranteed by the law of motion for human capital. Rearranging terms,

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

or now replacing $h_0 \equiv z^\lambda h_p^\nu$,

$$(10) \quad z^{1 - \lambda(1 - \alpha)} h_p^{-\nu(1 - \alpha)} \leq F(S),$$

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

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

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

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

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

and S is determined by (10) at equality. The full solution is given by Lemmas 1-3 and we obtain Proposition 2. Otherwise $S = 0$ and the solution is given by Lemma 1. $\square$

A.2 Identifying b_{select} and λ

Applying (6) on (13) in Proposition 2 yields

$$(11) \quad \log e_{i,x} = \log C(x; S_i) + \frac{Z_i}{1 - \alpha} + b_{\text{select}} \cdot S_{Pi} + \text{error}.$$

where $e_{i,x}$ is individual i 's earnings at x years of experience, and the selection effect of parents' schooling on children's earnings, b_{select} , was defined in (15b). Since Z_i is uncorrelated with S_{Pi} , if we knew the function $C(\cdot)$ we could regress log-earnings on parents' schooling, controlling for experience and own schooling, to identify b_{select} .

A byproduct of Corollary 1 is that the schooling-specific experience effect $\varphi_{s,x}$ should capture the functions $C(x;s)$, $F(s)$ and $Z_{S_i=s}$ (Equation 17). If we were able to separate the two, either (15b) or (11) could be used to infer b_{select} . However, this is generally infeasible for both theoretical and empirical reasons. Theoretically, $\log C(0;s)$ diverges to $-\infty$ in the model, but $x = 0$ is required for identification. And in the data, not only do we not observe zero earnings but neither do we observe early-experience earnings for individuals with low levels of schooling.

To gain some traction independently of the structural estimation in Section 4, take an n th-order approximation around (17):

$$(12) \quad \log C(x;s) + \frac{\log F(s)}{(1-\alpha)[1-\lambda(1-\alpha)]} + \frac{\beta_{h_p Z} Z_s}{1-\lambda(1-\alpha)} \\ = A_0 + \sum_{j=1}^n A_j^s (s - \hat{s})^j + \sum_{m=1}^n \sum_{j=0}^m A_{mj}^x (s - \hat{s})^j (x - \hat{x})^{m-j} + \dots$$

where $(\hat{x}, \hat{s})$ are some reference levels of experience and schooling, and (A_j^s, A_{mj}^x) are some constants. Then instead of including a full set of schooling-specific experience effects as in (16), regress

$$(13) \quad \log e_{i,x}^* = a_0 + \sum_{j=1}^n a_j^s S_i^j + \sum_{m=1}^n \sum_{j=0}^m a_{mj}^x S_i^j x^{m-j} + b S_{P_i} + \tilde{u}_{i,x}$$

where (a_j^s, a_{mj}^x) are coefficients to be estimated. If the approximation is accurate, the estimate $\hat{b}$ would still capture b_{spill} . More importantly, the first sum captures all the variation coming from $F(s)$ and Z_s , which are only functions of s , and the second sum the variation from $C(\cdot)$. Thus

COROLLARY 1: IDENTIFYING SELECTION. *Suppose the assumptions in Corollary 1 hold, the approximation (12) is accurate, and we estimate (13). Then*

1. *Regressing $\sum_{j=1}^n \hat{a}_j^s S_i^j$ on S_{P_i} admits an unbiased estimate for $b_{\text{select}} - b_{\text{spill}}$.*
2. *Regressing $(\log e_{i,x}^* - \sum_{m=1}^n \sum_{j=0}^m a_{mj}^x S_i^j x^{m-j})$ on S_{P_i} admits an unbiased estimate for b_{select} .*

Once we know b_{spill} from Corollary 1, Part 1 can be used to back out b_{select} . Part 2 implies that once we filter out an individual's schooling-specific experience profile, the coefficient on parents' schooling S_P captures only b_{select} .² This may be somewhat surprising, but a direct consequence of

²Both Parts 1 and 2 only work because Z is uncorrelated with S_{P_i} .

Proposition 2: (ν, λ) only affect the child's initial level of human capital, of which effect is absorbed into the length of schooling, which in turn only affects the shape of experience profiles.³

According to our model then, the difficulty in separating the two effects stems from how we control for own schooling S . In Corollary 1, the coefficient on parents' schooling identifies b_{spill} because when holding own schooling S fixed, the regression reveals the relationship that z must have with S_P to be consistent with the observed level of S . But if we could directly control for how schooling affects earnings as in Corollary 1, the regression reveals the sample correlation that z must have with S_P . This intuition is visualized in Appendix Figure A.1. The parental spillover is understood as a child of a more educated parent getting more out of the same level of schooling (i.e., the quality of schooling). And within a sample of children with the same level of schooling, we can recover the combinations of parents' schooling and children's earnings that are consistent with the observed level of schooling, assuming that the spillover effect is subsumed in the choice of schooling.

In Table 3, we showed that the estimates of b_{spill} are robust to alternative specifications. In particular, columns (7)-(8) included 3rd and 4th-order polynomials of (x, S_i) . We use these estimates to compute the implied values of b_{select} via the two methods suggested by Corollary 1.

A formal test rejects equality of parts 1 and 2, but the magnitudes are strikingly similar. Similarly, the estimates barely change from a 3rd to 4th order polynomial, suggesting that higher order terms are quantitatively irrelevant. All in all, the results imply that the selection effect is almost 6 times larger than the spillover effect. However, as we find in Section 4, most of selection is in fact driven by preferences rather than abilities.

Similarly, we can use Corollary 1 to estimate λ :

COROLLARY 2 *Suppose the assumptions in Corollary 1 hold, the approximation (12) is accurate, and we obtain estimates from (13). If we regress*

$$(14) \quad \log e_{i,x}^* - \sum_{m=1}^n \sum_{j=0}^m \hat{a}_{mj}^x S_i^j x^{m-j} = \left[\vartheta_{S_{Pi}} + b_{S_{Pi}} \sum_{j=1}^n \hat{a}_j^S S_i^j \right] + \tilde{u}_{ix}$$

³Of course, the result is an outcome of our model. The advantage of this model implication is that we can establish a lower-bound as opposed to an upper-bound for causal effects. Note that our identification scheme cannot speak to the general returns to an individual's own schooling, but only his parent's schooling.

	1.	2.
3rd order	0.114 (37.72)	0.106 (117.40)
R^2	0.218	0.133
4th order	0.115 (38.00)	0.109 (119.11)
R^2	0.218	0.136
Sample	5,114	89,893

Table A.1: Implied Selection Coefficient

Selection coefficient defined in (15b), implied by Corollary 1. Columns (1) and (2) correspond to parts 1 and 2 of the corollary, respectively. The 1st (2nd) panel uses coefficients from a regression of (log) earnings on moms' years of schooling and a 3rd (4th) order polynomial of potential experience (age-6- S) and own schooling. HRS initial cohort, males born 1924-1941, ages 23-42. All columns include a linear and quadratic for potential experience (age-6- S). t -stats shown in parentheses.

where $\{\vartheta_{sp}\}$ is a set of parents' schooling fixed effects, and $\tilde{u}_{i,x}$ regression error, the estimates

$$\hat{b}_{S_p} = b_{S_p} = 1 / \{(1 - \alpha) [1 - \lambda(1 - \alpha)]\}.$$

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

$$(15) \quad \frac{\log z_i}{1 - \alpha} \propto \frac{\log F(S_i)}{(1 - \alpha)[1 - \lambda(1 - \alpha)]} + \frac{\nu \beta_{hpZ} Z_{S_i}}{1 - \lambda(1 - \alpha)}.$$

Applying this and $S_{pi} = \hat{S}_p$ in (13) yields

$$\log e_{i,x} - \log C(x; S_i) = K_{\hat{S}_p} + \frac{\log F(S_i)}{(1 - \alpha)[1 - \lambda(1 - \alpha)]} + \frac{\nu \beta_{hpZ} Z_{S_i}}{1 - \lambda(1 - \alpha)},$$

among individuals with $S_{pi} = \hat{S}_p$, where $K_{\hat{S}_p}$ is some constant. The term in brackets in (14) apply only to individuals whose parents have the same level of schooling and the sum on the left-hand side in (14) captures the function $\log C$, while the sum on the right-hand side captures exactly (15). Thus b_{S_p} is estimated as claimed. $\square$

Corollary 1 shows that among children whose *parents* attained the same years of schooling, the magnitude of λ is identified by the within-group variation in a function of log-earnings and a function of own schooling. So we can use C and the weighted sum of $\log F, Z_S$ recovered from

(13) to perform the suggested regression. The intuition is similar as Corollary 1, except that we focus on the within-group variation among different groups.

For children with identical S_p , earnings differences cannot be attributed to causal effects from parents. The only way that schooling can have heterogeneous effects on earnings of individuals whose parents have the same S_p is through z 's influence on $F(S)$, which reveals λ (given α).

Intuitively, we can identify the three parameters $(\nu, \lambda, b_{\text{select}})$ because we have 3 sets of observations: earnings e and own schooling S , and parents' schooling S_p . Ignoring experience, the model imposes two non-linear, monotonic relationships $e = e(S, z)$ and $S = S(S_p, z)$, and given an empirical distribution of (h_p, z_p) , we additionally parameterized the dependence of z on z_p . Therefore, all three parameters can be separately identified.

A.3 Numerical Algorithm

We solve the model in discrete time. A period is 2 years. A mom has a child at age 24, and her human capital stays constant till the child leaves (finished education).

1. The periods are as follows:

Period	1	2	3	4	5	6	7	8	9	10	11	12	13
Son's age	0	2	4	6	8	10	12	14	16	18	20	22	24
Mom's Age	24	26	28	30	32	34	36	38	40	42	44	46	48

Periods 8-13 are the possible terminal nodes for the son to leave the mom, choosing $S \in \{8, 10, 12, 14, 16, 18\}$ years of schooling.

2. In period 1, we set the grid for (y, h_p, z_p) as follows. For each level of S_p in AHEAD, we obtain the empirical distribution of $(y_i(0), h_{p_i}, z_{p_i})$ as described in Section 4.1. Let μ_y and σ_y^2 denote the mean and variance of $\log y_i(0)$. For computation, we set a grid of size N_y for assets and choose grid points $\bar{y}_j$ so that $\Phi(\log \bar{y}_1) = 1 - \Phi(\log \bar{y}_{N_y}) = 1/2N_y$ and $\Phi(\log \bar{y}_{j+1}) - \Phi(\log \bar{y}_j) = 1/N_y$. For z_p , we choose the grid points so that an equal number of observations are captured in each interval.
3. From period 4 onward, we set the lower bound of the y -grid to 0 and allow the distance

between grids to grow at a cubic rate. The child's ability z -grid is set similarly. Note that from period 4 onward, the parent's ability is no longer a state.

4. From period 5 onward, children's human capital grids are separately set for each grid point on the z grid. The lower- and upper-bounds in period 4 are determined by the minimum and maximum level levels of h_P . From period 5 onward, we allow the upper-bound of the h -grid to expand assuming that $\bar{y}_{N_y}/5$ was invested into accumulating human capital. Interim grid points are set so that the distance between grids grows at a cubic rate.
5. The model is solved by backward induction. At any point that the child leaves school and becomes independent, first solve for $V_{S'}$ in (5). We can then solve for the optimal intergenerational transfer choice in (4) using the endogenous grid method (EGM, Carroll 2006).
6. Solving (2)-(3) involves two choice variables (y', h') , the savings and human capital choice for 2 years ahead, and also a discrete choice. To facilitate computation, we use the multivariate "G²EGM" developed by Druedahl and Jørgensen (2017). Since we assume that the random component in (4) is logistic, the expectation in 3 can be solved for using the logsum formula in McFadden (1973).
7. Finally (1) can be solved by standard EGM yielding the optimal choice for y' , using a transition matrix from z_P to z to compute expectations.

Computing Model Moments Given our distributional assumptions on mom's schooling, learning abilities and preferences for schooling, we can compute the exact model implied moments as follows.

1. In period 1, we know the empirical distribution of moms over the (y, h_P, z_P) -grid from women in AHEAD, as described above and in Section 4.1. The optimal choice of y' from each node will lie in an interval $[\bar{y}_j, \bar{y}_{j+1})$ where $\bar{y}_j$ is the j th point on next period's y -grid. So we assume that fractions (π_j, π_{j+1}) land on $(\bar{y}_j, \bar{y}_{j+1})$, respectively, where π_j is the linear interpolation weight on $\bar{y}_j$.
2. In periods 4, for each node on the (y, h_P, z) grid, we can solve for children's initial human capital h exactly. The optimal choices of (y', h') will lie in a rectangle spanned by $[\bar{y}_j, \bar{y}_{j+1})$

and $[\bar{h}_k, \bar{h}_{k+1})$ where $(\bar{y}_j, \bar{h}_k)$ are the j th and k th points on next period's y - and h -grids, respectively. So we assume that a fraction π_{jk} lands on $(\bar{y}_j, \bar{h}_k)$ where π_{jk} is the bilinear interpolation weight on $(\bar{y}_j, \bar{h}_k)$, and similarly for the three other points. Note that the states (h_p, z) are permanent and do not change with age.

3. In periods 5-7, we follow the same method as above to compute the fraction of mom-son pairs landing on next period's (y, h) -grid, for each node on the (y, h, h_p, z) grid.
4. In periods 8-12, for each node on the (y, h, h_p, z) grid, we first use the conditional choice probabilities (CCP) to compute the fraction of mom-son pairs that decide to exit schooling using the CCP formula in McFadden (1973). For those that continue, we follow the same method as above to compute the fraction of mom-son pairs landing on next period's (y, h) -grid.

All model moments are computed by aggregating over the empirical distribution of the initial state (y, h_p, z_p) , using the z_p -to- z transition probabilities, interpolation weights and CCPs as weight for all later ages.

A.4 Tables and Figures Not in the Text

Mom's S_P	Fraction (%)	Child's S	Average S
≤ 5	12.25	≤ 11	6.34
		≥ 12	13.30
6-7	12.10	≤ 11	8.02
		≥ 12	13.48
8	21.90	≤ 12	10.67
		≥ 13	15.21
9-11	13.66	≤ 12	10.88
		≥ 13	15.27
12	30.32	≤ 12	11.17
		≥ 13	15.34
≥ 13	9.78	≤ 12	11.18
		≥ 13	15.87

Table A.2: Summary 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. In the third column, $\bar{S}$ denotes the average years of schooling attained in each category. These numbers can be compared with the target moments in Table 5.

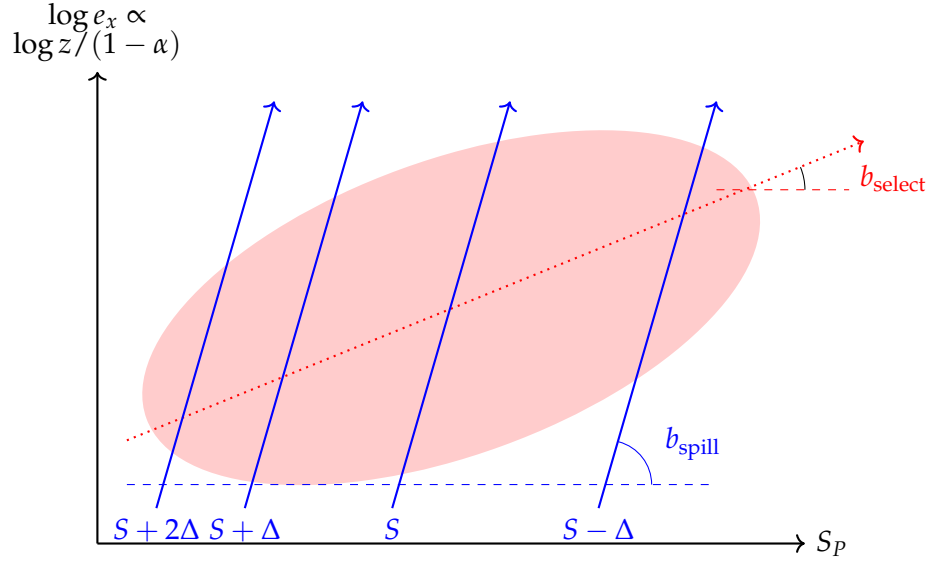


Figure A.1: Intuition for Identifying Ability Selection and Spillovers.

x -axis: parents' human capital, or their schooling levels, y -axis: children's log abilities, or log-earnings controlling for own schooling. The light gray area represents the distribution of abilities and parents' schooling, (z, S_P) . According to our model, selection is captured by the population correlation between parent's schooling and earnings, controlling for schooling (given knowledge of the earnings function C , defined in Proposition 2). This is captured by the slope of the gray dotted line. Spillovers are captured by the relationship between earnings and abilities among children with the same level of schooling, which is captured by the slope of the dark gray lines. Note that conditional on abilities, schooling is decreasing in parents' schooling.

Mom's S_P	Child's S	Fraction (%)	Average Earnings at Age			
			23-27	28-32	33-37	38-42
≤ 5	≤ 11	57.70	1.00	1.28	1.55	1.83
	≥ 12	42.30	1.03	1.58	2.13	2.68
6-7	≤ 11	40.97	1.03	1.29	1.55	1.80
	≥ 12	59.03	1.07	1.61	2.14	2.67
8	≤ 12	70.96	1.14	1.53	1.93	2.33
	≥ 13	29.04	1.02	1.65	2.25	2.85
9-11	≤ 12	59.57	1.20	1.61	2.02	2.43
	≥ 13	40.43	1.10	1.74	2.35	2.96
12	≤ 12	40.07	1.32	1.76	2.19	2.63
	≥ 13	59.93	1.24	1.92	2.56	3.21
≥ 13	≤ 12	23.66	1.48	1.94	2.39	2.85
	≥ 13	76.35	1.44	2.18	2.87	3.56
(S, S_P) correlation and OLS:			0.48	0.42		
Mincer coefficients on (S, S_P) :			0.04	0.02		

Table A.3: Implied 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 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.

Child's S					8	10	12	14	16	18
Fraction (%)					15.56	9.00	34.10	15.04	13.77	12.53
Mom's Age	30	32	34	36	38	40	42	44	46	48
UE (%)	14.90	10.17	8.53	7.20	5.09	4.58	1.40	0.72	0.21	-

Table A.4: Implied Moments, Continued

The top panel shows the fraction of children in the model who attain S years of schooling. The bottom panel shows the fraction of moms in the model who make zero earnings.

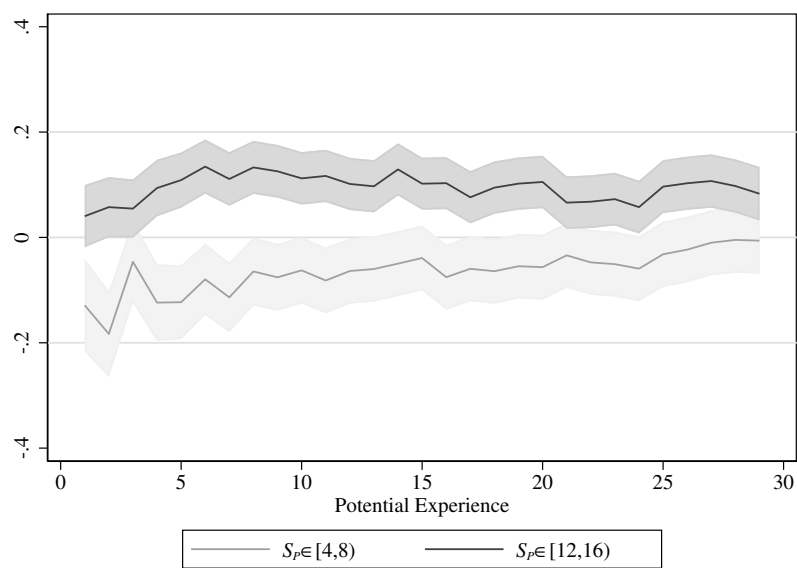


Figure A.2: Gaps between Average Earnings Profiles, All

Gaps between the earnings profiles of children by different levels of moms' schooling: 1924-1941 birth cohort. The y -axis is average log annual earnings in 2008 USD. Mothers' schooling levels are divided into 4 to 7 years and 12 or more years, and the light and dark lines correspond to the difference in the profiles compared to those whose moms attained 8 to 11 years of schooling.

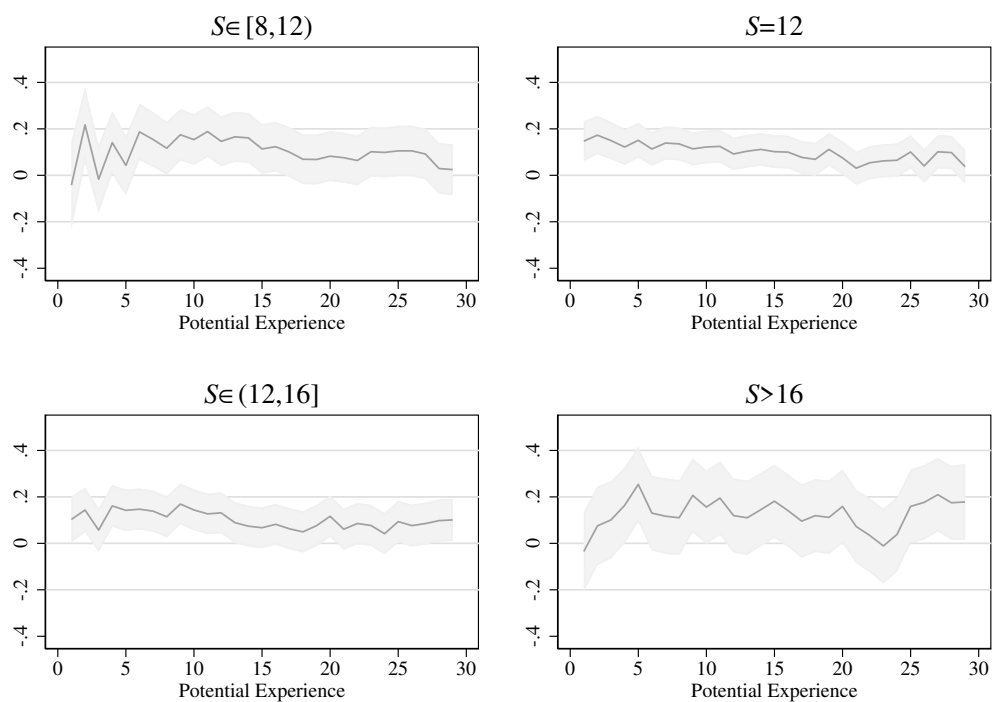


Figure A.3: Gaps between Average Earnings Profiles, All

Gaps between the earnings profiles of children by different levels of moms' schooling: 1924-1941 birth cohort. The y -axis is average log annual earnings in 2008 USD. Mothers' schooling levels are divided into 8 years or below, and more than 8 years. The lines correspond to the latter minus the former.

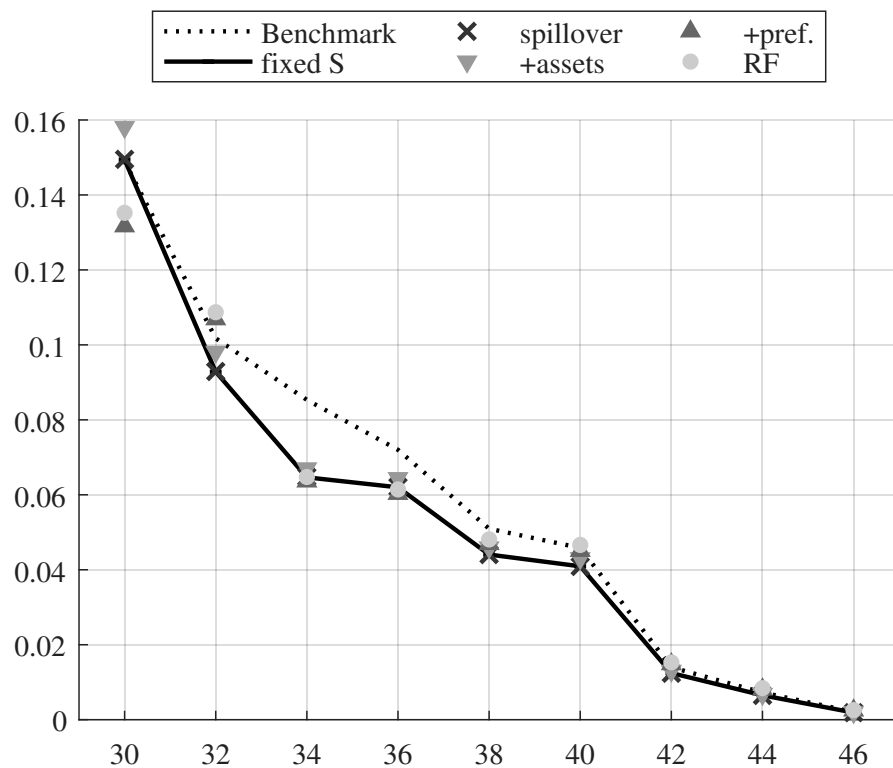


Figure A.4: Moms' Average Time Spent with Children

x -axis: mom's age; y -axis: average fraction of time spent with children. Refer to Section 5.1 for details of each scenario.