Education, Self-Selection and Intergenerational Transmission of Abilities

Adalbert Mayer*

February, 2008

This paper shows that the relationships between earnings and college education across generations can be explained by the intergenerational transmission of two distinct abilities and subsequent self-selection of educational attainment. It is not necessary to invoke credit constraints. The model offers a way to reconcile the predictions of economic theory with empirical results that stress the importance of ability transmission across generations and question the importance of credit constraints for determining college attendance.

(JEL J31, I20)

*Department of Economics, Texas A&M University, College Station, TX 77843-4228 (e-mail: amayer@econmail.tamu.edu) This paper is based on chapter two of my dissertation at the University of Rochester. Therefore, I want to thank my advisors, Mark Bils and especially Lance Lochner, as well as Gordon Dahl, for their continual advice and encouragement. I benefited greatly from conversations with Marine Carrasco, Donald Deere, Shakeeb Khan, John Moroney, Claudia Sanchez-Vela, Anthony Smith, Chris Taber, and Manuelita Ureta. I received many useful comments from Nathan Grawe, and seminar participants at the Human Capital Workshop of the Federal Reserve Bank in Cleveland, the University of Augsburg, Texas A&M, Purdue, Rochester, Southern Methodist, Universite de Quebec a Montreal, Western Ontario and the 2004 SED meetings. I would like to thank the editors Francis Lui and Isaac Ehrlich for several important suggestions. All remaining mistakes are mine.

I. Introduction

The persistence of earnings and education across generations has been well documented. The current conventional wisdom is that the correlation of earnings of fathers and sons is above 0.4 (Solon, 1999). Mulligan (1997) reports estimates of the correlation between parents' and child's schooling between 0.19 and 0.45. It is also understood that intergenerational persistence of these two variables is interconnected.[1]

Grawe and Mulligan (2002) distinguish between two approaches to study intergenerational links. One approach is to statistically differentiate between various sources of persistence (e.g. genetics vs. environment). This route is taken by Bowles and Gintis (2001) or Black et al. (2005). The other (structural) approach - exemplified by Becker and Tomes (1979, 1986) - relies on economic theory and recognizes that intergenerational correlations are determined by decisions of agents. Becker and Tomes assume the transmission of a single ability from parents to children. Subsequently, parents invest in the human capital of their children. This investment decision can be affected by credit constraints.

In this paper, I present an alternative model that explains the observed relationships between college education and earnings across generations through the intergenerational transmission of two distinct abilities and subsequent self-selection of educational attainment. The cost of obtaining an education does not propagate intergenerational persistence. The model is motivated by three important lines of empirical research that show: (1) the existence of multiple skills and self-selection in the labor market; [2] (2) the importance of intergenerational transmission of ability;[3] and (3) the relatively minor role played by credit constraints at the college level.[4]

I combine the intergenerational transmission of abilities with a self-selection framework in the spirit of Roy (1951). Each individual is endowed with two distinct abilities. One ability (brain) is useful in occupations that are obtained after receiving a college education; the other ability (brawn) is

[1] See Piketty (2000), Bowles and Gintis (2001), and Mulligan (1997).
[2] See Carneiro, et al. (2005), Heckman and Scheinkman (1987), Heckman and Sedlacek (1990), Sattinger (1993), and Willis and Rosen (1979).
[3] See Black et al.(2005), and Plug and Vijverberg (2003).
[4] See Carneiro and Heckman (2002), Cameron and Heckman (1999), and Cunha et al.(2007).

utilized in occupations that do not require prior college training. The potential earnings in a college or non-college career are determined by these abilities. Each person chooses his career in order to maximize his earnings.[5] The two abilities are, in part, transmitted from parents to their children.[6]

I use data from the Panel Study of Income Dynamics (PSID) to estimate the model and compare the predictions of the parameterized model to the data. The transmission of abilities and subsequent self-selection of educational attainment are able to explain the observed relationships between earnings and education within and across generations. Notably, the model generates the positive relationship between the earnings of a father and the likelihood that his son goes to college conditional on the education level of the father. Individuals who are talented in both sectors choose the sector with the highest variation in earnings – here, the college sector. Abilities are persistent across generations. Therefore, a positive correlation between the two abilities implies that the probability that a son chooses to obtain a college education increases in both abilities of the father. Consequently, greater parental ability and earnings are associated with a higher probability that the son attends college, independent of the education level of the father. A one sector model (like Becker and Tomes, 1986) requires credit constraints to explain this relationship between father's earnings and son's educational attainment. This is important because credit constraints suggest policy remedies to increase efficiency while heterogeneity in abilities does not.

In the next section, I introduce the model and derive some of its theoretical implications. In section III, I present the data and estimate the model parameters. In section IV, I examine how well the parameterized model replicates features of the data. Section V compares the intergenerational self-selection model and the classic model by Becker and Tomes. Section VI concludes.

---

[5] Alternatively, a self-selection model with two abilities can be viewed as a model of heterogeneity in the ability to benefit from education (see Caneiro, Heckman, and Vytlacil, 2005). Individuals are endowed with a base ability and differ in their capacity to benefit from investment in human capital. The optimal amount of education is chosen based on these two characteristics. If both the base ability and the capacity to benefit from education are persistent across generations, the resulting properties are equivalent to those of the framework here.

[6] I do not distinguish different transmission channels, like genetics or cultural influences. I envision a broad concept of ability. It includes all traits rewarded in the labor market like: intelligence, sense of duty, determination, physical attributes, etc. I consider abilities at the time of the decision to attend college; these abilities may be the result of prior investment in human capital. I cannot provide insights about the role of credit constraints in the formation of these abilities. For more on skill formation over the life-cycle see Cunha et al. (2007).

II. The Intergenerational Self-Selection Model

In this section, I present an intergenerational self-selection model, based on Roy (1951).[7] Heckman and Honore (1990) provide a formal discussion of Roy's ideas. I study the decision whether to obtain a college education.[8]

Each individual chooses to work in one of two sectors $k \in \{H, C\}$. Sector H (for high school) represents all occupations that can be obtained without a college education, while sector C consists of jobs requiring a college education. The net lifetime earnings of individual i of generation g in sector k are determined by his level of ability[9] relevant for the sector, $\tilde{s}_{ki}{}^{g}$, the task price paid for this ability $\tilde{\pi}_{k}{}^{g}$, and the costs to enter the sector, K_{ki}^{g} :

$$\tilde{w}_{ki}^{g} = \tilde{\pi}_{k}^{g}\tilde{s}_{ki}^{g} - K_{ki}^{g}.$$

I model the costs of entering the college sector to be proportional to the potential earnings, reflecting the importance of opportunity costs. Therefore,

$$K_{ki}^{g} = k_{ki}^{g}\tilde{\pi}_{k}^{g}\tilde{s}_{ki}^{g}$$

and

$$\tilde{w}_{ki}^{g} = \frac{\tilde{\pi}_{k}^{g}\tilde{s}_{ki}^{g}}{\tilde{c}_{ki}^{g}}, \qquad \text{with } \tilde{c}_{ki}^{g} = \frac{1}{1-k_{ki}^{g}}.$$

If the high-school sector is chosen, no costs occur and $\tilde{c}_{Hi}^{g} = 1$. The focus of this paper is the determination of earnings and education due to differences in abilities. To clearly distinguish this channel from explanations of educational attainment due to cost differences, I assume that costs are the same for each member of a generation, $\tilde{c}_{ki}^{g} = \tilde{c}_{k}^{g}$. Transforming the above equation by taking logs gives:

$$\ln \tilde{w}_{ki}^{g} = \ln \tilde{\pi}_{k}^{g} + \ln \tilde{s}_{ki}^{g} - \ln \tilde{c}_{k}^{g} \qquad \text{or} \qquad w_{ki}^{g} = \pi_{k}^{g} + s_{ki}^{g} - c_{k}^{g},$$

---

[7] The model can be generalized to allow for choices among many different sectors. Like Roy (and most of the literature on the self-selection of educational attainment) I consider two sectors.

[8] Roy's framework has been used previously to study this choice, see Willis and Rosen (1979) Carneiro, Hansen, and Heckman (2003), or Heckman, Lochner, and Todd (2006). Like these papers, I model only the discrete choice of college attendance and ignore quality differences in educational attainment.

[9] The sector specific abilities may include intelligence, work ethic, or physical attributes; all of which may be the result of investment in human capital prior to the decision whether to attend college.

where $x = \ln \tilde{x}$ for all *x* variables.

An individual chooses the sector, *J,* that provides the highest earnings; the maximized earnings are denoted by v_i^g.

$$v_i^g = \max_{J \in \{H,C\}} \{w_{Hi}^g, w_{Ci}^g\}.$$

Each individual has one child. The transmission of abilities between generations is described by the following first order autoregressive process:

(1) $$s_{ki}^{g+1} = b_k s_{ki}^g + u_{ki}^{g+1}.$$

The coefficient b_k captures the persistence of abilities. To simplify the notation, I drop the subscript *i* for the individual. The term u_k^g is a random draw that represents the evolution of the ability useful in sector *k* from one generation to the next. I assume that u_k^g is i.i.d. across individuals and generations. The two sector-specific abilities may be correlated; this is modeled by having u_k^g drawn from a bivariate normal distribution with mean zero and covariance matrix Σ:

$$\left(u_H^g, u_C^g\right)' \sim N\left(0, \Sigma\right) \qquad \text{with } \Sigma = \begin{bmatrix} \sigma_H^2 & \sigma_{HC} \\ \sigma_{HC} & \sigma_C^2 \end{bmatrix}.$$

The correlation between u_H and u_C is given by: $\rho = \sigma_{HC} / \left(\sigma_H \sigma_C\right)$. The AR(1) structure for the evolution of skills leads to a stationary distribution of abilities[10] for each generation and the overall population observed at each given point in time:

(2) $$\left(s_H^g, s_C^g\right)' \sim N\left(0, \begin{bmatrix} \dfrac{\sigma_H^2}{1-b_H^2} & \dfrac{\sigma_{HC}}{1-b_H b_C} \\ \dfrac{\sigma_{HC}}{1-b_H b_C} & \dfrac{\sigma_C^2}{1-b_C^2} \end{bmatrix}\right) = N\left(0, \begin{bmatrix} \bar{\sigma}_H^2 & \bar{\sigma}_{HC} \\ \bar{\sigma}_{HC} & \bar{\sigma}_C^2 \end{bmatrix}\right)$$

The relationship between the skill distributions of fathers and sons is characterized by b_k and σ_k. $\bar{\sigma}_k^2 = \sigma_k^2 / \left(1 - b_k^2\right)$ is the variance of the skill s_k^g for a cross-section of the population.[11] The persistence

[10] The stationary skill distribution requires the assumption that any changes in π_k^g do not affect human capital investment that leads to the formation of skills.

parameter b_k characterizes the relationship between the population-wide variance of a skill and the variance of the son's skill given the skill level of his father. It is not possible to separately identify the task price paid for ability in a given sector and the mean of the distribution of the respective ability. Therefore, I normalize the mean ability within each sector to zero, $E\left(s_k^g\right)=0$. The task price, π_k^g, captures the mean of the respective ability and price paid for the ability.

For any given generation, the properties of the standard Roy Model hold (see Heckman and Honore, 1990). The high school sector is chosen if it promises higher lifetime earnings than the college sector, $s_H+\pi_H^g>s_C+\pi_C^g-c_C^g$. The distribution of skills in each generation is described by equation (2) and does not change over time. If the prices, π_k^g, and costs, c_k^g, are constant, there are also no changes in the distribution of earnings and sector choices. The fraction of workers in each sector remains fixed. If the relative price for one skill increases, the fraction of people working in the respective sector increases as well.

The model predicts the relationships between ability distributions, sector choices and earnings distributions across generations. I focus on two generations, fathers *(g=f)* and sons *(g=s)*. Note that the task price level, π_k^g, may be different for fathers and sons.

The persistence of abilities described by equation (1) leads to persistence in earnings across generations; this is captured in Proposition 1.

Proposition 1:

The son's earnings increase in the father's earnings: $\frac{dE\left(v^s\right)}{dv^f}>0$.

Proof: See Appendix 2

The intergenerational transition probabilities between sectors give the odds that a son chooses a certain sector given the sector of his father. They can be obtained from the sons' ability distributions

[11] Also, ρ stands for the correlation between u_H and u_C, while $\bar{\rho}$ represents the correlation between s_k^g and s_k^g.

conditional on their fathers' sector choice. Proposition 2 states that the model implies persistence in sector choice, given the same level of persistence in both skills, i.e. $b_H=b_C=b$.[12]

Proposition 2: *The effect of father's sector choice on the son's sector choice.*

If $b_H=b_C=b>0$ then:

a) *The probability of choosing a certain sector is higher, if the father already chose this sector.*

$$\Pr\left(J^s = l \middle| J^f = l\right) > \Pr\left(J^s = l \middle| J^f = k\right) \qquad \text{for } l \neq k$$

b) *Conditional on the earnings of the father, the probability of choosing a certain sector is higher, if the father already chose this sector.*

$$\Pr\left(J^s = l \middle| J^f = l, v^f = v^*\right) > \Pr\left(J^s = l \middle| J^f = k, v^f = v^*\right) \qquad \text{for } l \neq k$$

c) *The persistence in sector choices increases in the ability-persistence*

$$\frac{d\left[\Pr\left(J^s = l \middle| J^f = l\right) - \Pr\left(J^s = l \middle| J^f = k\right)\right]}{db} > 0$$

Proof: See Appendix 2

Intuitively, the difference between the abilities for sector *H* and sector *C* is partially transmitted to the next generation. This transmission process is described by $s_H^s - s_C^s = b\left(s_H^f - s_C^f\right) + \left(u_H^s - u_C^s\right)$. Fathers in sector *H* have relatively high values of $\left(s_H^f - s_C^f\right)$ and thus their sons tend to have higher values of $(s_H^s - s_C^s)$ than the sons of fathers in sector *C*. This still holds true when conditioning on the earnings of the father. If the level of persistence of abilities is higher, more of the difference in abilities is transmitted.

The allocation of more or less talented workers into the sectors depends on the variance and covariance of the abilities s_H^g and s_C^g.[13] Given the assumption of $\overline{\sigma}_C^2 > \overline{\sigma}_H^2$ (which is empirically

[12] Simulations of the model suggest that, for most parameter combinations, propositions 2 and 3 are also true when b_H and b_C are substantially different.

confirmed below), the fathers in sector C have above average talent for sector C. Consequently, intergenerational persistence of abilities implies that the sons of college educated fathers have above average talent for the college sector:

$$E\left[s_C^s \middle| J^f = C\right] > 0 > E\left[s_C^s \middle| J^f = H\right].$$

The talent for sector H of fathers in sector H can be - depending on the correlation between the abilities in the two sectors - above or below average. Again, this feature of the fathers is transmitted to the sons.

$$E\left[s_H^s \middle| J^f = H\right] > 0 > E\left[s_H^s \middle| J^f = C\right] \qquad \text{if } \bar{\rho} < \frac{\bar{\sigma}_H}{\bar{\sigma}_C}$$

$$E\left[s_H^s \middle| J^f = H\right] < 0 < E\left[s_H^s \middle| J^f = C\right] \qquad \text{if } \bar{\rho} > \frac{\bar{\sigma}_H}{\bar{\sigma}_C}$$

For a sufficiently high correlation between the two skills, the sons of college educated fathers have above average talent for the high-school sector, as well.

Proposition 3 relates the earnings of the father to the sector choice of the son. If the variation in the college skill is higher than the variation in the high-school skill, there is a positive relationship between the earnings of college educated fathers and the likelihood that their sons attend college.

Proposition 3: *The effect of the father's earnings on the son's sector choice*

If $\sigma_C > \sigma_H$ *and* $b_H = b_C = b$ *then:*

a) $\Pr\left(J^s = C \middle| J^f = C\right)$ increases in v^f

b) If $\bar{\rho} > \frac{\bar{\sigma}_H}{\bar{\sigma}_C}$ then $\Pr\left(J^S = C \middle| J^f = H\right)$ increases in v^f

If $\bar{\rho} < \frac{\bar{\sigma}_H}{\bar{\sigma}_C}$ then $\Pr\left(J^S = C \middle| J^f = H\right)$ decreases in v^f

c) $\Pr\left(J^S = C \middle| J^f = C\right)$ increases more in v^f than $\Pr\left(J^S = C \middle| J^f = H\right)$

Proof: See Appendix 2

[13] See Heckman and Honore (1990) or Sattinger (1993).

For a high correlation, ρ, the most talented people of both ability types tend to end up in sector *C*. The higher the earnings of a father in sector *H*, the more likely he also has a high ability in sector *C*. This makes it more likely for sons of the high earnings fathers in sector *H* to switch to job *C*. While higher earnings of the father in sector *H* relate to ability *C* of the son indirectly, earnings of the father in sector *C* translate directly into a higher s_C^s of the son. Therefore, the likelihood that a son attends college increases more in the earnings of a college educated father, than in the earnings of a high-school educated father.

III. Data and Estimation

I use data from the Panel Study of Income Dynamics (PSID) conducted by the Survey Research Center, Institute for Social Research at the University of Michigan. The PSID is a longitudinal study of a representative sample of the US population. It began in 1968 and includes information about various economic and demographic variables. All members of the original families and their offspring are interviewed annually.

I analyze the core sample, selected to be representative of the US population, and consider only males. I match fathers and sons to obtain father-son pairs. These pairs consist of fathers who were between 22 and 65 years old in 1970 and sons who were between 1 and 19 years old in 1970. I categorize all fathers and sons according to years of schooling at age 25. Those with more than 12 years of schooling are considered part of the college sector; all others are in the non-college sector.

Using annual wage information about the earnings from 1968 to 1976 for the fathers and from 1992 to 2001 for the sons, I construct an age and year adjusted value representing lifetime wages of each individual *i*. I only consider wages of heads of household who worked at least 100 hours in the respective year. For each year, I regress the natural logarithm of wages, y_{ti}, on a constant, race and region controls, age and age squared:

$$\ln y_{ti} = a_{1t} + \beta X_t + a_{2t} age_{ti} + a_{3t} age_{ti}^{\ 2} + e_{ti} .$$

I then use $v_i = \frac{1}{N}\sum_{t=1}^{N} e_{ti}$ (for all N available years) as a measure of normalized lifetime-wages.[14]

Table 1 describes the main features of the sample. Allowing for multiple sons per father, I obtain a sample of 853 sons and 469 fathers (some fathers appear more than once in the sample).[15] The return to each year of education (as measured by OLS on the years of schooling) has increased between the generations. The fraction of workers who attended college increased from below 40 percent to more than one half. The intergenerational correlation in wages is .3. The intergenerational correlation in years of education is .45.

Identification and Estimation

The parameters of the model are reflected in a number of data features, providing more information than the standard static Roy Model.[16] I observe the sector choices of the father-son pairs; the mean and variance of the wages of both generations conditional on the sector choices of the father-son pairs; and intergenerational persistence of wages conditional on the sector choices of the father-son pairs. I use all the resulting 21 moment conditions (see Table A1) to estimate the 9 parameters of the model.

Even though these data features depend simultaneously on all the parameters of the model, there are some obvious relationships that illustrate how each parameter is identified. Identification of σ_H, σ_C, b_H, and b_C is provided by the variation of wages conditional on sector choice and the intergenerational persistence in wages. The relative task prices, $\pi_C^f - \pi_H^f$ and $\pi_C^s - \pi_H^s$, and the cost of attending college, c^f and c^s, are reflected by the fraction of sons and fathers in each sector and by the wage differences between the two sectors. The relationship between the fathers' wages and educational attainment of the sons is directly related to ρ.

14 Averaging reduces effects of measurement error which biased early estimates of intergenerational persistence. I do not control for other characteristics (like occupation or marital status) that that may drive the intergenerational relationships I try to explain. I also used annual labor income instead of hourly wages; discounted wages over time; changed the years considered and other selection criteria. The resulting data patterns are very similar.

15 Restricting the sample to only one son per family does not change the results.

16 Heckman and Honore(1990) show that the identification of the Roy model in a cross section is theoretically possible, given appropriate functional form assumptions. However, it is very difficult to obtain robust results from a practical standpoint. The intergenerational relationships considered here offer additional means of identification.

The model is characterized by 9 parameters: ρ, σ_H, σ_C, b_H, b_C, $\pi_C^f - \pi_H^f$, $\pi_C^s - \pi_H^s$, c^f, and c^s.[17] I use a Simulated Generalized Method of Moments approach to estimate these parameters. I choose the parameter values that minimize the weighted sum of squared differences between the data moments and the moments resulting from simulations of the model. The values of the simulated and actual moments are displayed in Table A1.

The estimation results are presented in Table 2.[18] The last column gives the value of the objective function. A test of the 12 over-identifying restrictions rejects the specification of the model (the p-value is below .01). An exact fit of the model would imply that the functional form assumptions are exactly met and that the model captures all the mechanisms connecting earnings and education across generations. This is not the case. However, in section IV, I show that the model is able to generate most of the relationships between education and earnings across generations. The estimated correlation between the skills, ρ, is 0.52. The estimates for the conditional standard deviations of the high-school skills (σ_H) and college skills (σ_C) are 0.38 and 0.56. Conditional on the skill of the father, the variance of the college skill is higher than the variance of the non-college skill. The level of persistence for the college skill is higher than for the non-college skill, $b_C = .52$ and $b_H = .11$. Together the parameters b_k and σ_k characterize the variance of skill *k* in a cross-section, $\overline{\sigma}_k^2 = \sigma_k^2 / \left(1 - b_k^2\right)$. The abilities used in occupations requiring a college education vary more across the population than the abilities required in other jobs.

I estimate that $\pi_C^g - \pi_H^g$ increased between generations, while c^g decreased. We observe a higher rate of college enrollment for the generation of the sons than for the generation of the fathers. Theoretically, this can arise from two factors: an increase in relative task price of the college skill or a decrease of the cost of attending college. The estimate of the change in $\pi_C^g - \pi_H^g$ that generates the

[17] Note, I constrain the estimates for the variances to be non-negative. The correlation is constrained to be between -1 and +1. The other parameters are unrestricted.

[18] Standard errors are obtained in a similar fashion to GMM, the efficiency of the estimator is, however, slightly reduced by replacing closed-from expressions of the moments by simulated values. I simulate 10,000 father-son pairs which results in an efficiency of 96% relative to a classic GMM estimation (Gourieroux and Monfort,1996).

observed growth in the college wage premium explains only part of the increase in college attendance. To fit the data, the cost of college attendance has to decrease simultaneously. One explanation for this finding is that governmental subsides have reduced the cost of attending college for the generation of the sons. It is also possible that the sons work more while in school, which decreases forgone earnings.

The estimates of the absolute levels of $\pi_C^g - \pi_H^g$ and c^g depend on the rate of college enrollment and the college wage premium. They also depend on the remaining parameters of the model. All parameters of the model are estimated simultaneously. I estimate $\pi_C^g - \pi_H^g$ to increase from -.1026 for the generation of the fathers to .048 for the generation of the sons.[19] The estimates of the cost of attending college are 12% of lifetime wages for the generation of the fathers and a negligible amount for the generation of the sons.[20] This last estimate is lower than expected. There are obvious monetary costs of attending college, such as forgone earnings or tuition. Nevertheless, a low value for c^s provides the best overall fit of the model. One possible explanation for the low estimate of c^s is that the cost of attending college reflects other factors besides monetary costs. In my model the cost captures factors that affect the decision to attend college but are not reflected in subsequent wages. A taste for college would decrease the cost of attending college. Moreover, given the standard errors reported in Table 2, it is not possible to reject more substantial costs, c^s.

As a robustness check, I use outside information to determine the values of c^f and c^s. I use forgone earnings while in college to approximate the cost of college. I calibrate $c^f = .2$ and $c^s = .2$.[21] I also consider a specification with $c^f = .2$ and $c^s = .1$. I impose these values as restrictions on the model and estimate the remaining 7 parameters. The results are displayed in Table A2. The specification with a lower cost of college for the sons fits the data better. Using the parameters from the restricted

---

[19] The negative estimate for $\pi_C^f - \pi_H^f$ indicates that the majority of workers in the generation of the fathers has higher lifetime earnings in the high-school sector. These workers choose to forgo a college education. The model still implies higher wages for college educated workers than for workers without college education (see Table 5).

[20] I do not impose any restrictions on the estimates of c^f and c^s.

[21] Let T denote the number of years from high school graduation to retirement, S the years spent in college, and r the discount rate. The increase in earnings necessary to compensate a person for forgone earnings due to college attendance is given by: $c = \ln\left(1 - e^{-rT}\right) - \ln\left(e^{-rS} - e^{-rT}\right)$. For T=45, S=4 and r=0.04, this is about 0.2.

estimations instead of the parameters from the unrestricted estimation does not change the implications of the model discussed in Section IV. For both robustness specifications, the predicted relationships between wages and education across generations are qualitatively identical to the predictions based on the unrestricted model. For the specification with a lower cost of college for the sons, the predictions are also quantitatively very similar to those implied by the unrestricted estimates. When using the parameters from the robustness specifications instead of the parameters from the unrestricted estimation, the model can still generate the data features discussed in Section IV without appealing to credit constraints.

IV. Quantitative assessment

In this section, I compare the predictions of the model to various features of the data. I show that the parameterized intergenerational self-selection model can explain the observed connections of education and earnings within and across generations both qualitatively and quantitatively.[22] I obtain the predictions of the model by simulating it with the parameter estimates from Table 2: ρ=.5154, $\sigma_H = .3776$, $\sigma_C = .5624$, $b_H = .1118$, $b_C = .5249$, $\pi_C^f - \pi_H^f = -.1026$, $\pi_C^s - \pi_H^s = .048$, $c^f = .1161$, and $c^s = .0003$. In Tables 3 through 7 the values based on PSID data are seen on the left and the predictions of the model are on the right.

Table 3 shows the results of regressions of the lifetime-wages of sons on their own education level and lifetime-wages and education level of their fathers.[23] The model replicates the magnitudes for all combinations of independent variables. Column 1 displays the OLS return to college. Column 2 shows the intergenerational wage elasticity. As suggested by Proposition 1 the model generates the observed intergenerational persistence. The positive effect of father's college attendance on the wage of

[22] I estimate the 9 parameters of the model by fitting it to 21 parameters. An alternative approach would be to estimate the model by fitting fewer moments and test the ability of the model to fit the additional moments. The goal of this section is to show that the model generates the observed data patters. Therefore, I assess the fit of the model based on characteristics determined by these 21 moments. In addition, the estimation results are sensitive to the choice of the 9 moments.
[23] Note, lifetime-wages are adjusted of race and region. They are not adjusted for features that are related to the choice modeled, like occupation or industry.

the son is reported in column 3. Column 4 displays the effect of the father's wage on the son's wage after controlling for son's education. As predicted by the model, the effect of father's wage is reduced. Similarly, controlling for education of the father reduces the effect of father's wage (column 5). In column 6, I include controls for college attendance for both, the father and the son. All three coefficients are positive and the model replicates their relative magnitudes.

The Roy Model offers not only predictions about the conditional means but also about the distributions of wages conditional on sector choice. Table 4 shows the mean and standard deviation of the sons' lifetime-wages, given the sector choices of the father and the son. The model replicates the relative sizes of the standard deviations of the sons' wages for the different sector choice combinations. It captures that the sons' earnings are highest when both members of the father-son pair attended college and lowest when both did not attend college. The predictions of the model and the data differ for the sons of college educated fathers without college education. The model predicts lower wages and a lower variation in wages for these workers than observed in the data. Table 5 shows the mean and standard deviation of the lifetime-wages of the father by career choice of the father-son pairs. The model replicates the relative magnitudes of the means and standard deviations. College educated fathers whose sons attended college as well have the highest wages. The variation in wages is also the highest for this group. The lowest wages (and variation in wages) are observed for high-school fathers of sons without college education.

As seen in Table 6, the model replicates the sector choice decisions of father-son pairs. As a result of the changes of the cost of attending college and the relative task prices, the fraction of college educated workers has increased between the generations. The sons of college educated fathers are more likely to attend college than the sons of fathers without college education, a feature implied by Proposition 2 a). This is also seen in Table 7, which shows the results of probit regressions of college attendance of the son on earnings of the father, college attendance of the father and their interaction. Proposition 2 b) states that the positive affect of father's education on son's college attendance persists after controlling for father's wages. Column 3 of Table 7 confirms this prediction.

In column 1 of Table 7, I report the positive relationship between the earnings of the father and the probability that the son goes to college. As seen in column 3, this relationship is weaker but still clearly positive, when conditioning on the education level of the father. As suggested by Proposition 3 the model replicates these features. Proposition 3 c) suggests that the effect of fathers earnings on educational attainment of the son is stronger for college educated fathers than for fathers without college education. To tests this prediction, I include the interaction between father's earnings and father's education in column 4 of Table 7. While the point estimate for the interaction term is positive, it is not significant and its magnitude is smaller than predicted by the model.

Using only nine free parameters, the model is able to replicate many more features of the data. However, there are two notable discrepancies between the data and the predictions of the model. First, the wages of sons without college education whose fathers went to college are predicted lower than observed in the data. Second, the role of father's education for the importance of the effect of father's earnings on son's education is predicted higher than observed. This suggests that the intergenerational transmission of abilities and subsequent self-selection of educational attainment is a mechanism that is important for – but does not completely explain – the observed relationships between education and earnings across generations.

V. Comparison to Becker and Tomes (1986)

A model like Becker and Tomes (1986) is also able to explain many of the observed intergenerational connections of education and earnings. (A simplified version of the Becker-Tomes model is reviewed in Appendix 1) However, my model and the Becker-Tomes model have important differences.

The Becker-Tomes model without credit constraints predicts the intergenerational persistence in earnings and educational attainment. It can also explain the positive correlations between father's education and son's income, or father's income and education of the son (See Tables 3 and 7).

The positive relationship between parental income and educational attainment of the son after conditioning on parental education (column 3 Table 7) can only be explained when allowing for two intergenerational channels. Becker and Tomes do so by allowing for credit constraints. The model presented here offers an alternative explanation of this data pattern that does not imply the presence of credit constraints.

There are two data features that cannot be explained by a model like Becker-Tomes. First, it cannot generate the positive effect of father's education on wages of the son after controlling for wage of the father and education of the son. This effect can be seen in column 6 of Table 3. Second, the Becker-Tomes model does not predict that the influence of father's earnings on son's education is stronger for college fathers than for non-college fathers. This effect is reported in column 4 of Table 7. However, the effect is weaker than predicted by the self-selection model and not statistically significant.

Although some of the predictions of the self-selection model presented here and the Becker-Tomes model are similar, there are important differences in the underlying mechanisms. Persistence in ability is a natural starting point for both models. However, a single intergenerational transmission channel cannot generate the observed data patterns. Therefore, both models offer two channels to link generations. Becker and Tomes consider the intergenerational transmission of one ability and propose credit constraints as a second link between generations – parental income directly affects educational attainment of their offspring.

The mechanism proposed here is the transmission of two distinct abilities and subsequent self-selection of educational attainment. The related literature offers some support for this approach. Evidence for the existence of multiple skills and self-selection in the labor market is provided by – among others – Willis and Rosen (1979), Heckman and Sedlacek (1990), and Heckman and Scheinkman(1987) (see Sattinger (1993) for further references). Carneiro, Heckman, and Vytlacil (2005), who also compare individuals with and without college education, find that "comparative advantage is a central feature of modern labor markets". The importance of intergenerational ability transmission has been confirmed more recently. Black et al. (2005) find that "the high correlations between parents' and children's education are due to family characteristics and inherited ability and not education spillovers." Plug and Vijverberg (2003) report that IQ is an important determinant of educational attainment. They compare biological and adopted children to disentangle the effects of nature and nurture, and argue that 55 to 60 percent of parental ability is genetically transmitted. Moreover, there is evidence that credit constraints are of minor importance for the decision whether to attend college. Carneiro and Heckman (2002) suggest that only a small fraction of recent high-school graduates are affect by credit constraints. Cameron and Heckman (1999) and Cunha et al. (2007) show that much of the correlation between parental earnings and the educational attainment of children can be explained by observed ability differences.

The self-selection model presented here has different policy implications than the Becker-Tomes model. If credit constraints are important, intergenerational persistence is in part driven by the cost of educational attainment and can be affected by policies like student loans or tuition subsidies.[24] Moreover, credit constraints imply a distortion in the human capital investment decision. The self-selection model suggests an alternative: the intergenerational transmission of two distinct abilities related to different education levels. Differences in educational attainment are the result of ability differences and policies that encourage educational attainment would not improve efficiency.

VI. Conclusion

I propose a self-selection model with the intergenerational transmission of two distinct abilities. It illustrates how the transmission of abilities and subsequent self-selection of educational attainment can generate the observed connections between college education and earnings across generations – without invoking credit constraints. Therefore, this model offers an alternative to the intergenerational transmission mechanism proposed in Becker and Tomes (1986). If transmission of abilities is indeed the crucial link between generations, subsidizing college attendance is not an effective means to increase upward mobility or to improve the lot of low-income workers in general. Rather, the focus should lie on improving the abilities that are already formed by the time a person decides whether or not

[24] Note, the data analyzed is generated in a world were such policies are present. The question addressed here is: Are there credit constraints that are not overcome by current policies?

to attend college. Examples for such policies are early childhood intervention programs or investments in the (pre-) school education of children from low income families. Cunha et al. (2007) report that such programs can be helpful, especially when implemented early in the lifecycle of children.

Moreover, the model offers a novel way of using data on two generations to ascertain the extent of self-selection in the labor market. Theory and intergenerational data serve as an alternative to instruments or exclusion restrictions. This insight can be useful for other models of schooling with multiple skills.

Appendix 1

Intergenerational ability transmission with one sector

Here, I sketch implications of a model[25] with intergenerational transmission of a single ability and subsequent investment in human capital. Credit constrains may influence investment in human capital. Each individual of generation, *g,* is endowed with one ability, E_g, which is transmitted from parent to child,

$$E_{g+1} = bE_g + v_{g+1} \ .$$

The ability endowment, E_g, and investment in human capital by the parents, x_{g-1}, determine the human capital, H_g, of an individual,

$$H_g = \Psi\left(x_{g-1}, E_g\right) \ .$$

Becker and Tomes assume that endowment and investment in human capital are complimentary,

$$\frac{\partial H_g}{\partial x_{g-1}} > 0 \, , \ \frac{\partial H_g}{\partial E_g} > 0 \, , \ \frac{\partial^2 H_g}{\partial x_{g-1} \partial E_g} > 0 \cdot$$

The earnings of an individual, Y_g, are determined by his human capital level and market luck, l_g,

$$Y_g = H_g + l_g \, .$$

Without credit constraints, x_{g-1} is chosen so that the marginal return of the investment in human capital is equal to the market interest rate. If credit constraints are important, high income parents invest more than low income parents. It is straight forward to see that this model generates persistence in earnings and educational attainment (investment in human capital).

The model requires credit constraints to generate that both education and income of a child increase in parental income after conditioning on parental education. To see this, consider two dynasties *A* and *B*. In the absence of credit constraints $x_g^A = x_g^B$ implies $E_g^A = E_g^B$ and $H_g^A = H_g^B$.

[25] Becker and Tomes (1986) provide a more thorough discussion of such a model.

Therefore, $Y_g^A > Y_g^B$ is the result of $l_g^A > l_g^B$ which is not transmitted across generations and we would expect $E\left(x_{g+1}^A\right) = E\left(x_{g+1}^B\right)$ and $E\left(Y_{g+1}^A\right) = E\left(Y_{g+1}^B\right)$.

The one sector model cannot generate a positive effect of parental education on earnings of the child after conditioning for parental income and education of the child. Again, consider the dynasties A and B. If we observe $Y_g^A = Y_g^B$, $x_g^A > x_g^B$, and $x_{g+1}^A = x_{g+1}^B$, credit constraints are not an issue, as both dynasties have the same income. The human capital investment, x_{g+1}, is determined by the endowment of the child, E_{g+1}. Consequently, $E_{g+1}^A = E_{g+1}^B$ and $H_{g+1}^A = H_{g+1}^B$. This implies $E\left(Y_{g+1}^A - Y_{g+1}^B \middle| Y_g^A = Y_g^B, x_g^A > x_g^B, x_{g+1}^A = x_{g+1}^B\right) = 0$.

Appendix 2

Proofs of Propositions

Proof of Proposition 1

The earnings of the son are: $v^s = \max\left\{b_H s_H^f + u_H^s, b_C s_C^f + u_C^s\right\}$. Given the normality assumption, the expected value of the unobserved skill of the father is increasing in the observed skill,

$$\frac{dE\left(s_H^f \middle| s_H^f > s_C^f\right)}{ds_C^f} > 0 .$$

■

To simplify notation, let $u \equiv u_H - u_C$, $X^g \equiv s_H^g - s_C^g$, and $\bar{\pi}^g \equiv \pi_H^g - \pi_C^g$, with $g = s, f$.

Hence: $X^s = bX^f + u$. From (2) we know that: $X^g \sim N\left(0, \bar{\sigma}_H^2 + \bar{\sigma}_C^2 - 2\bar{\sigma}_{HC}\right)$ and $u \sim N\left(0, \sigma_H^2 + \sigma_C^2 - 2\sigma_{HC}\right)$. Let $f_X(.)$ and $f_u(.)$ denote the density functions of the distributions of X^g and *u*.

Proof of Proposition 2

The distribution of X^s is given that the father is in sector *H* is given by:

$$g_X^{s|H}\left(X^s\right) = \int_{\bar{\pi}^f}^{\infty} \frac{f_X\left(X^f\right)}{\Pr\left(J^f = H\right)} f_u\left(X^s - bX^f\right) dX^f ,$$

if the father chooses sector *C* it is:

$$g_X^{s|C}\left(X^s\right) = \int_{-\infty}^{\bar{\pi}^f} \frac{f_X\left(X^f\right)}{\Pr\left(J^f = C\right)} f_u\left(X^s - bX^f\right) dX^f .$$

The cumulative distribution function of X^s, with the father in sector *H* is given by:

(A 1) $$G_X^{s|H}\left(X^s\right) = \int_{\bar{\pi}^f}^{\infty} \frac{f_X\left(X^f\right)}{\Pr\left(J^f = H\right)} F_u\left(X^s - bX^f\right) dX^f ,$$

with a college educated father it is:

(A 2) $$G_X^{s|C}\left(X^s\right) = \int_{-\infty}^{\bar{\pi}^f} \frac{f_X\left(X^f\right)}{\Pr\left(J^f = C\right)} F_u\left(X^s - bX^f\right) dX^f .$$

The probabilities of following the father to sector H and switching to sector H are given by:

$$\Pr\left(J^s = H \middle| J^f = H\right) = \Pr\left(X^s > \overline{\pi}^s \middle| J^f = H\right) = 1 - G_X^{s|H}\left(\overline{\pi}^s\right)$$

and

$$\Pr\left(J^s = H \middle| J^f = C\right) = \Pr\left(X^s > \overline{\pi}^s \middle| J^f = C\right) = 1 - G_X^{s|C}\left(\overline{\pi}^s\right).$$

a) We need to show that $G_X^{s|C}\left(\overline{\pi}^s\right) - G_X^{s|H}\left(\overline{\pi}^s\right) > 0$. F_u is increasing. Therefore $F_u\left(X^s - bX^f\right)$ decreases in X^f. Consequently, (A 1) and (A 2) imply that $G_X^{s|C}\left(\overline{\pi}^s\right) - G_X^{s|H}\left(\overline{\pi}^s\right) > 0$.

■

b) We need to condition on the sector choice and the income of the father.

$$\Pr\left(J^s = H \middle| J^f = H, v^f = v^*\right) - \Pr\left(J^s = H \middle| J^f = C, v^f = v^*\right) > 0$$

or

$$G_X^{s|C,v*}\left(\overline{\pi}^s\right) - G_X^{s|H,v*}\left(\overline{\pi}^s\right) > 0.$$

The cumulative distribution function of X^s, with the father in sector *H* is given by:

(A 1)
$$G_X^{s|H}\left(X^s\right) = \int_{\overline{\pi}^f}^{\infty} \frac{f_X\left(X^f \middle| s_H^f = v* - \pi_H^f\right)}{\Pr\left(J^f = H, v*\right)} F_u\left(X^s - bX^f\right) dX^f,$$

with a college educated father it is:

(A 2)
$$G_X^{s|C}\left(X^s\right) = \int_{\overline{\pi}^f}^{\infty} \frac{f_X\left(X^f \middle| s_C^f = v* - \pi_C^f\right)}{\Pr\left(J^f = C, v*\right)} F_u\left(X^s - bX^f\right) dX^f.$$

The same argument as in part a) applies.

■

c) As b increases $F_u\left(X^s - bX^f\right)$ decreases more in X^f and the difference between the values $F_u\left(X^s - bX^f\right)$ conditional on the sector choice of the father increases. The effect described to in part a) is magnified.

■

Proof of Proposition 3

The distribution of X^f conditional on the skill level in sector C is given by:

$$F\left(X^f \middle| s_C^f\right) \sim N\left(\left[\frac{\sigma_{HC}}{\sigma_C^2} - 1\right] s_C^f ; \sigma_H^2\left(1-\rho^2\right)\right).$$

The distribution conditional on the skill level and the sector choice is given by:

$$G\left(X^f \middle| J^f = C; s_C^f\right) = G\left(X^f \middle| s_C^f ; X^f < \bar{\pi}^f\right) = \frac{F\left(X^f \middle| s_C^f\right)}{F\left(\bar{\pi}^f \middle| s_C^f\right)}.$$

The probability of choosing sector C given that the father chooses sector C and had skill level s_C^f is given by this cumulative distribution function $G(.)$ evaluated at the difference in task prices:

$$\Pr\left(J^s = C \middle| J^f = C; s_C^f\right) = \int_{-\infty}^{\bar{\pi}^s} g\left(X^s \middle| J^f = C; s_C^f\right) dX^s = G\left(\bar{\pi}^s \middle| J^f = C; s_C^f\right).$$

a) We know that if $G\left(X^f \middle| A\right) > G\left(X^f \middle| B\right)$ then $G\left(X^s \middle| A\right) > G\left(X^s \middle| B\right)$, hence it is sufficient to show that $\frac{dG\left(X^f \middle| J^f = C; s_C^f\right)}{ds_C^f} > 0$. The derivate of $F\left(X^f \middle| s_C^f\right)$ with respect to s_C^f is given by:

$$F'\left(X^f \middle| s_C^f\right) = \frac{dF\left(X^f \middle| s_C^f\right)}{ds_C^f} = \left(1 - \frac{\sigma_{HC}}{\sigma_C^2}\right) f\left(X^f \middle| s_C^f\right) > 0.$$

The derivative of the distribution function conditional on the choice of sector C is given by:

$$\frac{dG\left(X^f \middle| s_C^f ; X^f < \bar{\pi}^f\right)}{ds_C^f} = \frac{F'\left(X^f \middle| s_C^f\right) F\left(\bar{\pi}^f \middle| s_C^f\right) - F\left(X^f \middle| s_C^f\right) F'\left(\bar{\pi}^f \middle| s_C^f\right)}{\left[F\left(\bar{\pi}^f \middle| s_C^f\right)\right]^2}.$$

The numerator is positive if: $\left(1 - \frac{\sigma_{HC}}{\sigma_C^2}\right) f\left(X^f \middle| s_C^f\right) F\left(\bar{\pi}^f \middle| s_C^f\right) > \left(1 - \frac{\sigma_{HC}}{\sigma_C^2}\right) F\left(X^f \middle| s_C^f\right) f\left(\bar{\pi}^f \middle| s_C^f\right)$

or

$$\frac{f\left(X^f \middle| s_C^f\right)}{F\left(X^f \middle| s_C^f\right)} > \frac{f\left(\bar{\pi}^f \middle| s_C^f\right)}{F\left(\bar{\pi}^f \middle| s_C^f\right)}.$$

Since $F(.)$ describes a normal distribution and $X^f < \bar{\pi}^f$ (Sector C is chosen) this is true. ■

b) $\Pr\left(J^s = C \middle| J^f = H; s_H^f\right) = 1 - G\left(-\pi^s \middle| J^f = H; s_H^f\right)$

Need to show that:

$$\frac{dG\left(-X^f \middle| J^f = H; s_H^f\right)}{ds_H^f} \quad \begin{cases} > 0 & \text{if } \sigma_{HC} < \sigma_H^2 \\ < 0 & \text{if } \sigma_{HC} > \sigma_H^2 \end{cases}$$

The derivate of $F\left(X^f \middle| s_H^f\right)$ with respect to s_H^f is given by:

$$F'\left(-X^f \middle| s_H^f\right) = \frac{dF\left(-X^f \middle| s_H^f\right)}{ds_H^f} = \left(1 - \frac{\sigma_{HC}}{\sigma_H^2}\right) f\left(-X^f \middle| s_H^f\right) \quad \begin{cases} > 0 & \text{if } \sigma_{HC} < \sigma_H^2 \\ < 0 & \text{if } \sigma_{HC} > \sigma_H^2 \end{cases}$$

The derivative of the conditional distribution function is:

$$\frac{dG\left(-X^f \middle| s_H^f; -X^f < -\bar{\pi}^f\right)}{ds_H^f} = \frac{F'\left(-X^f \middle| s_H^f\right) F\left(-\bar{\pi}^f \middle| s_H^f\right) - F\left(-X^f \middle| s_H^f\right) F'\left(-\bar{\pi} \middle| s_H^f\right)}{\left[F\left(-\bar{\pi}^f \middle| s_H^f\right)\right]^2}$$

Consequently with $-X^f < -\bar{\pi}^f$ (father chooses sector *H*) and $\sigma_{HC} < \sigma_H^2$ we obtain:

$$\frac{f\left(-X^f \middle| s_H^f\right)}{F\left(-X^f \middle| s_H^f\right)} > \frac{f\left(-\bar{\pi}^f \middle| s_H^f\right)}{F\left(-\bar{\pi}^f \middle| s_H^f\right)}.$$

And therefore

$$\frac{d\Pr\left(J^s = C \middle| J^f = H; s_H^f\right)}{ds_H^f} = \frac{d\left(1 - G\left(-\bar{\pi}^s \middle| J^f = H; s_H^f\right)\right)}{ds_H^f} < 0.$$

If $\sigma_{HC} > \sigma_H^2$ we obtain $\dfrac{f\left(-X^f \middle| s_H^f\right)}{F\left(-X^f \middle| s_H^f\right)} < \dfrac{f\left(-\bar{\pi} \middle| s_H^f\right)}{F\left(-\bar{\pi} \middle| s_H^f\right)}$ with

$$\frac{d\Pr\left(J^s = C \middle| J^f = H; s_H^f\right)}{ds_H^f} = \frac{d\left(1 - G\left(-\bar{\pi}^s \middle| J^f = H; s_H^f\right)\right)}{ds_H^f} > 0.$$

■

c) For the case $\sigma_{HC} < \sigma_H^2$ the statement is obviously true. If $\sigma_{HC} = \sigma_H^2$ then (A2) holds with equality. As σ_{HC} increases or σ_H^2 decreases the effect of s_H^f on $\Pr\left(J^s = C\right)$ increases, until it is equal to the effect of s_C^f.

■

References

Becker, Gary S., and Nigel Tomes. 1979. "An equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy,* Vol. 87, 1153-89.

Becker, Gary S., and Nigel Tomes. 1986. "Human Capital and the rise and fall of families." *Journal of Labor Economics*, Vol. 4, S1-S39.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *American Economic Review*, Vol. 95, 437 – 449.

Bowles, Samuel and Herbert Ginitis. 2001. "The Inheritance of Economic Status: Education, Class and Genetics." In Marcus Feldman ed. *Genetics Behavior and Society* Oxford, Elsevier.

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman 2003. "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice." *International Economic Review*, Vol. 44, 361-422.

Carneiro, Pedro and James J. Heckman, 2002. "The Evidence on Credit Constraints in Post-Secondary Schooling." *The Economic Journal,* Vol. 112, Issue 482, 705-734.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil. 2005. "Understanding what Instrumental Variables Estimate: Estimating Marginal and Average Returns to Education." Unpublished Manuscript, Univ. Chicago, Dept. Econ.

Cameron, Steven V. and James J. Heckman. 1999. "Can Tuition Policy Combat Rising Wage Inequality?" In M. Kosters ed. *Financing College Tuition.* American Enterprise Institute, Washington D.C.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2007. "Interpreting the Evidence on Life Cycle Skill Formation" in E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, Amsterdam: Elsevier Science.

Gourieroux, Christian, and Alain Monfort. 1996. Simulation based econometric methods. New York: Oxford University Press.

Grawe, Nathan D. and Casey B. Mulligan. 2002. “Economic interpretations of intergenerational correlations.” *Journal of Economic Perspectives*, Vol. 16, 45-58.

Heckman, James J., and Bo E. Honore. 1990. “The Empirical Content of the Roy Model.” *Econometrica,* Vol. 58, 1121-1149.

Heckman, James J., Lance Lochner, and Petra E. Todd. 2006. “Earnings Functions, Rates of Return, and Treatment Effects:The Mincer Equation and Beyond.” In E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, Amsterdam: Elsevier Science.

Heckman, James J. and Jose Scheinkman. 1987. “The Importance of Bundling in a Gorman-Lancaster Model of Earnings” *The Review of Economic Studies*, Vol. 54, No. 2., 243-255.

Heckman, James J. and Guilherme L. Sedlacek. 1990. “Self-Selection and the Distribution of Hourly Wages.” *Journal of Labor Economics*, Vol. 8, S329-S363.

Mulligan, Casey B. 1997. Parental Priorities and Economic Inequality. Chicago: University of Chicago Press.

Piketty, Thomas. 2000. “Theories of Persistent Inequality and Intergenerational Mobility.” In A.B. Atkinson, F. Bourguignon (eds.), *Handbook of Income Distribution.* Amsterdam: North Holland.

Plug, Erik, and Wim Vijverberg. 2003. “Schooling, Family Background, and Adoption: Is it Nurture or is it Nature?” *Journal of Political Economy*, Vol. 111, 611-641.

Roy A. D. 1951. “Some thoughts on the distribution of earnings.” *Oxford Economic Papers* Vol. 3, 135-146.

Sattinger, Michael. 1993. “Assignment Models of the Distribution of Earnings” *Journal of Economic Literature,* Vol. 31, 831-880.

Solon, Gary. 1999. "Intergenerational Mobility in the Labor Market." In O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics,* Vol. 3a, Amsterdam: North Holland.

Willis, Robert J. and Sherwin Rosen. 1979. "Education and Self-Selection. *Journal of Political Economy* Vol. 87, 7-36.

Table 1

Data description

Father – Son Pairs
from the PSID

		Fathers (Weighted by number of sons)	Sons (Multiple sons per father)
Observations		469	853
Age (in 1970)	min	22	1
	max	65	19
	average	39.9	10.1
	std. dev.	8.52	5.37
Lifetime Hourly Wage std. dev.		0.47	0.52
Lifetime Hourly Wage Mean if edu>12		0.29	0.17
Lifetime Hourly Wage Std. Dev. If edu>12		0.46	0.53
Years of Education	mean	11.64	13.40
Years of Education	std. dev.	3.42	2.11
Fraction with more than 12 years of education		0.36	0.51
Return to one extra year of education, OLS		0.07	0.10
(s.e.)		0.00	0.01
Intergenerational correlation lifetime wage		0.3	
Intergenerational correlation education		0.45	

Table 2

Results from Structural Estimation

Parameter	ρ	σ_H	σ_C	b_H	b_C	$\pi_C^f - \pi_H^f$	$\pi_C^s - \pi_H^s$	c^f	c^s	Objective Function Value
Estimate	0.5154	0.3776	0.5624	0.1118	0.5249	-0.1026	0.0480	0.1161	0.0003	37.47
s.e.	0.132	0.014	0.035	0.067	0.042	0.068	0.080	0.070	0.076	

Note: The model is estimated by minimizing the weighted difference between the 21 data moments and the moments obtained from simulations of the model (see Table A1).
I use simulations with 10,000 father-son pairs.

Table 3

Regression of Lifetime-Wage of Sons on College Dummy and Fathers' features

Comparison of data and values implied by the model

Dependent Variable: Lifetime-Wage Son

	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6	Simulation 1	Simulation 2	Simulation 3	Simulation 4	Simulation 5	Simulation 6
College Son	0.35			0.28		0.25	0.41			0.33		0.32
s.e.	*0.03*			*0.03*		*0.04*						
Lifetime-Wage Father		0.34		0.25	0.26	0.21		0.39		0.27	0.34	0.26
s.e.		*0.04*		*0.04*	*0.04*	*0.04*						
College Father			0.28		0.17	0.1			0.26		0.11	0.04
s.e.			*0.04*		*0.04*	*0.04*						

Note: Columns 1 through 6 are separately estimated regression models with the same dependent variable and varying sets of independent variables.

Table 4

Lifetime-Wage of Sons by Sector Choice

Comparison of data and values implied by the model

Data

		Father: No College	Father: College	Father: Total
Son	No College	**-0.21**	**-0.01**	**-0.18**
	s.d.	0.44	0.51	0.46
	College	**0.08**	**0.25**	**0.17**
	s.d.	0.48	0.57	0.53
	Total	**-0.10**	**0.18**	**0**
	s.d.	0.48	0.56	0.53

Simulation

		Father: No College	Father: College	Father: Total
Son	No College	**-0.22**	**-0.20**	**-0.22**
	s.d.	*0.38*	*0.38*	*0.38*
	College	**0.09**	**0.29**	**0.19**
	s.d.	*0.47*	*0.52*	*0.50*
	Total	**-0.09**	**0.17**	**0**
	s.d.	*0.44*	*0.53*	*0.49*

Table 5

Lifetime-Wage of Fathers by Sector Choice

Comparison of data and values implied by the model

Data

		Father: No College	Father: College	Father: Total
Son	No College	**-0.20**	**0.14**	**-0.14**
	s.d.	0.38	0.43	0.42
	College	**-0.08**	**0.34**	**0.13**
	s.d.	0.40	0.45	0.48
	Total	**-0.16**	**0.29**	**0**
	s.d.	0.39	0.46	0.47

Simulation

		Father: No College	Father: College	Father: Total
Son	No College	**-0.22**	**0.05**	**-0.17**
	s.d.	*0.37*	*0.42*	*0.4*
	College	**-0.06**	**0.36**	**0.15**
	s.d.	*0.37*	*0.47*	*0.47*
	Total	**-0.16**	**0.28**	**0**
	s.d.	*0.44*	*0.52*	*0.47*

Table 6

Fraction of Father/Son Pairs by Sector Choice

Comparison of data and values implied by the model

		Father		
		No College	College	Total
Son	No College	**0.40**	**0.09**	**0.49**
	College	**0.25**	**0.26**	**0.51**
	Total	**0.64**	**0.36**	**1**

Data

		Father		
		No College	College	Total
Son	No College	**0.38**	**0.09**	**0.47**
	College	**0.27**	**0.26**	**0.53**
	Total	**0.64**	**0.36**	**1**

Simulation

Table 7

Probit Regression of College Son on Fathers' features

Comparison of data and values implied by the model

Dependent Variable: College Son

	Data 1	Data 2	Data 3	Data 4	Simulation 1	Simulation 2	Simulation 3	Simulation 4
Lifetime-Wage Father	0.86		0.53	0.50	1.04		0.80	0.71
s.e.	*0.10*		*0.11*	*0.14*				
College Father		0.95	0.74	0.73		0.87	0.58	0.56
s.e.		*0.09*	*0.10*	*0.11*				
Lifetime-Wage Father * College Father				0.10				0.21
s.e.				*0.23*				

Note: Columns 1 through 4 are separately estimated regression models with the same dependent variable and varying sets of independent variables.

Table A1

Data Moments vs. Moments implied by estimated model

Moment	Data	Simulation
$E\left(D^f\right)$	0.355	0.357
$E\left(D^s\right)$	0.508	0.528
$E\left(D^f D^s\right)$	0.263	0.268
$E\left(D^f v^f\right)$	0.102	0.104
$E\left(D^f v^s\right)$	0.065	0.062
$E\left(D^s v^f\right)$	0.068	0.083
$E\left(D^s v^s\right)$	0.086	0.105
$E\left(v^f v^s\right)$	0.074	0.089
$E\left(D^f D^s v^f\right)$	0.089	0.100
$E\left(D^f D^s v^s\right)$	0.066	0.079
$E\left(D^f v^f v^s\right)$	0.040	0.054
$E\left(D^s v^f v^s\right)$	0.034	0.061
$E\left(D^f D^s v^s v^f\right)$	0.036	0.053
$E\left(v^f v^f\right)$	0.219	0.223
$E\left(v^s v^s\right)$	0.277	0.253
$E\left(D^f v^f v^f\right)$	0.103	0.112
$E\left(D^f v^s v^s\right)$	0.125	0.118
$E\left(D^s v^f v^f\right)$	0.125	0.136
$E\left(D^s v^s v^s\right)$	0.157	0.160
$E\left(D^f D^s v^f v^f\right)$	0.084	0.097
$E\left(D^f D^s v^s v^s\right)$	0.100	0.101

Note: I normalize $E\left(v^g\right)=0$,

and define $D^g = 1$ if College educated.

$D^g = 0$ if High-School educated.

Table A2

Results from Structural Estimation

Values for cost of college obtained through separate calibration

Imposed cost of college	Parameter	ρ	σ_H	σ_C	b_H	b_C	$\pi_C^f - \pi_H^f$	$\pi_C^s - \pi_H^s$	Objective Function Value
$c^f = .2$	Estimate	0.151	0.414	0.540	0.160	0.533	0.096	0.191	47.16
$c^f = .2$	s.e.	0.112	0.014	0.027	0.060	0.044	0.049	0.032	
$c^f = .2$	Estimate	0.254	0.400	0.551	0.163	0.516	-0.074	0.109	42.38
$c^f = .1$	s.e.	0.089	0.013	0.024	0.064	0.042	0.042	0.031	

Note: The model is estimated by minimizing the weighted difference between the 21 data moments and the moments obtained from simulations of the model.
I use simulations with 10,000 father-son pairs.
I calibrate the cost of attending college.