For Online Publication

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

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

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

where

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

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

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

$$
q_2(a) = \frac{w}{r} \cdot q(a) \tag{26}
$$

$$
m(a) = \alpha_2 \left[\kappa q(a) z\right]^{\frac{1}{1-\alpha}} \tag{27}
$$

$$
h(a) = h_S + \frac{r}{w} \cdot \left[\int_{6+S}^{a} q(x)^{\frac{\alpha}{1-\alpha}} dx\right] \cdot (\kappa z)^{\frac{1}{1-\alpha}} \tag{28}
$$

and

$$
\frac{wh(a)n(a)}{\alpha_1} = \frac{m(a)}{\alpha_2}, \tag{29}
$$

where

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

Proof. Given that equation (5) holds at equality, dividing by (6) leads to equation (29), so once we know the optimal path of $h(a)$ and $m(a)$, $n(a)$ can be expressed explicitly. Plugging (5) and the guess for the value function into equation (7), we obtain the linear, non-homogeneous first order differential equation

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

to which (26) is the solution. Using this result in (5)-(6) yields the solution for m, (27). Substituting (26), (27) and (29) into equation (1b) trivially leads to (28). □

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

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

$$q_1(a) = e^{r(a-6)} q_0 \tag{30}$$

$$m(a)^{1-\alpha_2} = \alpha_2 e^{r(a-6)} \cdot q_0 z \tag{31}$$

$$h(a)^{1-\alpha_1} = h_0^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2} \cdot \left[e^{\frac{\alpha_2 r(a-6)}{1-\alpha_2}} - 1 \right] \cdot (\alpha_2 q_0)^{\frac{\alpha_2}{1-\alpha_2}} z^{\frac{1}{1-\alpha_2}}. \tag{32}$$

Proof. Since $n(a) = 1$ during the schooling phase, using the guess for the value function in (7) we have

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

to which solution is (30). Then equation (31) follows directly from (6), and using this in (1b) yields the first order ordinary differential equation

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

to which (32) 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*

$$
\begin{aligned}
q_0 &= \frac{e^{-rS}}{\alpha_2^{\alpha_2}} \cdot \left([\kappa q(6+S)]^{1-\alpha_2} z^{\alpha_1} \right)^{\frac{1}{1-\alpha}} && (33)\\
h_S &= \frac{\alpha_1}{w} \cdot [\kappa q(6+S) z]^{\frac{1}{1-\alpha}}. && (34)
\end{aligned}
$$

Proof. The value matching for this problem boils down to setting $n(6+S) = 1$ in the working phase, which yields (34).[30] The smooth pasting condition for this problem is

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

Using the guesses for the value functions, we have

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

and by replacing h_S with (34) we obtain (33). □

This proves proves Proposition 2, and the solutions for $n(a)h(a)$ and $m(a)$ during the working phase in Lemma 1 proves Corollary 1. We must still show Proposition 1.

Proof of Proposition 1. The length of the schooling period can be determined by plugging equations (33)-(34) into (32) evaluated at age $6+S$:

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

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

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

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

$$
z^{1-\lambda(1-\alpha)} h_P^{-\nu(1-\alpha)} \leq F(S), \tag{35}
$$

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

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

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

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

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

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

and S is determined by (35) at equality. The full solution is given by Lemmas 1-3 and we obtain Proposition 2 and Corollary 1. Otherwise $S = 0$ and the solution is given by Lemma 1. □

B Analytical Characterization of the Extended Model

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

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

$$V(a,h) = \frac{w}{r} \cdot q(a)h + D_W(a) \tag{36}$$

and the solution is characterized by

$$n(a)h(a) = \left[\frac{\alpha_W}{r} \cdot q(a)z\right]^{\frac{1}{1-\alpha_W}} \tag{37}$$

$$h(a) = h_J + \left(\frac{\alpha_W}{r}\right)^{\frac{\alpha_W}{1-\alpha_W}} \cdot \left[\int_{6+S+J}^{a} q(x)^{\frac{\alpha_W}{1-\alpha_W}} dx\right] \cdot z^{\frac{1}{1-\alpha_W}}, \tag{38}$$

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

$$V(a,h) = e^{r(a-6-S)} q_S \cdot \frac{h^{1-\alpha_W}}{1-\alpha_W} + e^{-r(6+S+J-a)} D(J, h_J) \tag{39}$$

where

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

while human capital evolves as

$$h(a)^{1-\alpha_W} = h_S^{1-\alpha_W} + (1-\alpha_W)(a-6-S)z. \tag{40}$$

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

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

$$h_J = \left[\frac{\alpha_W}{r} \cdot q(6+S+J) z \right]^{\frac{1}{1-\alpha_W}}. \tag{42}$$

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

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

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

$$\frac{z}{h_S^{1-\alpha_W}} \le G(J) \equiv \left[\frac{\alpha_W}{r} \cdot q(6+S+J) - (1-\alpha_W)J \right]^{-1}. \tag{43}$$

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

$$G(0) < \frac{z}{h_S^{1-\alpha_W}} \qquad \Leftrightarrow \qquad \frac{r}{\alpha_W q(6+S)} < \frac{z}{h_S^{1-\alpha_W}}, \tag{44}$$

and J is determined by (43) at equality. Otherwise $J = 0$.

Now if S were discrete, as in the model we estimate, we only need to solve for h_S, the level of human capital at age $6+S$. Then we can solve for $V(h_0, z; s)$ for all 6 possible values of s, using Lemmas 2 and 4 for the schooling and working phases, respectively. But it is also possible to characterize the unconstrained continuous choice of S, even though a closed form solution does not exist in general. We only need consider new value matching and smooth pasting conditions.

Lemma 5: Schooling Phase, Extended *The length of schooling, S, and level of human capital at age $6+S$, h_S, are determined by*

1. *if $J = 0$,*

$$\epsilon + (1-\alpha_2)\left[\frac{\alpha_2^{\alpha_2} w}{r}\cdot q(6+S) z h_S^{\alpha_1}\right]^{\frac{1}{1-\alpha_2}} = w\cdot\left(h_S + (1-\alpha_W)\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S) z\right]^{\frac{1}{1-\alpha_W}}\right) \tag{45}$$

$$h_S^{1-\alpha_1} \le h_0^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2}\cdot\left(1-e^{-\frac{\alpha_2 rS}{1-\alpha_2}}\right)\cdot\left[\frac{\alpha_2 w}{r}\cdot q(6+S) h_S^{\alpha_1}\right]^{\frac{\alpha_2}{1-\alpha_2}} z^{\frac{1}{1-\alpha_2}} \tag{46}$$

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

$$q_0 = \frac{we^{-rS}}{r}\cdot q(6+S) h_S^{\alpha_1}. \tag{47}$$

2. *if $J > 0$,*

$$\epsilon + (1-\alpha_2)\left(\alpha_2^{\alpha_2} we^{-rJ}\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S+J) z\right]^{\frac{1}{1-\alpha_W}}\cdot h_S^{\alpha_1-\alpha_W}\right)^{\frac{1}{1-\alpha_2}}$$
$$= we^{-rJ}\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S+J) z\right]^{\frac{1}{1-\alpha_W}} \tag{48}$$

$$h_S^{1-\alpha_1} \le h_0^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2}\cdot\left(1-e^{-\frac{\alpha_2 rS}{1-\alpha_2}}\right) \tag{49}$$
$$\cdot\left(\alpha_2 we^{-rJ}\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S+J) z^{\alpha_W}\right]^{\frac{1}{1-\alpha_W}} h_S^{\alpha_1-\alpha_W}\right)^{\frac{\alpha_2}{1-\alpha_2}}\cdot z^{\frac{1}{1-\alpha_2}}$$

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

$$q_0 = we^{-r(S+J)}\cdot\left[\frac{\alpha_W^{\alpha_W}}{r}\cdot q(6+S+J) z^{\alpha_W}\right]^{\frac{1}{1-\alpha_W}}\cdot h_S^{\alpha_1-\alpha_W}. \tag{50}$$

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

$$\epsilon - m(6+S) + e^{rS} q_0 z m(6+S)^{\alpha_2} = wh_S\left[1-n(6+S)\right] + \frac{w}{r}\cdot q(6+S) z\left[n(6+S) h_S\right]^{\alpha_W}$$
$$e^{rS} q_0 h_S^{-\alpha_1} = \frac{w}{r}\cdot q(6+S).$$

Hence (47) follows from the smooth pasting condition. Likewise, (45) follow from plugging $n(6+S)$, $m(6+S)$ from Lemmas 2 and 4 and q_0 from (47) in the value matching condition. Lastly, (46) merely states that the optimal h_S must be consistent with optimal accumulation in the schooling phase, $h(6+S)$.

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

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

Hence (50) follows from plugging q_S and h_S from (41)-(42) in the smooth pasting condition. Likewise, (48) follow from plugging $n(6+S) = 1$, $m(6+S)$ from Lemma 2, and q_0 from (50) in the value matching condition. Again, (49) requires consistency between h_S and $h(6+S)$. $\square$

For each case where we assume $J = 0$ or $J > 0$, it must also be the case that condition (44) does not or does hold.

C Numerical Algorithm

For the purposes of our estimated model in which S is fixed, the solution method in Appendix B is straightforward. We need not worry about value-matching conditions and only need to solve the smooth-pasting conditions given S, which are equations (46) and (49), to obtain h_S. Note that there is always a solution to (46) or (49)—i.e., we can always define a function $h_S(S)$ as a function of S. This is seen by you rearranging the equations as (bold-face for emphasis)

$$1 = \left(\frac{h_0}{\mathbf{h_S}}\right)^{1-\alpha_1} + \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2} \cdot \left(1 - e^{-\frac{\alpha_2 rS}{1-\alpha_2}}\right) \cdot \left[\frac{\alpha_2 w}{r} \cdot q(6+S)\right]^{\frac{\alpha_2}{1-\alpha_2}} z^{\frac{1}{1-\alpha_2}} \cdot \mathbf{h_S}^{-\frac{1-\alpha}{1-\alpha_2}} \tag{51}$$

$$\begin{aligned}
1 = \left(\frac{h_0}{\mathbf{h_S}}\right)^{1-\alpha_1} &+ \frac{(1-\alpha_1)(1-\alpha_2)}{r\alpha_2} \cdot \left(1 - e^{-\frac{\alpha_2 rS}{1-\alpha_2}}\right) \\
&\cdot \left(\alpha_2 w e^{-rJ} \left[\frac{\alpha_W^{\alpha_W}}{r} \cdot q(6+S+J) z^{\alpha_W}\right]^{\frac{1}{1-\alpha_W}}\right)^{\frac{\alpha_2}{1-\alpha_2}} \cdot z^{\frac{1}{1-\alpha_2}} \cdot \mathbf{h_S}^{-\frac{1-\alpha+\alpha_2\alpha_W}{1-\alpha_2}},
\end{aligned} \tag{52}$$

respectively. Hence, for any given value of S, both RHS's begin at or above 1 at $h_S = h_0$, goes to 0 as $h_S \to \infty$, and is strictly decreasing in h_S. The solution $h_S(S)$ to both (51) and (52) are such that

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

The rest of the model can be solved by Lemmas 2 and 4, and we can use Lemma 4 to determine J. Depending on whether condition (44) holds, we may have two solutions:

1. If only one solution satisfies (44), it is the solution.
2. If both satisfy (44), compare the two value functions at age 6 given S and candidate solutions

$J_1 = 0$ and $J_2 > 0$ from Lemma 4 using the fact that the function D_W in (36) can be written

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

and

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

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

Computing Model Moments Given our distributional assumptions on mother's schooling, learning abilities and tastes for schooling, we can compute the exact model implied moments as follows. We set grids over h_P ,z, and S, with $N_{h_P} = 17$, $N_z = 100$ and $N_S = 6$ nodes each.

1. Construct a grid over all observed levels of S_P in the data. This varies from 0 to 16 with mean 9.26 and standard deviation 3.52. Save the p.m.f. of S_P to use as sampling weights.

2. Assuming $\beta = 0.06$, construct the h_P-grid which is just a transformation of the S_P-grid according to (9).

3. For each node on the h_P-grid, construct z-grids according to (20), according to Kennan (2006). This results in a total of $N_{h_P} \times N_z$ nodes and probability weights, where for each h_P node we have a discretized normal distribution.

4. For each (h_P, z) compute the pecuniary of choosing $S \in \{8, 10, 12, 14, 16, 18\}$ (solve for $V(S;6,h_0)$ according to the above) and compute the fraction of individuals choosing each schooling level using the CCP's in (22)-(23).

All moments are computed by aggregating over the $N_{h_P} \times N_z \times N_S$ grids using the product of the empirical p.m.f. of h_P, the discretized normal p.d.f. of z, and CCP's of S as sampling weights.

D Formal Description of Experiments in Section 5

Formally, for any initial condition $x = (S_P, \log z, \xi)$, the model implied schooling and age-a earnings outcomes can be written as functions of x, $S = S(x)$, $E(a) = \tilde{E}(x;S;a)$.[31] Then schooling following a j-year increase in S_P, holding (z,ξ) constant, is

$$S_v^j(x) \equiv S(S_P + j, z, \xi). \tag{53}$$

[31] Since although the parent variable in the initial condition is h_P, it is defined as $\log h_P = \beta S_P$ in (9).

Age a earnings following a j-year increase in S_P, holding (z, ξ) and S constant, is

$$E_0^j(x;a) \equiv \tilde{E}(S_P + j, z, \xi; S; a)|_{S=S(S_P,z,\xi)} \tag{54}$$

i.e., the schooling choice is fixed as if mother's education is S_P, but the earnings outcome, or amount of human capital accumulated, is computed assuming that mother's education is $S_P + j$. This captures the spillover effect that is independent of quantity (schooling) adjustment. Now if we define

$$E(x;a) \equiv \tilde{E}(S_P, z, \xi; S; a)|_{S=S(S_P,z,\xi)},$$

i.e. the earnings outcome when both the amount of human capital accumulation and the schooling choice are computed from the same level of S_P, we can write the total spillover effect as

$$E_v^j(x;a) \equiv \tilde{E}(S_P + j, z, \xi; a), \tag{55}$$

which also includes the substitution effect between the length and quality of schooling. Selection on abilities and tastes associated with a j-year increase in S_P can be written as

$$\begin{aligned}
\Delta_z^j &\equiv \exp\left[(\rho_{zh_P}\sigma_z/\sigma_{h_P}) \cdot \beta j\right] \\
\Delta_\xi^j(S_P) &\equiv \left\{\Delta_\xi^j(S;S_P)\right\}_S \equiv \left\{\delta_S \gamma_{h_P} \exp(\beta S_P)\left[\exp(\beta j) - 1\right]\right\}_S,
\end{aligned}$$

respectively, where $\Delta_\xi^j(S_P)$ is a 6-dimensional vector for each level of schooling $S \in \{8, \ldots, 18\}$. The first expression follows since $(S_P, \log z)$ are joint-normal, and the second from the definition of tastes in (21). Then schooling and age a earnings following a j-year increase in S_P, including partial selection effects on z or ξ, are

$$S_z^j(x) \equiv S(S_P + j, z \cdot \Delta_z^j, \xi), \qquad E_z^j(x;a) \equiv E(S_P + j, z \cdot \Delta_z^j, \xi; a), \tag{56}$$

$$S_\xi^j(x) \equiv S(S_P + j, z, \xi + \Delta_\xi^j(S_P)), \qquad E_\xi^j(x;a) \equiv E(S_P + j, z, \xi + \Delta_\xi^j(S_P); a). \tag{57}$$

Outcomes incorporating all spillover and selection effects following a j-year increase are

$$S_{rf}^j(x) \equiv S(S_P + j, z \cdot \Delta_z^j, \xi + \Delta_\xi^j(S_P) + \Delta_{z\xi}^j(z)) \tag{58a}$$

$$E_{rf}^j(x;a) \equiv E(S_P + j, z \cdot \Delta_z^j, \xi + \Delta_\xi^j(S_P) + \Delta_{z\xi}^j(z); a), \tag{58b}$$

where $\Delta_{z\xi}^j(z) \equiv \left\{\Delta_{z\xi}^j(S;z)\right\}_S = \left\{\delta_S \gamma_z z \left[\Delta_z^j - 1\right]\right\}_S$ is a compounded selection effect on tastes that comes from (z, ξ) being correlated, even conditional on h_P. We coin this the "reduced form" effect since by construction,

$$\int S_{rf}^j(x) d\Phi(\hat{S}_P = S_P, \hat{z}, \hat{\xi}) = \int S(x) d\Phi(\hat{S}_P = S_P + j, \hat{z}, \hat{\xi}),$$

where Φ is the joint distribution over x, and $\hat{x}$ are dummies for integration.

The first row of Table 9 is obtained by integrating the change from $S(x)$ in (53) and (56)-(58) over the population distribution Φ, when $j = 1$. The second row is the outcome of

$$\log\left[\sum_{a=14}^{R}\left(\frac{1}{1+r}\right)^{a-14}\int E_k^1(x;a)d\Phi(x)\right] - \log\left[\sum_{a=14}^{R}\left(\frac{1}{1+r}\right)^{a-14}\int E(x;a)d\Phi(x)\right]$$

for $k \in \{0, \nu, z, \xi, rf\}$. Figure 3 is obtained by plotting

$$\log\left[\int E_k^1(x;a)d\Phi(x)\right] - \log\left[\int E(x;a)d\Phi(x)\right], \qquad \text{for } k \in \{0, \nu, z, \xi, rf\},$$

and each bar in the left and right panels of Figure **??** plots, respectively,

$$\log\left[\int S_k^1(x)d\Phi(\hat{S}_P \in M, \hat{z}, \hat{\xi})\right] - \log\left[\int S(x)d\Phi(\hat{S}_P \in M, \hat{z}, \hat{\xi})\right]$$

for the distribution in change in schooling S, and

$$\log\left[\sum_{a=14}^{R}\frac{\int E_k^1(x;a)d\Phi(\hat{S}_P \in M_s, \hat{z}, \hat{\xi})}{(1+r)^{a-14}}\right] - \log\left[\sum_{a=14}^{R}\frac{\int E(x;a)d\Phi(\hat{S}_P \in M_s, \hat{z}, \hat{\xi})}{(1+r)^{a-14}}\right],$$

for the distribution in the change in average lifetime earnings, for $k \in \{0, \nu, z, \xi, rf\}$.

The tables and figures in Section 5.2 are similarly computed by choosing the right change in j for each affected mother's cohort, and integrating over the affected population.

E Tables and Figures not in text

HC Production	α	E slopes across S given E
	α_1	E slopes across S_P given E
	α_W	E slope controller
Spillovers	ν	E levels across S_P
	λ	E levels across S
	b	S level controller
Abilities	ρ_{zh_P}	Mincer coefficient β_2
	μ_z	E level controller
	σ_z	E level variation
Tastes	$\delta_S, \zeta_h, \zeta_c$	S levels
	γ_h, γ_z	S levels across S_P and E
	σ_ξ	S level variation

Table 11: Identification

(S_P, S, E) stand for mother's schooling, and the individuals' schooling and earnings levels, respectively. Taste heterogeneity picks up the residual unobserved heterogeneity not captured by the simple model.

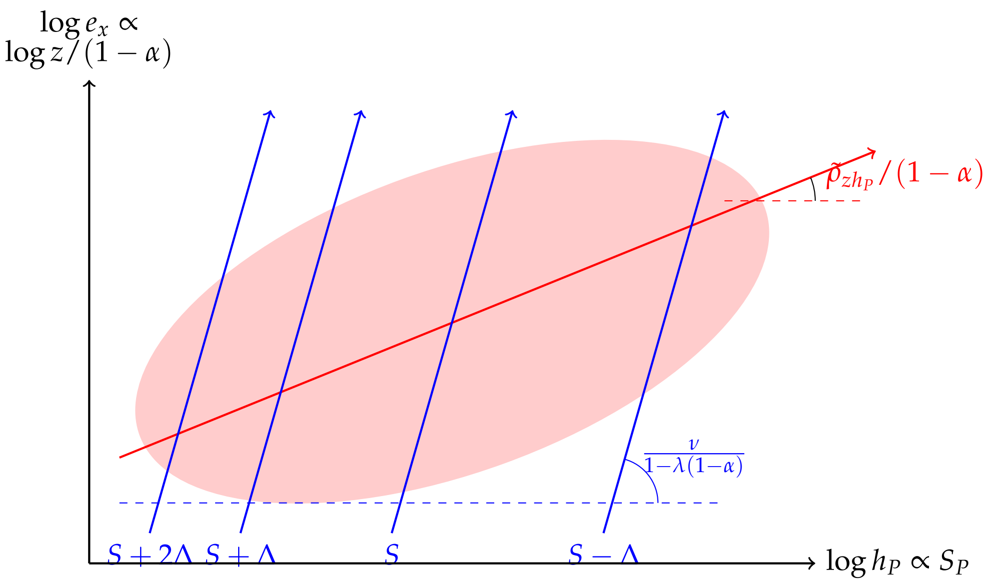


Figure 5: Intuition for Identifying Ability Selection and Spillovers.

x-axis: parents' human capital, or their schooling levels, y-axis: children's log abilities, or log earnings controlling for own schooling. The pink area represents the distribution of abilities and parents' schooling, (h_P, z). According to our model, selection is captured by the population correlation between parent's schooling and earnings, controlling for schooling. This is captured by the slope of the red 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 blue lines. Note that conditional on abilities, schooling is decreasing in parents' schooling.

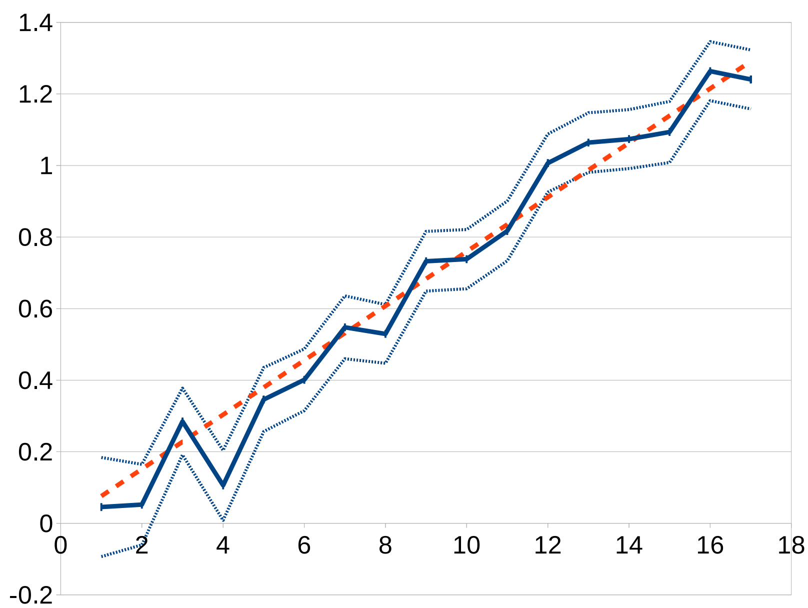
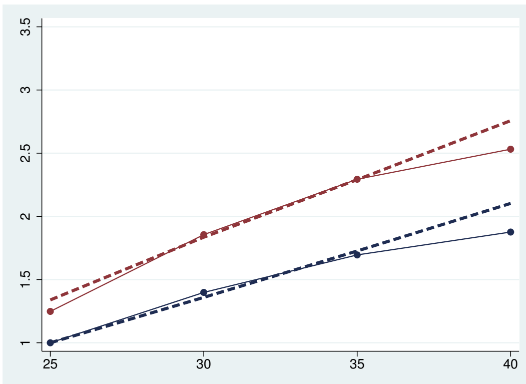
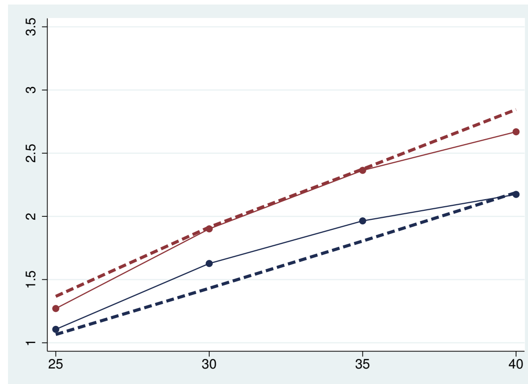


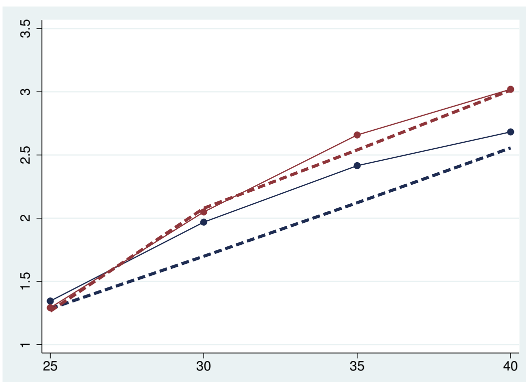
Figure 6: Mincerian Return to Schooling, Linear vs. Dummies.



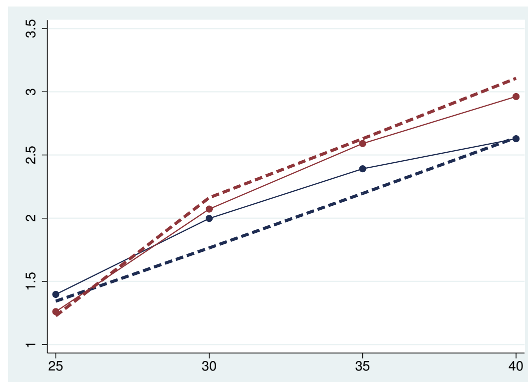
(a) $S_P < 5$



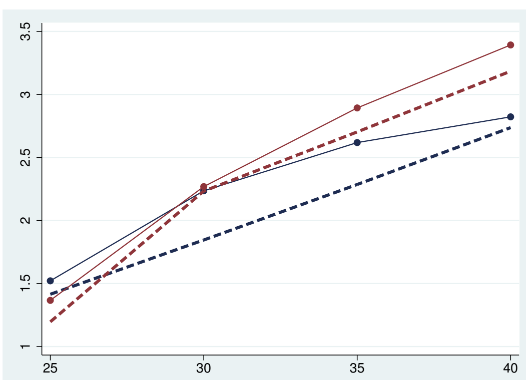
(b) $S_P \in \{6, 7\}$



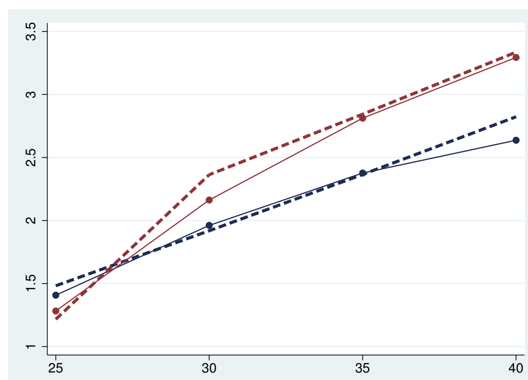
(c) $S_P = 8$



(d) $S_P \in [9, 12)$



(e) $S_P = 12$



(f) $S_P > 12$

Figure 7: Model Fit

y-axis: normalized average earnings, x-axis: ages 25,30,35,40. Solid and dashed lines are, respectively, the data and model moments implied by the GMM parameter estimate values. The red lines on top correspond to individual's with $S < 12$ for the first row of plots, and $S \leq 12$ for the rest. The blue lines on the bottom correspond to the converse.