The Return to College: Selection Bias and Dropout Risk*

Lutz Hendricks†        Oksana Leukhina‡

April 5, 2011

Very preliminary and incomplete.

Abstract

We study two long-standing questions: (i) What part of the measured return to education is due to selection? (ii) The ex post return to schooling appears higher than the return to most financial assets. How large are the contributions of various frictions to the “high” return to schooling? We focus in particular on the roles of college dropout risk, borrowing constraints, and learning about ability.

We develop and calibrate a model of school choice. Key model features are: (i) ability heterogeneity, (ii) students learn about their abilities while in college, (iii) borrowing constraints, (iv) dropping out of college is a choice.

Preliminary results indicate that the probability of graduating from college increases strongly with ability. Most college dropouts are students of intermediate abilities who try college in part to learn about their abilities and in part because of the option value of receiving a large earnings gain upon graduation. Ability selection accounts for about 80% of the measured college wage premium.


Key words: Education. College dropout risk.

*The usual disclaimer applies.
†University of North Carolina, Chapel Hill, CESifo, Munich, and CFS, Frankfurt; lutz@lhendricks.org
‡University of Washington; oml@u.washington.edu
1 Introduction

The question:

1. Part of the college wage premium is an ability premium. How large is this part?

2. The ex post return to college is high relative to the returns earned by financial assets (Heckman, Lochner, and Todd, 2008). What is the contribution of college completion risk, ability selection, and borrowing constraints towards sustaining the high rate of return? How large is the ex ante rate of return for persons of various abilities?

The approach: We develop a model of school choice with the following features:

1. Individuals differ in their abilities. Abilities affect adult earnings and the likelihood of graduating from college.

2. Individuals are uncertain about their abilities when choosing whether or not to attend college.

3. College students are borrowing constrained.

4. College has a consumption value.

5. In contrast to much of the literature, our model does not feature a “psychic cost” of schooling.

We include these model features for the following reasons:

1. Ability heterogeneity: it is the basis for selection bias.

2. Uncertain abilities: one-third of students attempting 4 year colleges drop out without earning a bachelor’s degree. One possible reason is that students learn about their abilities or graduation prospects as they move through college (Manski, 1989).

3. Borrowing constraints: A large literature discusses their importance. Even if most students have access to sufficient borrowing so they can finance college expenditures, borrowing constraints affect selection into college by ability (Belley and Lochner, 2007).
4. Consumption value of college: In the data, some students attend college even though their probability of graduating, conditional on observable characteristics, is low.

5. No psychic cost: Many existing models of school choice attribute a large share of individual variation to a “psychic cost.” Examples include Cunha, Heckman, and Navarro (2005) and Navarro (2008). We agree with Heckman, Lochner, and Todd (2006) that “explanations based on psychic costs are intrinsically unsatisfactory” (p. 436).

We calibrate the model for white men in the 1960 birth cohort. Our main data source is the NLSY79, which provides us with schooling, cognitive test scores, and partial earnings histories. We complete the earnings histories using CPS data.

**Results.** Preliminary results are as follows:

1. About 80% of the measured college wage premium is due to ability selection.

2. Uncertainty about individual abilities and thus college completion prospects accounts for a large share of college dropouts. This is consistent with the argument proposed by Manski (1989).

3. College graduation prospects vary strongly with ability. Students in the lowest ability decile have a less than 3% chance of graduating from college, while student in the top decile have a 90% chance of graduating.

## 2 The Model

### 2.1 Model Outline

We study a partial equilibrium model of school choice. We follow a single cohort, starting at the date of high school graduation, through college (if chosen), work, and retirement.

Figure 1 summarizes the agent’s life-cycle. At the start of age 1, all agents graduate from high school. They are endowed with assets $k_1$, ability $a$, a signal about ability $m$, and a net price of attending college $q$. Ability is not observed until the agent starts working.
Agents choose whether to start working as high school graduates immediately or to attempt college. Agents are not allowed to return to school after they started work.

Working agents choose a consumption path to maximize lifetime utility subject to a lifetime budget constraint that equates the present value of income to the present value of consumption spending.

While in college, students accumulate college credits \( n \). Once a student reaches \( n_{\text{grad}} \) credits he graduates and works as a college graduate. The accumulation of credits is stochastic. More able students accumulate credits faster.

In each period, students pay a tuition cost \( q \), they pay for consumption \( c_F \), they attempt \( n_c \) credits and succeed in a random subset. They update their beliefs about their abilities and how long it will take to graduate. They decide whether to continue studying next period or drop out and work as a college dropout. A student who fails to achieve enough credits by the end of year \( T_c \) must drop out of college and start working. The cost of college evolves randomly.

The details are described next. We motivate our assumptions in Section 2.8.

### 2.2 Endowments

At birth each person draws the following endowments:

1. an ability \( a \) that takes on value \( \hat{a}_i \) with probability \( p_{a,i}; i = 1, ..., N_a \). Ability is not known to the agent until he starts working.
2. a signal \( m \) that takes on value \( \hat{m}_i \) with probability \( p_{m,i}(\hat{a}_j); i = 1, ..., N_m \).
3. a net price of attending college \( q \) that takes on value \( \hat{q}_i \) with probability \( p_{q,i}; i = 1, ..., N_q \).
4. initial assets \( k_1 \) that are drawn from a continuous distribution. \( k_1 \) is not correlated with ability, conditional on \( m \). It is therefore not useful for updating beliefs about \( a \).
5. \( n = 0 \) completed college credits.
Figure 1: Model Timing

(a) Choices at age 1

\[ V_W(k_1, 0, m, HS, 1) \]

Draw endowments: \( k_1, n_1 = 0, m, q_1 \)

↓

Work as HS graduate

Try college

\[ V_C(k_1, 0, m, q_1, 1) \]

↓

Learn ability

\[ V(k_1, a, HS, 1) \]

(b) Timing while in college

\[ V_C(k_t, n_t, m, q_t, t) \]

Choose \( c_{Ft}, k_{t+1} \)

↓

Draw \( n_{t+1}, q_{t+1} \)

Graduate

\[ V_W(k_{t+1}, n_{t+1}, m, CG, t+1) \]

↓

Learn ability

\[ V(k_{t+1}, a, n_{t+1}, CG, t+1) \]

Drop out

\[ V_W(k_{t+1}, n_{t+1}, m, CD, t+1) \]

↓

Learn ability

\[ V(k_{t+1}, a, n_{t+1}, CD, t+1) \]

Study in \( t + 1 \)

\[ V_C(k_{t+1}, n_{t+1}, m, q_{t+1}, t + 1) \]
2.3 Preferences

Agents enter the model at age 1 and live until age $T$. They consume two goods: a market good $c_F$ and a non-market good $c_L$. Expected utility is given by

$$E_0 \sum_{i=1}^{T} \beta^i u(c_t)$$

(1)

where $c_t = [c_{Ft}^\rho + c_{Lt}^\rho]^{1/\rho}$ and $u(c) = c^{1-\phi}/(1-\phi)$.

2.4 Choices at Age 1

At the beginning of life, the agent chooses whether to attempt college or work as a high school graduate. We summarize the value of working by the value function $V_W(k_\tau, n, m, s, \tau)$ where $\tau$ is the age at which work starts, $k_\tau$ denotes the level of assets, $n$ is the number of completed college credits, $s$ is the level of completed schooling. $s$ takes on the values HS for high school graduates, CD for college dropouts and CG for college graduates. Working as a high school graduate yields value $V_W(k_1, 0, m, HS, 1)$. Section 2.5 describes how $V_W$ is determined.

We summarize the value of starting year $t$ of college by $V_C(k, n, m, q, t)$. Section 2.6 describes how $V_C$ is determined. An agent chooses to start college if $V_C(k_1, 0, m, q, 1) > V_W(k_1, 0, m, HS, 1)$.

2.5 Work

Upon completing schooling, the worker learns her ability. At this point all uncertainty has been resolved and the ability signal $m$ no longer matters.

The worker’s value $V(k_\tau, a, n, s, \tau)$ is determined as follows. Let $e^{a+\mu sn}Y(s, \tau)$ denote the present value of lifetime earnings. The worker chooses market consumption to solve

$$V(k_\tau, a, n, s, \tau) = \max_{\{c_{Fi}\}} \sum_{i=\tau}^{T} \beta^{t-i} u(c[c_{Fi}, \hat{c}_L])$$

(2)

subject to the budget constraint

$$e^{a+\mu sn}Y(s, \tau) + Rk_\tau = \sum_{i=\tau}^{T} c_{Fi}R^{t-i}$$

(3)
where $R$ is the gross interest rate. The worker buys market goods, $c_F$, at price 1. He receives a fixed amount of the non-market good, $\hat{c}_L$, for free. We discuss the role of the non-market good in Section 2.8.

Before ability is revealed, the value of working is given by

$$V_W(k_\tau, n, m, s, \tau) = \mathbb{E}_a \left\{ V(k_\tau, a, n, s, \tau) | n, m, \tau \right\}$$

(4)

$$= \sum_{i=1}^{N_a} V(k_\tau, \hat{a}_i, s, \tau) \mathbb{P}(\hat{a}_i | n, m, \tau)$$

(5)

where $\mathbb{P}(\hat{a}_i | n, m, \tau)$ is the agent’s belief about her ability, which we derive in Section 2.7.

### 2.6 College

The value of being in college at age $t$, $V_C(k_t, n_t, m, q, t)$, is determined as follows. A student enters the period with assets $k_t$ and earns capital income $Rk_t$. Assets may be negative. She then pays tuition $q$. She chooses consumption $c_{Ft}$ so that next period’s assets are determined by the budget constraint $k_{t+1} = Rk_t - c_{Ft} - q$. She receives a fixed amount of non-market consumption, $\bar{c}_L$, for free. Borrowing is constrained by $k_{t+1} \geq k_{\text{min},t+1}$. If the student lacks the funds to pay for tuition, so that $Rk_t < q + k_{\text{min},t+1}$, we set $V_C(k_t, n_t, m, q, t) = -\infty$.

After choosing consumption, the student attempts $n_c$ courses and completes each with probability $Pr_c(a)$. More able students are more likely to pass a course: $Pr_c'(a) > 0$. Using the number of credits completed, $n_{t+1}$, the student updates her beliefs about $a$. She then draws a new college cost $q'$ from the transition matrix $Pr_q(q'|q)$ and decides whether to work or study in period $t+1$.

The option of studying is not available if

1. $n_{t+1} \geq n_{\text{grad}}$: the student graduates from college and works as a college graduate with continuation value $V_W(k_{t+1}, n_{t+1}, m, CG, t+1)$.

2. $n_{t+1} < n_{\text{grad}}$ and $t = T_c$: the student fails to earn enough credits in the last year of college. She must work as a college drop-out with continuation value $V_W(k_{t+1}, n_{t+1}, m, CD, t+1)$.
3. $k_{t+1}$ is too low to pay for tuition next period.

Otherwise the student chooses to remain in college if the continuation value in college is greater than that of working as a college dropout: $V_W(k_{t+1}, n_{t+1}, m, CD, t + 1) > V_C(k_{t+1}, n_{t+1}, m, q_{t+1}, t + 1)$. The Bellman equation is therefore given by

$$V_C(k_t, n_t, m, q_t, t) = \max_{k_{t+1} \geq k_{\text{min}, t+1}} u(c[Rk_t - k_{t+1} - q_t, \bar{c}_L]) + \beta \sum_{q_{t+1}} Pr_q(q_{t+1} | q_t) \sum_{m_{t+1}} Pr(n_{t+1} | n_t, m_t) V_{EC}(k_{t+1}, n_{t+1}, m, q_{t+1}, t + 1)$$ (6)

where $V_{EC}(k, n, m, q, t) = V_W(k, n, m, CG, t)$ if the student graduates from college, $V_{EC}(k, n, m, q, t) = V_W(k, n, m, CD, t)$ if the student is forced to drop out of college, and $V_{EC}(k, n, m, q, t) = \max \{ V_C(k, n, m, q, t), V_W(k, n, m, CD, t) \}$ if the student can choose whether to work or study next period.

### 2.7 Probabilities

We now derive the probabilities governing how students accumulate credits and form beliefs about their abilities. The probability of passing a course, $Pr_c(a)$, is an exogenous, increasing function of ability. All other probabilities and beliefs are found using Bayes’ Rule.

The probability of passing $j$ courses in one year is Binomial. Denote it by $Pr_n(j|a)$. Then

$$Pr(n_{t+1} | n_t, m_t) = \sum_{i=1}^{N_a} Pr_n(n_{t+1} - n_t | \hat{a}_i) Pr(\hat{a}_i | n_t, m_t)$$ (7)

The agent’s beliefs about her ability follow from Bayes’ Rule:

$$Pr(\hat{a}_i | n_t, m_j, t) = \frac{p_{a,i} p_{m,j}(\hat{a}_i) Pr(n_t | \hat{a}_i, t)}{Pr(n_t, m_j | t)}$$ (8)

Pr$(n_t | \hat{a}_i, t)$ is given by Binomial formula for $n_t$ successes out of $(t - 1)n_c$ draws. Also from Bayes’ Rule, we have $Pr(n, m | t) = \sum_i p_{a,i} Pr(n_t, m | \hat{a}_i, t)$ and $Pr(n_t, m | \hat{a}_i, t) = Pr(m | \hat{a}_i) Pr(n_t | \hat{a}_i, t)$. 8
2.8 Discussion of Model Assumptions

We motivate key model assumptions.

1. Heterogeneity in assets.

   Our analysis focuses on ability selection and the risk of dropping out of college. The literature emphasizes borrowing constraints as a potential friction to ability selection.

2. The non-market consumption good.

   A long-standing puzzle is: why don’t college students consume more? Many students do not take out available and subsidized loans to smooth consumption between college and work periods. A second puzzle is: why do so many students drop out of college? Some low ability students attempt college, even though their ex ante probability of graduating is low. Our model can potentially resolve these puzzles.

   In our model, college has a consumption value. The idea is that attending college provides non-market consumption, such as socializing with other students, that cannot be purchased. Free non-market consumption makes college attractive, even for those who do not expect to graduate.

   Upon graduation from college, the non-market good is no longer available ($\hat{c}_L = 0$). As a result, the marginal utility of the market good increases. Even if individuals smooth the marginal utility of market consumption, consumption spending jumps upon graduation.

   A model with a single consumption good that is purchased at all stages of life would predict, counterfactually, that high asset students smooth market consumption between study and work periods. It would also predict that all high ability students max out available loans and parental transfers.

3. Individual abilities are not perfectly known.

   Manski (1989) argues that learning about ability may explain why many students drop out of college before earning a degree. We wish to investigate the quantitative importance of this explanation.

   However, the reader should keep in mind an important caveat. Since our model lacks a “psychic cost,” learning about ability takes on its role in the calibration: it absorbs
the effects of unmodeled heterogeneity. In the robustness analysis, we plan to explore how our results change when a psychic cost is added to the model.

4. Students accumulate grades as they advance through college. A simpler alternative would be to model dropping out as a random function of ability. However, modeling dropping out as a choice has important benefits. We can use data on the timing and the characteristics of dropouts in the calibration.

3 Calibration

We calibrate the model parameters to match moments for white men born in 1960. Our main data source is the NLSY79, which is a representative, ongoing sample of persons born between 1957 and 1964. We retain all white men who participated in the ASVAB battery of aptitude test. We include members of the supplemental samples, but use weights to offset the oversampling of low income persons. Appendix B provides additional detail.

We use cognitive test scores are noisy measures of individual abilities. Specifically, we use the 1980 Armed Forces Qualification Test (AFQT) percentile rank (variable R1682). The AFQT aggregates a battery of aptitude test scores into a scalar measure. The tests cover numerical operations, word knowledge, paragraph comprehension, and arithmetic reasoning (see NLS User Services 1992 for details). We remove age effects by regressing AFQT scores on the age at which the test was administered (in 1980). We transform the residual so that it has a standard Normal distribution. In mapping AFQT scores to the model, we assume that AFQT scores are noisy measures of the agent’s ability signal: $AFQT_i = m_i + \varepsilon_{AFQT,i}$ where $\varepsilon_{AFQT,i} \sim N(0, \sigma_{AFQT})$ is a measurement error term.

We adopt the following normalizations:

1. The period is one year.

2. The unit of account is $10,000 in year 2000 prices. We use the Consumer Price Index (all wage earners, all items, U.S. city average) reported by the Bureau of Labor Statistics to convert dollar figures into units of account.

3. The non-market good is only consumed while in college: $\dot{c}_L = 0$. 
We set the following parameters based on outside evidence:

1. Demographics: Agents enter the model at age 19, they retire at age 65 ($T_R = 65 - 19$) and die at age 80 ($T = 80 - 19$).

2. Preferences: Utility is logarithmic in the composite consumption good ($\phi = 1$). The discount factor is $\beta = 0.98$. The elasticity of substitution between the two consumption goods is set to 4 ($\rho = 0.75$). We lack evidence on this parameter and plan to explore alternative values.

3. Prices: The gross interest rate is $R = 1.03$.

4. College: Most students graduate from college in either four or five years (Bowen, Chingos, and McPherson, 2009). We therefore set the maximum duration of college to $T_C = 5$. The number of credits needed to graduate is set to $n_{\text{grad}} = 8$. In each year, students attempt $n_c = 3$ credits. This number is set so that graduating in 3 years is possible, but not common, which accords with the data.

In reality, students typically needs 120 credits, which are obtained from 40 courses of 3 credits each. Increasing the number of credits a student takes would increase the number of ability signals she receives in each period, which may affect the rate of learning. It is, however, computationally costly.

5. Initial assets: We plan to estimate the distribution from data on parental transfers, grants, and loans. For now, we use a uniform distribution on a 9 point grid between 0 and $\$60,000$. We assume that assets are not correlated with $m$, but plan to relax this assumption later.

6. Borrowing constraints: For college students of the NLSY79 cohorts most loans taken out during college are Stafford loans (see Johnson 2010). Until 1986, students could borrow $\$2,500$ in each year of college up to a total of $\$12,500$. We ignore the restriction that loan amounts cannot exceed college related expenditures. We set $k_{\text{min},t} = -\min(\$12,500, \$2,500t)$, where dollar values are converted into units of account.

7. The direct cost of college is set to $\$3,000$. Bowen, Chingos, and McPherson (2009) show that, for public universities, the cost of tuition net of scholarships and grants is small.
The following parameters are calibrated jointly:

1. The distribution of abilities approximates a Normal distribution with mean 0 and standard deviation $\sigma_a$ on 5 point grid.

2. Ability signals: $p_{m,i}(\hat{a}_j)$ approximates a Normal distribution with mean $\hat{a}_j$ and standard deviation $\sigma_m$.

3. The standard deviation of noise in AFQT scores: $\sigma_{AFQT}$.

4. The lifetime earnings functions $Y(s, \tau)$.

5. The probability of passing a credit is given by $[1+\gamma_1 e^{-\gamma_2 a}]^{-1}$. $\gamma_1$ and $\gamma_2$ are calibrated.

6. The consumption value of college, $\bar{c}_L$.

### 3.1 Data and Calibration Targets

This section describes the data sources and calibration targets. We do not give a structural interpretation to any of the data moments. Instead, we compute comparable summary statistics from simulated life histories of model agents.

#### 3.1.1 NLSY79

We take several data moments from Hendricks and Schoellman (2011). They are estimated from white men in the NSLY79 sample, which covers the cohorts born between 1957 and 1964. Table 1 shows summary statistics for this sample. From the NLSY79 we construct the following moments:

1. The “wage return to AFQT,” which is estimated by regressing log wages at 20 years of experience on standard Normal AFQT scores. The regression coefficient is $\beta_{AFQT} = 0.104$ (s.e. 0.017).

2. The joint distribution of AFQT scores and schooling. For each AFQT quartile, Table 2 reports the fraction of high school graduates who complete a given level of schooling.

3. The fraction of persons who attain each school level (see Table 1).
Table 1: Summary statistics for the NLSY79 sample


<table>
<thead>
<tr>
<th></th>
<th>School Attainment</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;HS</td>
<td>HS</td>
<td>SC</td>
<td>C</td>
</tr>
<tr>
<td>Avg. school</td>
<td>9.5</td>
<td>11.7</td>
<td>13.3</td>
<td>17.0</td>
</tr>
<tr>
<td>Real wage at age 40</td>
<td>13.9</td>
<td>17.2</td>
<td>22.0</td>
<td>39.0</td>
</tr>
<tr>
<td>Adj. wage at age 40</td>
<td>13.9</td>
<td>17.2</td>
<td>22.6</td>
<td>35.3</td>
</tr>
<tr>
<td>AFQT percentile</td>
<td>0.13</td>
<td>0.34</td>
<td>0.50</td>
<td>0.75</td>
</tr>
<tr>
<td>N</td>
<td>163</td>
<td>642</td>
<td>644</td>
<td>493</td>
</tr>
</tbody>
</table>

Source: Hendricks and Schoellman (2011)

4. The mean log lifetime earnings for each school group. The corresponding model object is $E\{a + \mu_s n + \log(Y(s, \tau))|s\}$

**Lifetime earnings.** To estimate lifetime earnings, we assume that workers’ earnings are given by

$$g(t|n, s, \tau) = a + \mu_s n + w_{s,t} + g(t|s, \tau)$$

(9)

where

$$g(t|s, \tau) = \begin{cases} f(t - \tau|s) & \text{if } t \leq T_R \\ f_R(t|s) & \text{if } t > T_R \end{cases}$$

(10)

We lack data to estimate how the age earnings profile varies with $\tau$. We therefore assume that the experience earnings profile is the same for all $\tau$. In retirement, the agent receives a fixed age retirement income profile.

The early years of the age earnings profile are estimated from NLSY79 data. We keep all men with known schooling and earnings in each year. We regress log earnings (above a threshold of $1,000 in year 2000 prices) on an age quartic, cohort dummies, and the unemployment rate. The latter absorbs cyclical year effects. We compute the implied age earnings profile for the 1960 birth cohort, with a starting age at $1960 + \tau_s$, where $\tau_s$ is a typical age at which a person with schooling level $s$ starts to work. We set $\tau_s$ to 19 for high school graduates, 21 for college dropouts, and 23 to for college graduates.
Table 2: AFQT scores and schooling. NLSY79 data.

<table>
<thead>
<tr>
<th>AFQT quartile</th>
<th>&lt;HS</th>
<th>HS</th>
<th>SC</th>
<th>C+</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.861</td>
<td>0.417</td>
<td>0.184</td>
<td>0.010</td>
</tr>
<tr>
<td>2</td>
<td>0.118</td>
<td>0.338</td>
<td>0.317</td>
<td>0.105</td>
</tr>
<tr>
<td>3</td>
<td>0.021</td>
<td>0.194</td>
<td>0.309</td>
<td>0.294</td>
</tr>
<tr>
<td>4</td>
<td>0.000</td>
<td>0.051</td>
<td>0.190</td>
<td>0.591</td>
</tr>
<tr>
<td>Fraction</td>
<td>0.060</td>
<td>0.323</td>
<td>0.328</td>
<td>0.289</td>
</tr>
<tr>
<td>N</td>
<td>163</td>
<td>642</td>
<td>644</td>
<td>493</td>
</tr>
</tbody>
</table>

Source: Hendricks and Schoellman (2011)

Individuals in the NLSY79 are only observed until at most age 48 (for the 1960 cohort). We need to fill in the remaining years until retirement. We use CPS data to estimate a cohort invariant age earnings profile for each school group using individuals aged $\tau_s$ through $T_R$. This uses a similar regression to the one we apply to the NLSY79 data. We find the level parameter so that the means of the CPS profile matches that of the NLSY profile for the last 5 years observed in the NLSY. In the model, the estimated profiles correspond to $E\{a + \mu_s n | s, \tau_s\} + f(t - \tau_s | s)$.

We use a similar regression to estimate the age retirement income profile, $E\{a + \mu_s n | s, \tau_s\} + f_R(t | s)$. [Details to be written.]

The results are shown in Figure 2. Each panel shows the log age earnings profiles estimated from the NLSY and CPS samples for the 1960 cohort, the age quartic estimated from the CPS over the entire working age range, and the age earnings profile used as a model input. The latter multiplies the profile conditional on working by the fraction working, shown in Figure 3. The Figure assumes that the agent starts at age $\tau = \tau_s$.

We can now compute $\log(Y(s, \tau)) + E\{a + \mu_s n | s, \tau_s\}$ as the present value of the estimated age earnings profile. Note that this identifies $Y(s, \tau)$ up to a scale factor that is the same for all $\tau$. The calibration algorithm searches over these scale factors, which we denote by $\hat{Y}(s)$. For each guess, the household problems are solved. The algorithm iterates until the simulated data match the estimated lifetime earnings levels.
Figure 2: Age earnings profiles
3.2 Model Parameters

Table 3 shows the calibrated model parameters. Central for our findings is the dispersion of abilities, $\sigma_a$. A one standard deviation increase in ability raises earnings by 27%. Together with a relatively small value of signal noise ($\sigma_m = 0.12$), this implies strong school sorting by ability.

3.3 Model Fit

Table 4 compares the model outcomes with the calibration targets for each school group. Figure 2 shows the relationship between schooling and AFQT scores. For each school group, the Figure shows the fraction of persons that belong to each AFQT quartile. We compare the model with NLSY79 data. The model comes close to replicating all calibration targets.
Table 3: Calibrated parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_a$</td>
<td>Standard deviation of ability</td>
<td>0.29</td>
</tr>
<tr>
<td>$\sigma_m$</td>
<td>Standard deviation of signal</td>
<td>0.12</td>
</tr>
<tr>
<td>$\sigma_{AFQT}$</td>
<td>Standard deviation of AFQT</td>
<td>0.20</td>
</tr>
<tr>
<td>$\bar{c}_L$</td>
<td>Consumption in college</td>
<td>1.35</td>
</tr>
<tr>
<td>$\bar{Y}_{HS}$</td>
<td>Lifetime earnings factor</td>
<td>1.29</td>
</tr>
<tr>
<td>$\bar{Y}_{CG}$</td>
<td>Lifetime earnings factor</td>
<td>0.75</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Earnings gain for each completed credit</td>
<td>0.020</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>Governs probability of passing a course</td>
<td>3.42</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>Governs probability of passing a course</td>
<td>2.77</td>
</tr>
</tbody>
</table>

Table 4: Model fit

<table>
<thead>
<tr>
<th></th>
<th>School group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HS</td>
</tr>
<tr>
<td>Fraction</td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>0.34</td>
</tr>
<tr>
<td>Model</td>
<td>0.35</td>
</tr>
<tr>
<td>Lifetime earnings</td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>66.5</td>
</tr>
<tr>
<td>Model</td>
<td>66.7</td>
</tr>
<tr>
<td>Mean AFQT percentile</td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>0.34</td>
</tr>
<tr>
<td>Model</td>
<td>0.28</td>
</tr>
</tbody>
</table>
Figure 4: Joint distribution of schooling and AFQT
Table 5: Outcomes by schooling

<table>
<thead>
<tr>
<th>School group</th>
<th>HS</th>
<th>CD</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean log ability</td>
<td>-0.25</td>
<td>-0.02</td>
<td>0.28</td>
</tr>
<tr>
<td>Mean log lifetime earnings</td>
<td>4.20</td>
<td>4.46</td>
<td>4.94</td>
</tr>
</tbody>
</table>

4 Results

4.1 Selection Bias

Our first question is: What fraction of the college wage premium represents selection bias? Table 5 shows the selection bias implied by the model. In the model, the mean log wage of school group $s$ is given by a skill price plus the mean ability of workers in that group, $E(a|s)$. Selection bias accounts for $E(a|CG) - E(a|HS) = 0.6$. To put this into perspective, Table 5 also shows mean log lifetime earnings, discounted to the age at which agents start working, for each school group. College graduates earn 0.72 log points more than high school graduates. 83% of this gap (0.6/0.72) is due to selection bias.

The ability gap between college graduates and high school graduates is much larger than many previous estimates. For example, the ability gap in Hendricks and Schoellman (2011) is only 0.25.

School choice at age 1. To illustrate selection into college, Figure 5 shows the decision whether or not to try college at age 1. The top panels show the values of choosing work as a high school graduate and of choosing college at age 1. Both are shown in consumption equivalents. The bottom panel shows the difference between the two values. Students with more assets and with better ability signals find it profitable to try college. The ability signal matters more for the choice in the sense that all agents with the highest ability signal choose college regardless of asset levels, while students with the lowest signal do not. Assets only matter for students with intermediate ability signals.

An interesting feature is the strong asymmetry in the value gain from attempting college. High ability agents reap large gains from attending college. Agents receiving the highest are willing to pay about $60,000 for the opportunity to attend college. By contrast, agents
in the lowest ability group would attend college in exchange for a payment of only $3,000. Of course, most of these students would drop out of college after the first year. This is one reason why the required payment is so small.

This finding raises the concern that college attendance in the model is unreasonably sensitive to changes in tuition. To address this issue, we compute the effect of reducing tuition by $1,000. A sizeable empirical literature estimates the effects of reducing tuition on college attendance. Dynarski (2003) summarizes this literature as well as her own estimates as follows: a $1,000 reduction in the cost of attending college (in 1998 prices) leads to a 6% to 8% increase in college attendance (a 3 to 4 percentage point increase while 50% of the sample attend college in the baseline case). In our model, reducing the cost of college by $1,000 increases college attendance by 9% (a 6.2 percentage point increase while 66% of the sample attend college in the baseline case).
4.2 Return to College and Ability

We next examine the rate of return to college for students of different ability levels. Table 6 summarizes the outcomes achieved, on average, by students of each ability level.

One striking finding is that very few students of low abilities graduate from college (<0.1% in the lowest ability group). By contrast, the probability of graduating is over 90% of the most able students. More able students also graduate earlier and drop out later.

Low ability students who attempt college generally do so on the basis of incorrect ability signals. The mean signal of college students in the lowest ability group is above 3. Conversely, high ability students who do not attempt college do so on the basis of low ability signals (mean 1.9). There is strong sorting by ability signal. The model thus supports Manski’s (1989) emphasis on the role of learning about ability.

Table 7 shows the incentives for attending college. For each ability level, the table shows the lifetime incomes earned for high school dropouts and for students who attend college until they either graduate or are forced to leave without a degree. Because the probability of graduating is low for students in the lower ability groups, attempting to complete college reduces their lifetime incomes.

Even though college graduates earn 80% more over their lifetimes than do high school graduates, the gain from completing college is only 12%. The large gap is due to selection.

4.3 Identification

We investigate whether the data moments used in the calibration identify the model parameters.

Lower ability dispersion. The main concern is whether the dispersion of abilities is identified. It is a major determinant of selection bias in the college wage premium. Lower values of $\sigma_a$ could imply less selection bias. To address this possibility, we recalibrate the model while fixing $\sigma_a$ at 0.15, which is the preferred value in Hendricks and Schoellman (2011).

We find that the model attains all calibration targets nearly as well as the baseline model. The exception is that the lifetime earnings gap between college graduates and high school
Table 6: Outcomes By Ability

<table>
<thead>
<tr>
<th>Ability</th>
<th></th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative fraction (pct)</td>
<td>17.0</td>
<td>42.5</td>
<td>71.6</td>
<td>91.0</td>
<td></td>
</tr>
<tr>
<td>Dropout probability</td>
<td>100.0</td>
<td>100.0</td>
<td>53.8</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>Dropout probability, $\tau = T_c + 1$</td>
<td>100.0</td>
<td>100.0</td>
<td>38.4</td>
<td>0.0</td>
<td></td>
</tr>
</tbody>
</table>

Fraction by school

<table>
<thead>
<tr>
<th>School</th>
<th></th>
<th>0.92</th>
<th>0.47</th>
<th>0.05</th>
<th>0.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>HS</td>
<td>Fraction</td>
<td>0.08</td>
<td>0.53</td>
<td>0.51</td>
<td>0.04</td>
</tr>
<tr>
<td>CD</td>
<td>Fraction</td>
<td>0.00</td>
<td>0.00</td>
<td>0.44</td>
<td>0.96</td>
</tr>
</tbody>
</table>

HS

<table>
<thead>
<tr>
<th></th>
<th>20.9</th>
<th>18.6</th>
<th>1.9</th>
<th>0.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction</td>
<td>61.7</td>
<td>76.8</td>
<td>95.5</td>
<td>118.7</td>
</tr>
<tr>
<td>Lifetime earnings</td>
<td>2.0</td>
<td>3.1</td>
<td>3.9</td>
<td>4.5</td>
</tr>
<tr>
<td>Mean signal</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

CD

<table>
<thead>
<tr>
<th></th>
<th>2.0</th>
<th>22.9</th>
<th>22.7</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction</td>
<td>59.8</td>
<td>73.7</td>
<td>89.6</td>
<td>117.8</td>
</tr>
<tr>
<td>Lifetime earnings</td>
<td>3.6</td>
<td>4.8</td>
<td>6.1</td>
<td>7.8</td>
</tr>
<tr>
<td>Mean signal</td>
<td>2.0</td>
<td>2.5</td>
<td>4.9</td>
<td>3.6</td>
</tr>
<tr>
<td>Age at start of work</td>
<td>0.0</td>
<td>0.4</td>
<td>3.4</td>
<td>3.8</td>
</tr>
<tr>
<td>Completed credits</td>
<td>n/a</td>
<td>n/a</td>
<td>100.8</td>
<td>132.8</td>
</tr>
<tr>
<td>Mean signal</td>
<td>n/a</td>
<td>n/a</td>
<td>6.2</td>
<td>7.9</td>
</tr>
<tr>
<td>Age at start of work</td>
<td>n/a</td>
<td>n/a</td>
<td>6.1</td>
<td>4.3</td>
</tr>
</tbody>
</table>

CG

<table>
<thead>
<tr>
<th></th>
<th>0.0</th>
<th>0.0</th>
<th>18.8</th>
<th>23.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction</td>
<td>n/a</td>
<td>n/a</td>
<td>100.8</td>
<td>132.8</td>
</tr>
<tr>
<td>Lifetime earnings</td>
<td>n/a</td>
<td>n/a</td>
<td>6.2</td>
<td>7.9</td>
</tr>
<tr>
<td>Mean signal</td>
<td>n/a</td>
<td>n/a</td>
<td>6.1</td>
<td>4.3</td>
</tr>
</tbody>
</table>


Table 7: Lifetime Incomes By Ability and College Choice

<table>
<thead>
<tr>
<th>Ability</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative fraction (pct)</td>
<td>17.0</td>
<td>42.5</td>
<td>71.6</td>
<td>91.0</td>
</tr>
<tr>
<td>Prob(CG)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.62</td>
<td>1.00</td>
</tr>
<tr>
<td>Y HS</td>
<td>61.7</td>
<td>76.8</td>
<td>95.5</td>
<td>118.7</td>
</tr>
<tr>
<td>Y try college</td>
<td>48.9</td>
<td>63.1</td>
<td>93.9</td>
<td>131.7</td>
</tr>
<tr>
<td>Y try / Y HS</td>
<td>0.79</td>
<td>0.82</td>
<td>0.98</td>
<td>1.11</td>
</tr>
<tr>
<td>Y CD</td>
<td>48.9</td>
<td>63.1</td>
<td>86.0</td>
<td>109.2</td>
</tr>
<tr>
<td>Y CD / Y HS</td>
<td>0.79</td>
<td>0.82</td>
<td>0.90</td>
<td>0.92</td>
</tr>
<tr>
<td>Y CG</td>
<td>61.9</td>
<td>77.5</td>
<td>98.9</td>
<td>131.7</td>
</tr>
<tr>
<td>Y CG / Y HS</td>
<td>1.00</td>
<td>1.01</td>
<td>1.04</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Notes: The table shows mean lifetime incomes, conditional on ability and completed schooling. “Y try college” denotes lifetime income conditional on trying college. Lifetime income consists of earnings net of tuition costs.

graduates is 15% lower than in the data. The implied selection bias, \(E(a|CG) - E(a|HS) = 0.25\), is only about half of the baseline result, but still economically significant. It is interesting to note that our model yields essentially the same as that of Hendricks and Schoellman (2011) with the same value of \(\sigma_a\), even though the two models are very different.

Learning about ability.

5 Conclusion

To be written.
References


A Appendix: CPS Data

We estimate the age-earnings profile using Current Population Survey for 1964-2008, restricting our attention to 18-65 white men born between 1901 and 1986, not living in group quarters, and not reporting any farm income. We further restrict our attention to individuals with available relevant demographic information, wage and business income and personal weights. Because income and wage information refer to earnings in the year prior to the survey, the work years we considered are 1963-2007. We subtract 1 from respondent’s age, so it refers to the year for which the earnings are reported.

We obtain the annual unemployment rate for all years from the Bureau of Labor Statistics. We identify the self-employed using the variable classwly (values 10,13,14). For the self-employed, we define earned income as the sum of their wage and business income. The survey questions is phrased in a way that business income refers to the owner’s earnings in this business. Note there is no double-counting (i.e. the part of the business income the owner pays himself as wages is not counted in wage income). We use CPI to convert all incomes into 2000 dollars.

Schooling. The measure of schooling attainment is inconsistent across surveys. Prior to 1992, we have information regarding the completed years of schooling (higrade). This variable specifies whether each given year was attempted and completed. Beginning in 1992, CPS reports education according to the highest degree attained (educ99). Hence, for the survey years prior to 1992, we define high school graduates as those completing 12 years (higrade=150), college dropouts as those with less than four years of college (151,...,181), and college graduates are those with 16+ years of schooling (190 and above). For the 1992 surveys and all subsequent surveys, we define high school graduates as those with HS diploma or GED (educ99=10), college dropouts as those with ”some college no degree,” ”associate degree/occupational program,” ”associate degree/academic program” (11,12,13). College graduates are those with bachelors degree, masters degree, professional degree, doctorate degree (14,...,17) We restrict our sample to high school graduates, college dropouts and college graduates.

Age earnings profiles. To estimate the age-earnings profiles, we further restrict the sample to those earning above the $1000 threshold. For each schooling group, we regress
log earnings on the age quartic, unemployment, and cohort dummies (there are 86 cohorts, one is omitted). To complete the age log earnings profile for the 1960 cohort, we multiply the profile predicted by the age quartic by the participation profile. We complete the participation profile for the 1960 cohort (this cohort reached the maximum age of 47) by using the cross-sectional participation profile from the last survey.

To estimate the earnings profile for ages 65 and over, we need to construct their earnings. We restrict the CPS sample to white men of ages 65 and over (89 is the maximum age available). We do so by summing wages and business income (for the self employed as discussed above), Social Security income (incss) and veterans benefits (incvet). We do not include welfare income, as we did not include that in calculating the profile of working age men. We further include in the retirement income all other government, military, or company/union pensions, except for the private pension or annuity accounts such as IRA, KEOGH, 401K (This income is simply the return on investing previously earned income, so we need to avoid the double counting). Hence, we need detailed information on the nature of the composition of one’s retirement income. Because this information is available beginning in 1988, we restrict our attention to CPS years 1988-2010. Variables increti1 and increti1 refer to the first and second additional retirement income. Variables srcreti1 and srcreti2 specify the type of the retirement plan that is making these payments respectively. We incorporate these retirement incomes into the total retirement income except those in the retirement plan categories 5 and 6.

We employ the same restrictions on the CPS samples as described above, except we work with individuals born between 1910 and 1940. Because the observations are fewer for older people, we define cohorts according to the 5 year birth interval: those born in 1910-1914 comprise the first cohort and so on. To estimate the age-earnings profiles for 65+, we further restrict the sample to those reporting earned/pension income above the $1000 threshold. Then for each schooling group, we regress log earnings on the age quartic, unemployment, and cohort dummies (there are 6 cohorts, one is omitted). To complete the age log earnings profile, we multiply the profile predicted by the age quartic by the participation profile of the first cohort. (There are very few people below the threshold, around 2% in the entire sample) We complete the participation profile (ages 65-73) by using the participation profile for the 3rd cohort.
B Appendix: NLSY79 Data

To be written.