

The Return to College: Selection Bias and Dropout Risk*

Lutz Hendricks[†] Oksana Leukhina[‡]

February 10, 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 is 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.

We find 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 40% of the measured college wage premium.

JEL: E24, J24 (human capital), I21 (analysis of education).

Key words: Education. Ability. Skill premium.

*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 40% of the measured college wage premium is due to ability selection. This finding is similar to Hendricks and Schoellman (2011).
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.

At the start of age 1, all agents graduate from high school. They are endowed with assets k_1 , ability a , and a signal about ability m . 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 credits n . Once a student reaches n_{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 x , 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 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 $p_{a,i}$; $i = 1, \dots, N_a$. Ability is not known to the agent until he starts working.
2. a signal m that takes on value \hat{m}_j with probability $p_{m,i}(\hat{a}_j)$; $j = 1, \dots, N_m$.
3. 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 .
4. $n = 0$ completed college credits.

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

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

where $c_t = [\alpha c_{Ft}^\rho + (1 - \alpha)c_{Lt}^\rho]^{1/\rho}$ 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 τ is the age at which work starts, k_τ 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 $V_C(k, n, m, t)$. Section 2.6 describes how V_C is determined. An agent chooses to start college if $V_C(k_1, 0, m, 1) > V_W(k_1, 0, m, HS, 1)$.

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 $e^{a+\mu_s n} Y(s, \tau)$ denote the present value of lifetime earnings. The worker chooses market consumption to solve

$$V(k_\tau, a, n, s, \tau) = \max_{\{c_{Fi}\}} \sum_{i=\tau}^T \beta^{i-\tau} u(c[c_{Fi}, \hat{c}_L]) \quad (2)$$

subject to the budget constraint

$$e^{a+\mu_s n} Y(s, \tau) + Rk_\tau = \sum_{i=\tau}^T c_{Fi} R^{\tau-i} \quad (3)$$

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

$$V_W(k_\tau, n, m, s, \tau) = E_a \{V(k_\tau, a, n, s, \tau)|n, m, \tau\} \quad (4)$$

$$= \sum_{i=1}^{N_a} V(k_\tau, \hat{a}_i, s, \tau) Pr(\hat{a}_i|n, m, \tau) \quad (5)$$

where $Pr(\hat{a}_i|n, m, \tau)$ 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 , $V_C(k_t, n_t, m, t)$, 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 x . She chooses consumption c_{Ft} so that next period's assets are determined by the budget constraint $k_{t+1} = Rk_t - c_{Ft} - x$. She receives a fixed amount of non-market consumption, \bar{c}_L , for free. Borrowing is constrained by $k_{t+1} \geq k_{min,t+1}$. If the student lacks the funds to pay for tuition, so that $Rk_t < x + k_{min,t+1}$, we set $V_C(k_t, n_t, m, t) = -\infty$.

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 and decides whether to work or study in period $t + 1$.

The option of studying is not available if

1. $n_{t+1} \geq n_{grad}$: 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. $n_{t+1} < n_{grad}$ 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: $V_W(k_{t+1}, n_{t+1}, m, CD, t + 1) > V_C(k_{t+1}, n_{t+1}, m, t + 1)$. The Bellman equation is therefore given by

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

where $V_{EC}(k, n, m, t) = V_W(k, n, m, CG, t)$ if the student graduates from college, $V_{EC}(k, n, m, t) = V_W(k, n, m, CD, t)$ if the student is forced to drop out of college, and $V_{EC}(k, n, m, t) = \max\{V_C(k, n, m, t), V_W(k, n, m, CD, t)\}$ 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

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

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

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

$Pr(n_t|\hat{a}_i, t)$ 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 $Pr(n_t, m|\hat{a}_i, t) = Pr(m|\hat{a}_i) Pr(n_t|\hat{a}_i, t)$.

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 ASWAB 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.
3. The non-market good is only consumed while in college: $\hat{c}_L = 0$.

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. The preference weight on the market good is set to $\alpha = 0.5$. (We plan to use expenditure shares to pin this down.)
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 $n_{grad} = 8$. 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: We allow students to borrow \$10,000 in each period. (We plan to incorporate borrowing opportunities in more detail later.)
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 σ_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 σ_m .
3. The standard deviation of noise in AFQT scores: σ_{AFQT} .
4. The lifetime earnings functions $Y(s, \tau)$.
5. The probability of passing a credit is given by $e^{-\gamma a}$. γ is calibrated.
6. The consumption value of college, \bar{c}_L .

Table 1: AFQT scores and schooling. NLSY79 data.

AFQT quartile	School Attainment			
	<HS	HS	SC	C+
1	0.861	0.417	0.184	0.010
2	0.118	0.338	0.317	0.105
3	0.021	0.194	0.309	0.294
4	0.000	0.051	0.190	0.591
Fraction	0.060	0.323	0.328	0.289
N	163	642	644	493

Source: Hendricks and Schoellman (2011)

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 NLSY79 sample, which covers the cohorts born between 1957 and 1964.

The first target is 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).

The second data moments is the joint distribution of AFQT scores and schooling. For each AFQT quartile, Table 1 reports the fraction of high school graduates who complete a given level of schooling.

The third data moment is the fraction of persons who attain each school level.

Lifetime earnings. We estimate lifetime earnings, conditional on schooling, using a sample of white men.

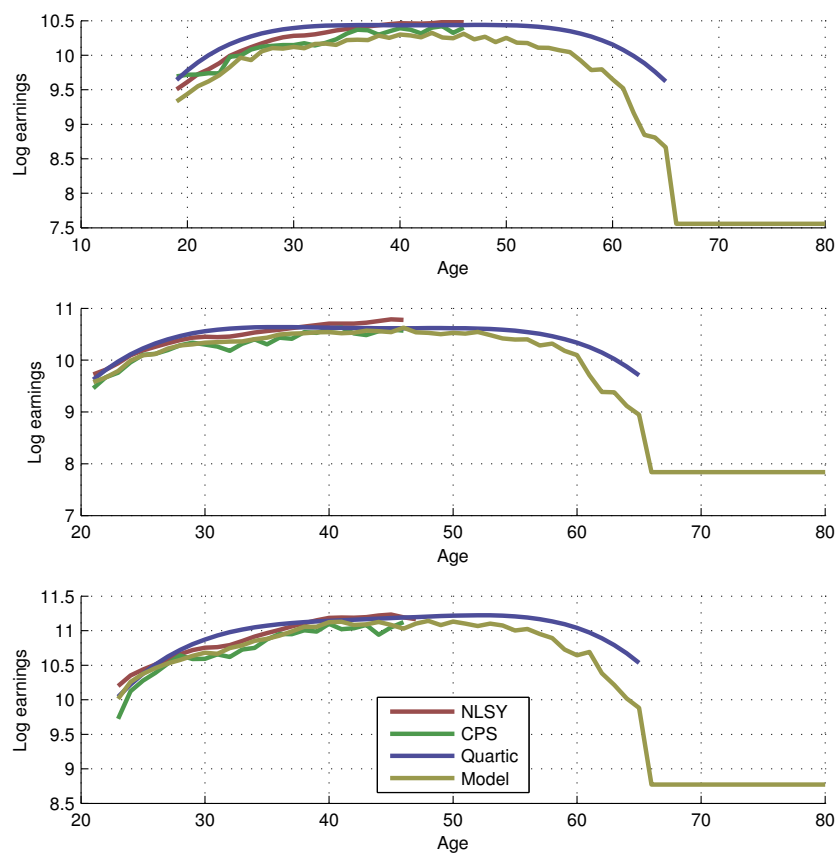
We keep all men with known schooling and earnings in each year. We regress log earnings (above a threshold of \$1,000 in year 2000 prices) on age quartic, cohort dummies, and the unemployment rate. We keep the implied age earnings profile for the 1960 birth cohort, with a starting age at $1960 + \tau_s$, where τ_s is a typical age at which a person with schooling level s starts to work. We set τ_s to 19 for high school graduates, 21 for college dropouts, and 23 to for college graduates. We then calculate the fraction of men receiving earnings above the threshold for each age and multiply the estimated age-earnings profiles by this fraction.

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 and after retirement. We use CPS data to estimate a cohort invariant age earnings profile for each school group. 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.

We use a similar regression to estimate the age retirement income profile. [Details to be written.] The results are shown in Figure 1. 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 2.

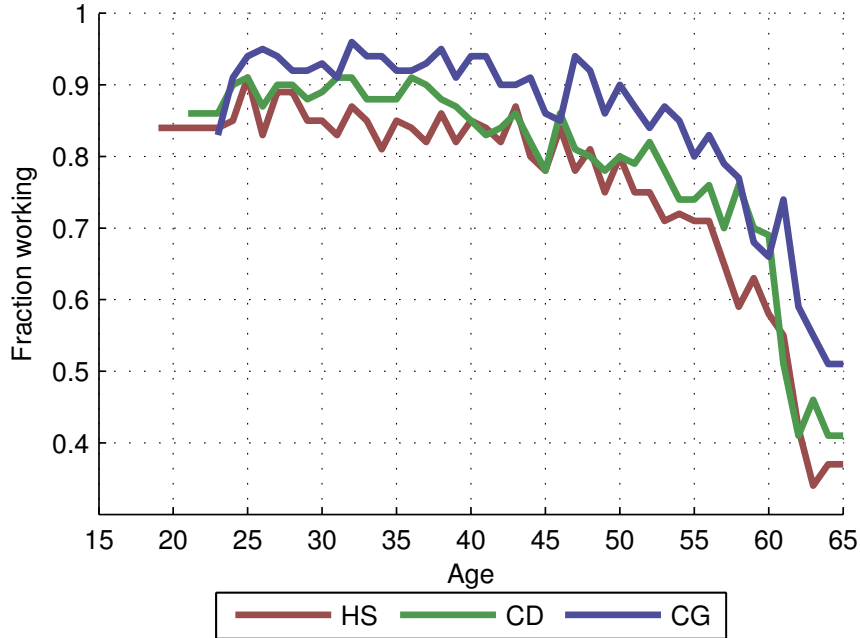
Mapping into model objects: We assume that the experience earnings profile is independent of the age at which work starts. A worker who graduates from college one year later effectively shifts the age earnings profile back in time by one year and thus loses the last year of earnings. The data identify $Y(s, \tau)$ up to the scale factor $E(e^{a+\mu_s n}|s)$. The calibration algorithm searches over these scale factors. For each guess, the household problems are solved. The algorithm iterates until the simulated data match the estimated lifetime earnings levels.

Figure 1: Age earnings profiles



Note: Annual earnings in multiples of \$10,000 in year 2000 prices.

Figure 2: Fraction working by age



4 Results

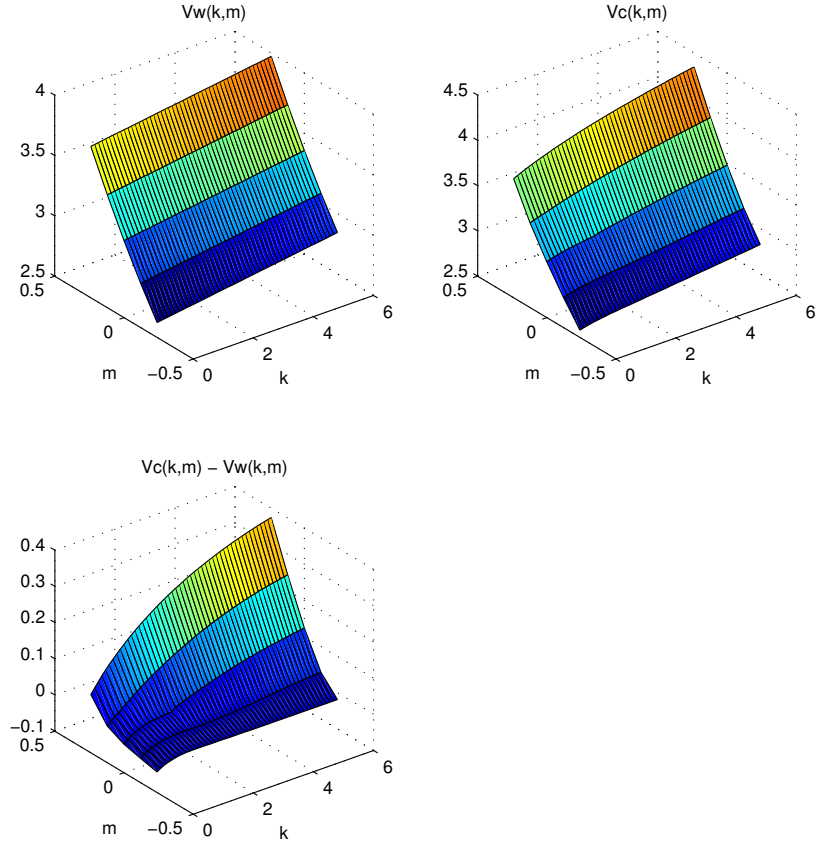
4.1 Selection Bias

Our first question is: What fraction of the college wage premium represents selection bias?

In our 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 $E(a|CG) - E(a|HS) = 0.25$. To put this into perspective, Hendricks and Schoellman (2011) estimate a college wage premium of 0.52 in the year 2000. Our model implies that almost half of the observed college wage premium reflects selection bias. Even though our model differs greatly from theirs, our finding is very close to that of Hendricks and Schoellman (2011).

To illustrate this result, Figure 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

Figure 3: College Choice At Age 1



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.

4.2 Return to College and Ability

We next examine the rate of return to college for students of different ability levels. Table 2 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 (3% 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.

Table 2: Outcomes By Ability

	Ability				
	1	2	3	4	5
Fraction (pct)	12.6	22.6	29.6	22.6	12.6
Dropout probability	97.2	85.7	60.0	30.7	10.3
HS					
Fraction	23.9	32.2	27.5	12.6	3.8
Lifetime earnings	73.3	82.4	92.7	104.4	117.6
Mean signal	1.2	1.3	1.5	1.7	1.9
Age at start of work	1.0	1.0	1.0	1.0	1.0
CD					
Fraction	13.4	30.0	36.2	16.9	3.5
Lifetime earnings	77.4	89.2	102.2	115.5	127.9
Mean signal	3.0	3.4	3.6	3.9	4.0
Age at start of work	3.0	3.4	3.7	3.8	3.5
Completed credits	1.3	1.9	2.4	2.5	2.1
CG					
Fraction	0.4	5.1	24.6	38.9	31.0
Lifetime earnings	112.1	125.8	142.0	160.3	181.4
Mean signal	3.2	3.5	3.8	4.1	4.4
Age at start of work	5.7	5.6	5.5	5.3	5.0

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.

5 Conclusion

To be written.

References

- BELLEY, P., AND L. LOCHNER (2007): “The Changing Role of Family Income and Ability in Determining Educational Achievement,” *Journal of Human Capital*, 1(1), 37–89.
- BOWEN, W. G., M. M. CHINGOS, AND M. S. MCPHERSON (2009): *Crossing the Finish Line: Completing College at America’s Public Universities*. Princeton University Press.
- CUNHA, F., J. J. HECKMAN, AND S. NAVARRO (2005): “Separating uncertainty from heterogeneity in life cycle earnings,” *Oxford Economic Papers*, 57, 191–261.
- HECKMAN, J. J., L. J. LOCHNER, AND P. E. TODD (2006): “Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond,” in *Handbook of the Economics of Education*, ed. by E. Hanushek, and F. Welch, vol. 1, chap. 7, pp. 307–458. Elsevier.
- (2008): “Earnings Functions and Rates of Return,” *Journal of Human Capital*, 2(1), 1–31.
- HENDRICKS, L., AND T. SCHOELLMAN (2011): “Student Abilities During the Expansion of U.S. Education, 1950-2000,” Mimeo. University of North Carolina.
- MANSKI, C. F. (1989): “Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon,” *Economics of Education Review*, 8(4), 305–312.
- NAVARRO, S. (2008): “Using Observed Choices to Infer Agent’s Information: Reconsidering the Importance of Borrowing Constraints, Uncertainty and Preferences in College Attendance,” Mimeo: University of Wisconsin.
- NLS USER SERVICES (1992): “NLS79 Profiles of American Youth. Addendum to Attachment 106,” http://www.nlsinfo.org/ordering/display_db.php3.

A Appendix: CPS Data

To be written.

B Appendix: NLSY79 Data

To be written.