Hendricks, Lutz; Leukhina, Oksana

**Working Paper**

# The Return to College: Selection Bias and Dropout Risk

CESifo Working Paper, No. 4733

**Provided in Cooperation with:**
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Hendricks, Lutz; Leukhina, Oksana (2014) : The Return to College: Selection Bias and Dropout Risk, CESifo Working Paper, No. 4733, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/96909

The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers
www.cesifo.org/wp

# The Return to College: Selection Bias and Dropout Risk

Lutz Hendricks
Oksana Leukhina

CESIFO WORKING PAPER NO. 4733
CATEGORY 5: ECONOMICS OF EDUCATION
APRIL 2014



CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 4733*

# The Return to College: Selection Bias and Dropout Risk

## Abstract

This paper estimates the effect of graduating from college on lifetime earnings. Motivated by the fact that nearly half of all college students fail to earn a bachelor's degree, we study a model of risky college completion. The central idea is that students drop out of college mainly because they fail to complete the requirements for earning a degree. This introduces two levels of ability selection that reinforce each other. (i) In college, low ability students typically do not succeed academically and drop out. (ii) At the college entry stage, their poor graduation prospects deter low ability students from even attempting college. Taken together, the two levels of selection generate a large ability gap between college graduates and high school graduates. We calibrate the model to data for men born around 1960 and find that ability selection accounts for nearly half of the college lifetime earnings premium.

JEL-Code: E240, J240, I210.

Keywords: education, college dropout risk.

*Lutz Hendricks*
*Department of Economics*
*University of North Carolina*
*USA - 27599-3305 Chapel Hill NC*
*lutz@lhendricks.org*

*Oksana Leukhina*
*Department of Economics*
*University of Washington*
*USA - Seattle, WA 98195-3330*
*oml@u.washington.edu*

October 1, 2013
For helpful comments we thank V. V. Chari, Patrick Kehoe, Rodolfo Manuelli, José-Victor Rios-Rull, Christoph Winter as well as seminar participants at Indiana University, University of Minnesota, University of Washington, Ohio State, the Federal Reserve Bank of Minneapolis, Simon Fraser University, Tel Aviv University, Washington State University, York University, the 2011 Midwest Macro Meetings, the 2011 SED Meetings, the 2011 Cologne Macro Workshop, the 2011 CASEE Human Capital Conference, the 2011 Barcelona Growth Workshop, the 2011 USC-Marshall mini macro labor conference, and the 2013 QSPS Summer Workshop.

# 1 Introduction

A large literature has investigated the causal effect of schooling on earnings.[1] In U.S. data, college graduates earn substantially more than high school graduates. However, part of this differential may be due to selection as students with superior abilities or preparation are more likely to graduate from college. While various approaches have been proposed to control for selection, no consensus has been reached about its importance.

In this paper, we offer a new approach which emphasizes the importance of *college completion risk*. The idea is that not all students are able to complete the coursework required for graduation, so that college completion is uncertain. Selection therefore occurs at two levels: (i) Recognizing that they are unlikely to graduate, students of lower abilities are less likely to attempt college. (ii) Those low ability students who attempt college likely fail to graduate. We show that the interaction between these two levels of selection generates large ability gaps between high school graduates and college graduates and accounts for our main finding: roughly half of the lifetime earnings gap between college graduates and high school graduates is due to selection.

Our emphasis on college completion risk is motivated by the following observations.[2]

1. Over their lifetimes, college graduates earn about $400,000 more than high school graduates, suggesting that the return to completing college may be large.

2. College dropouts earn only about $70,000 more than high school graduates, suggesting that the return to college accrues mainly to those who attain a degree.[3]

3. Even so, nearly half of those who start college fail to attain a degree, suggesting that dropout risk may be important for understanding the incentives for attending college (Bound, Lovenheim, and Turner, 2010).

4. College graduates have substantially higher cognitive test scores than do high school graduates, suggesting that ability differences may be important for college completion and for the college earnings premium.

[1] For a recent survey, see Oreopoulos and Petronijevic (2013).

[2] The observations are derived from NSLY79 data that are described in Section 3 and Appendix B. All dollar figures are in year 2000 prices.

[3] For evidence on sheepskin effects, see Jaeger and Page (1996).

We interpret dropping out of college as resulting mainly from a student's inability to complete the course work required for graduation. Based on High School & Beyond (HS&B) transcript data, we show that college dropouts earn one-third fewer credits in each year of college compared with college graduates (see Section 3). After four years in college, dropouts are still far from having completed the roughly 120 credits required for graduation. College dropouts also earn substantially poorer grades than do college graduates. How rapidly a student accumulates college credits plays a central role for dropout decisions in our model.

Our interpretation of dropping out as a failure is supported by other evidence. Pryor, Hurtado, Saenz, and Santos (2007) report that 98% of all freshmen entering colleges or universities plan to earn at least a bachelor's degree. Of course, far fewer attain this goal. Stinebrickner and Stinebrickner (2003) emphasize the importance of college grades for dropout decisions. Bound and Turner (2011) emphasize the role of college preparation for college completion.

Our notion of college completion risk differs from that of Keane and Wolpin (1997) and others following their approach.[4] In these models, all students can attain college degrees in a reasonable amount of time. Some students are exposed to shocks, such as wage offers, as they progress through college and therefore *choose* to forego the college wage premium. In our model, dropping out of college is largely the result of poor "grades" which signal low abilities and convince the students that completing college would be prohibitively expensive. This induces a stronger correlation between dropout behavior and abilities than in Keane and Wolpin style models.

**Approach:** Our paper unfolds as follows. In Section 2, we propose a model of college choice with ability heterogeneity and dropout risk. At high school graduation, agents are endowed with abilities that affect their chances of attaining college degrees and also their labor earnings. Following Manski and Wise (1983) and Manski (1989), we assume that students only observe noisy signals of their abilities. While in college, students take courses which add to their human capital. The probability of successfully completing a course depends on the student's ability (as in Garriga and Keightley 2007). As they progress through college, students gradually learn their abilities. Low ability students realize that graduat-

[4] Examples include Belzil and Hansen (2002), Keane and Wolpin (2001), Stange (2012) and Arcidiacono, Aucejo, Maurel, and Ransom (2012).

ing from college would take a long time and choose to drop out. After completing their education, individuals work until retirement. Their earnings depend on their educational attainment and ability.

In Section 3, we calibrate the model for men born around 1960. Our main data sources are the NLSY79, which provides us with schooling, cognitive test scores, and partial earnings histories, and High School & Beyond, from which we take transcript and financial variables.

**Findings:** We report our findings in Section 4. Our model implies that ability selection is important. We measure its contribution as the fraction of the lifetime earnings gap between college graduates and high school graduates that would remain, if both groups worked as high school graduates. In the main specification, this fraction is 48%.

To understand the intuition behind this result, it helps to contrast our model with the Roy model commonly used in the literature (e.g., Heckman, Lochner, and Taber 1998; Cunha, Heckman, and Navarro 2005). In the Roy model, students of any ability can graduate from college with certainty. In the simplest specification, the percentage wage gain from attending college is the same for all persons. The reason why high ability students are more likely to attend college is then that the absolute gain in lifetime earnings is increasing in ability. Therefore, any fixed costs associated with attending college (tuition or psychic costs) are less important for high ability students.

This type of selection is present in our model. However, with dropout risk, selection occurs at two additional levels. At college entry, low ability students are deterred by their bleak graduation prospects. While in college, low ability students fail to accumulate the number of credits needed to graduate and are forced to drop out. This implies a large ability gap between college graduates and high school graduates and therefore a large contribution of selection to the college premium.

We highlight a number of additional findings:

1. College graduation prospects and the earnings gains associated with entering college vary strongly with ability. The probability of graduating from college varies from 20% in the lowest ability decile to 89% in the highest. The inability to attain a degree is the main friction that prevents low ability students from entering college.

2. As a result, low ability students mainly view college as a consumption good. Their college entry decisions are quite sensitive to the direct costs of college. This feature al-

lows our model to account for the large effects of tuition changes on college enrollment estimated in the literature (see Section 5.1).

3. By contrast, high ability students view college mainly as an investment. Since they expect to graduate, their college entry decisions respond strongly to the wages earned by college graduates, but not to tuition changes.

4. Even though few students in our model, and in NLSY79 data, are close to their borrowing limits, relaxing these limits has important effects on college enrollment (Section 5.2). The reason is the option value of entering college. Some students lack the financial resources required for graduating from college. Since a large part of the return to college depends on graduation, these students either do not attempt college, or they enter college planning to drop out after a few semesters. Increased borrowing opportunities allow these students to attend college for several years without suffering very low consumption. If they earn fewer college credits than expected, they update their beliefs about their graduation prospects and drop out, incurring at most a small financial loss.

5. Dual enrollment programs allow high school students to take college level courses in order to provide them with better information about their college aptitudes. Our model implies that such programs have little effect on college entry decisions.

## 1.1 Related Literature

This paper relates to a vast literature that estimates returns to schooling. One strand of this literature uses econometric approaches, such as instrumental variables, to control for selection bias in wage regressions.[5] These efforts abstract from degrees and treat schooling as a continuous variable and are therefore silent about the college premium and completion risk.

A more recent literature has developed structural discrete choice models of schooling decisions. A large share of this is based on Roy models which abstract from college completion

[5] A seminal contribution is Willis and Rosen (1979). Card (1999) surveys this literature and discusses how its findings may be interpreted.

risk.[6]

Models with college completion risk have, for the most part, abstracted from heterogeneity in abilities that directly affect earnings. Examples include Altonji (1993), Caucutt and Kumar (2003), Akyol and Athreya (2005), Garriga and Keightley (2007), Chatterjee and Ionescu (2010), and Stange (2012).[7] These models cannot address the question how ability selection affects measured college wage premiums.

A number of recent papers feature both ability heterogeneity and college completion risk. As discussed earlier, we depart from models that build on Keane and Wolpin (1997) in the way we model academic achievement in college. We follow Garriga and Keightley (2007) in assuming that students need to earn college credits in order to graduate. A similar approach is taken by Eckstein and Wolpin (1999) who study high school dropouts, and by Trachter (2012) who studies the role of 2-year colleges as stepping stones towards a bachelor's degree. Relative to Trachter's study, our model features a richer specification of unobserved heterogeneity, which is important for estimating ability selection. We also calibrate our model using a richer set of empirical observations, in particular regarding the relationship between measured abilities, college outcomes, and earnings. In work in progress, Heckman and Urzua (2008) study a model of risky college completion where students learn about their abilities and schooling preferences.

# 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 ($t = 1$), through college (if chosen), work, and retirement. When entering the model, each high school graduate goes through the following steps:

1. He draws a type $j \in \{1, ..., J\}$ which determines his initial assets $k_1 = \hat{k}_j$, his ability signal $m = \hat{m}_j$, and a net price of attending college $q = \hat{q}_j$.

[6] Examples include Heckman, Lochner, and Taber (1998), Cunha, Heckman, and Navarro (2005), and Navarro (2008).

[7] It is possible to interpret some of the psychic costs in Stange's model as variation in returns to college. However, his model cannot quantify the contribution of ability selection to measured wage premiums.

2. He draws an ability $a$ that is not observed until the agent starts working. More able agents are more likely to graduate from college and earn higher wages in the labor market.

3. The student commits to a consumption level $c_{i,j}$ that remains fixed throughout college.

4. The student chooses between attempting college or working as a high school graduate ($s = HS$).

An agent who studies in period $t$ faces the following choices:

1. He consumes $c_{i,j}$, pays the college cost $\hat{q}_j$, and saves (or borrows) $k_{t+1} = Rk_t - c_{i,j} - \hat{q}_j$.

2. He attempts $n_c$ college credits and succeeds in a random subset, which yields $n_{t+1}$. More able students accumulate credits faster, as in Garriga and Keightley (2007).

3. Based on the information contained in the number of credits earned, the student updates his beliefs about $a$.

4. If the student has earned enough credits for graduation ($n_{t+1} \geq n_{grad}$), he must work in $t + 1$ as a college graduate ($s = CG$). If the student has exhausted the maximum number of years of study ($t = T_c$), he must work in $t+1$ as a college dropout ($s = CD$). Otherwise, he chooses between staying in college and working in $t + 1$ as a college dropout.

An agent who enters the labor market in period $t$ learns his ability $a$. He then chooses 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. Agents are not allowed to return to school after they start working.

The details are described next. We motivate our assumptions in Section 2.6.

## 2.2 Endowments

Agents enter the model at high school graduation (age $t = 1$) and live until age $T$. At age 1, a person is endowed with

1. $n_1 = 0$ completed college credits;

2. learning ability $a \in \{\hat{a}_1, ..., \hat{a}_{N_a}\}$ with $\hat{a}_1 = 0$ and $\hat{a}_{i+1} > \hat{a}_i$;

3. type $j \in \{1, ..., J\}$.

$a$ determines the person's productivity in school and at work. Normalizing the lowest ability level to zero simplifies the notation without loss of generality. Abilities are not observed by the agents until they start working. A person of type $j$ is endowed with the following:

1. A noisy signal $\hat{m}_j$ of the individual's true ability level $a$.

2. A net price of attending college $\hat{q}_j$. We think of this as capturing tuition, scholarships, grants, and other costs or payoffs associated with attending college.

3. Initial assets $\hat{k}_j \geq 0$. We think of these as capturing financial assets and parental transfers that are received regardless of whether the person attends college.

The distribution of endowments is specified in Section 3.

## 2.3 Work

We now describe the solution of the household problem, starting with the last phase of the household's life, work. Consider a person who starts working at age $\tau$ with assets $k_\tau$, ability $a$, $n_\tau$ college credits, and schooling level $s \in \{HS, CD, CG\}$. The worker chooses a consumption path $\{c_t\}$ for the remaining periods of his life $(t = \tau, ..., T)$ to solve

$$V(k_\tau, n_\tau, a, s, \tau) = \max_{\{c_t\}} \sum_{t=\tau}^{T} \beta^{t-\tau} u(c_t) + U_s \tag{1}$$

subject to the budget constraint

$$\exp(\phi_s a + \mu n_\tau + y_s) + R k_\tau = \sum_{t=\tau}^{T} c_t R^{\tau - t}. \tag{2}$$

Workers derive period utility $u(c_t) = \ln(c_t)$ from consumption, discounted at $\beta > 0$. $U_s$ captures the utility derived from job characteristics associated with school level $s$ that is common to all agents. The budget constraint equates the present value of consumption spending to lifetime earnings, $\exp(\phi_s a + \mu n_\tau + y_s)$, plus the value of assets owned at age $\tau$. $R$ is the gross interest rate. $y_s$ and $\phi_s > 0$ are schooling-specific constants.

Lifetime earnings are a function of ability $a$, schooling $s$ and accumulated credits $n_\tau$. A worker with ability $a = \hat{a}_1 = 0$ and no completed college credits earns $\exp(y_s)$. Each completed college credit increases lifetime earnings by $\mu > 0$ log points. This may reflect human capital accumulation. We impose $y_{CD} = y_{HS}$ and $\phi_{CD} = \phi_{HS}$ to ensure that attending college for a single period without earning credits does not increase earnings simply by placing a "college" label on the worker.

Our formulation allows for the effect of ability on lifetime earnings ($\phi_s$) to depend on schooling. If $\phi_{CG} > \phi_{HS}$, ability and schooling are complements: A high ability person gains more from obtaining a college degree than a low ability person. One possible reason is that high ability persons accumulate more human capital in college or on the job, as suggested by Ben-Porath (1967).[8]

Even though $y_s$ does not depend on $\tau$, staying in school longer reduces the present value of lifetime earnings by delaying entry into the labor market. Note that all high school graduates share $\tau = 1$ and $n_\tau = 0$, but there is variation in both $\tau$ and $n_\tau$ among college dropouts and college graduates.

Before the start of work, individuals are uncertain about their abilities. Expected utility is then given by

$$V_W(k_\tau, n_\tau, j, s, \tau) = \sum_{i=1}^{N_a} \Pr(\hat{a}_i | n_\tau, j, \tau) V(k_\tau, n_\tau, \hat{a}_i, s, \tau). \tag{3}$$

Our model of credit accumulation implies that the vector $(n_\tau, j, \tau)$ is a sufficient statistic for the worker's beliefs about his ability, $\Pr(\hat{a}_i | n_\tau, j, \tau)$, which implies that $(k_\tau, n_\tau, j, s, \tau)$ is the correct state vector.

## 2.4 College

Consider an individual of type $j$ who has decided to study in period $t$. He enters the period with assets $k_t$ and $n_t$ college credits. In each period, the student attempts $n_c$ credits and completes each with probability $\Pr_c(a)$ given by the logistic function

$$\Pr{}_c(a) = \gamma_{min} + \frac{\gamma_{max} - \gamma_{min}}{1 + \gamma_1 e^{-\gamma_2 a}}. \tag{4}$$

[8] High ability students also tend to favor more lucrative majors while in college (Arcidiacono, 2004).

We assume $\gamma_{max} = 0.98 > \gamma_{min}$ and $\gamma_1, \gamma_2 > 0$, so that the probability of earning credits increases with the student's ability. Based on the number of completed credits, $n_{t+1}$, the student updates his beliefs about $a$. Since $n_t$ is drawn from the Binomial distribution, it is a sufficient statistic for the student's entire history of course outcomes. It follows that his beliefs about $a$ at the end of period $t$ are completely determined by $n_t$ and $j$.

We assume that students commit to a constant consumption level for their entire college careers. We postpone the discussion of consumption choice for now and take consumption as fixed at $c_{i,j}$. While in college, assets evolve according to the budget constraint

$$k_{t+1} = Rk_t - c_{i,j} - \hat{q}_j. \tag{5}$$

The student earns gross interest $R$ on his assets (or debts), pays tuition $\hat{q}_j$ and consumes $c_{i,j}$. Since all students of type $(i,j)$ have the same initial assets $\hat{k}_j$ and expenditures $c_{i,j} + \hat{q}_j$, they also share the same asset levels in subsequent periods, which we denote by $k_{i,j,t}$. The flow utility of consumption is given by $u(c_{i,j})$.

The value of being in college at age $t$ is then given by

$$V_C(n,i,j,t) = u(c_{i,j}) + \beta \sum_{n'} \Pr(n'|n,j,t)\, V_{EC}(n',i,j,t+1), \tag{6}$$

where $\Pr(n'|n,j,t)$ denotes the probability of having earned $n'$ credits at the end of period $t$. This is computed using Bayes' rule from the students' beliefs about $a$. Since assets (or debts) are a function of $(i,j,t)$, they are not state variables. $V_{EC}$ denotes the value of entering period $t$ before the decision whether to work or study has been made. It is determined by the discrete choice problem

$$V_{EC}(n,i,j,t) = \max\{V_C(n,i,j,t) + \pi p_c, V_W(k_{i,j,t},n,j,s(n),t) + \pi p_w\} - \pi\bar{\gamma}, \tag{7}$$

where $p_c$ and $p_w$ are independent draws from standard type I extreme value distributions with scale parameter $\pi > 0$. $\bar{\gamma}$ is the Euler–Mascheroni constant, which is the mean of the standard type I extreme value distribution. Subtracting $\pi\bar{\gamma}$ effectively sets the means of $p_c$ and $p_w$ to $-\bar{\gamma}$, which simplifies the value functions. $s(n)$ denotes the schooling level associated with $n$ college credits (CG if $n \geq n_{grad}$ and CD otherwise).

The implied choice probabilities are given by

$$\Pr(\text{study}|n,i,j,t) = \frac{\exp(V_C(n,i,j,t)/\pi)}{\exp(V_C(n,i,j,t)/\pi) + \exp(V_W(k_{i,j,t},n,j,s(n),t)/\pi)}, \tag{8}$$

and the associated value function is[9]

$$V_{EC}(n,i,j,t) = \pi \ln\left(\exp\left(V_C\left(n,i,j,t\right)/\pi\right) + \exp\left(V_W\left(k_{i,j,t},n,j,s\left(n\right),t\right)/\pi\right)\right). \quad (9)$$

In evaluating $V_{EC}$ three cases can arise:

1. If $n \geq n_{grad}$, then $s\left(n\right) = CG$ and $V_C = -\infty$: the agent graduates from college with continuation value $V_W\left(k_{i,j,t},n,j,CG,t\right)$.

2. If $t = T_c$ and $n < n_{grad}$, then $s\left(n\right) = CD$ and $V_C = -\infty$: the student has exhausted the permitted time in college and must drop out with continuation value $V_W\left(k_{i,j,t},n,j,CD,t\right)$.

3. Otherwise the agent chooses between working as a college dropout with $s\left(n\right) = CD$ and studying next period.

## 2.5 Choices at High School Graduation

At high school graduation ($t = 1$), each student makes 2 choices: (i) whether to attempt college or work as a high school graduate and (ii) how much to consume in college.

**Consumption choice.** Before deciding whether to enter college, the student commits to a consumption level that remains fixed throughout college. Consumption is chosen from a discrete set of values $c_{i,j}$ that are indexed by $i = 1, ..., N_c$ and vary by type $j$. Consumption choice maximizes lifetime utility subject to $N_c$ independent preference shocks $p_i$, drawn from a standard type I extreme value distribution with scale parameter $\pi_c > 0$:

$$i = \arg\max_{\hat{i}} \left\{ V_C\left(0,\hat{i},j,1\right) + \pi_c(p_{\hat{i}} - \bar{\gamma}) \right\}. \quad (10)$$

Subtracting $\bar{\gamma}$ effectively sets the means of the preference shocks to $-\bar{\gamma}$, which simplifies the value functions. The implied choice probabilities are given by

$$\Pr\left(i|j\right) = \frac{\exp\left(V_C(0,i,j,1)/\pi_c\right)}{\sum_{\hat{i}=1}^{N_c} \exp\left(V_C(0,\hat{i},j,1)/\pi_c\right)}. \quad (11)$$

The main purpose of the preference shocks is to ensure that the objective function minimized by the calibration algorithm is smooth in the parameter values. The assumption that

[9] See Rust (1987) and Arcidiacono and Ellickson (2011).

consumption remains fixed while a student attends college drastically simplifies the model computation.[10] In the calibration, we set $N_c = T_c$ and fix $c_{i,j}$ such that a type $(i, j)$ student exactly exhausts his borrowing limits after $i$ periods in college. The motivation is that marginal utility is discontinuous at these points, so that students would choose these consumption levels with positive probability, even if consumption were continuous.

**College entry decision.** The college/work decision is made after consumption has been chosen. The agent solves

$$\max \left\{ V_C(0, i, j, 1) + \pi p_c,\ V_W(\hat{k}_j, 0, j, HS, 1) + \pi p_w \right\} - \pi\bar{\gamma}, \tag{12}$$

where $p_c$ and $p_w$ are two independent draws from a standard type I extreme value distribution with scale parameter $\pi > 0$. The probability of starting college is then given by

$$\Pr\left(\text{college}|i, j\right) = \frac{\exp(V_C(0, i, j, 1)/\pi)}{\exp(V_C(0, i, j, 1)/\pi) + \exp\left(V_W(\hat{k}_j, 0, j, HS, 1)/\pi\right)}. \tag{13}$$

## 2.6 Discussion of Model Assumptions

Our model assumptions attempt to capture key features that may be important for the main issues we wish to investigate: ability selection and the risk of dropping out of college.

We model *dropping out* of college as a *choice*. Similar to Garriga and Keightley (2007), students drop out if they receive poor "grades," which imply that graduating from college would take longer than previously expected. While we do not model this explicitly, we can think of the probability of completing a credit as a function of study effort, which is maximized out in the specification of $\Pr_c(a)$. Relative to the simpler alternative where dropping out is a shock (as in Caucutt and Kumar 2003 or Akyol and Athreya 2005), our approach has the benefit that we can use data on the characteristics and the timing of dropouts to help identify the frictions that prevent students from earning a degree, such as borrowing constraints or uncertainty about students' learning abilities. Relative to the literature that treats dropping out as an ex ante decision, we capture how the risk of failure affects the ex ante rate of return of college for students of various characteristics.

[10] If consumption were chosen in each period, the household problem would gain a state variable ($k_t$). Due to the borrowing constraints, the marginal value of $k_{t+1}$ is not continuous, so that first-order conditions could not be used to find the optimal consumption level.

Manski (1989) argues that *learning about ability* may explain why many students drop out of college. Stinebrickner and Stinebrickner (2012) present survey evidence suggesting that learning about ability is important for college dropout decisions at Berea College. The evidence presented by Arcidiacono, Bayer, and Hizmo (2008) suggests that college histories reveal individual abilities to the labor market. We wish to investigate the quantitative importance of this explanation. We therefore allow for the possibility that students observe only a noisy signal of their abilities.

We incorporate *heterogeneity in financial assets* and in the net cost of attending college to capture the role of borrowing constraints for college selection. Whether borrowing constraints are important remains controversial in the literature (see Cameron and Taber 2004, Belley and Lochner 2007, among others). In our model, the vast majority of students have access to sufficient funds to pay for college tuition. However, some are subject to soft borrowing constraints which limit the amount of consumption they can afford in college.

The work-study decisions of model agents are subject to *preference shocks* which are similar to the "psychic costs" commonly found in models of school choice (see Heckman, Lochner, and Todd 2006 for a discussion). The main purpose of the preference shock affecting the college entry decision is to regulate the association between agents' types and school choices. Without preference shocks, school sorting would be perfect in the sense that all agents of a given type $j$ would make the same college entry decision. This would bias our results in favor of large ability selection (see Hendricks and Schoellman 2011). The preference shocks affecting the college dropout decision mainly improve the model's ability to account for the timing of dropout decisions and for the dropout rates of high ability students. In Appendix D we show that our main result is robust against variation in the dispersion of the preference shocks.

# 3 Calibration

We calibrate the model parameters to data match moments for men born around 1960. The model period is one year. Our main data sources are the National Longitudinal Surveys (NLSY79) and High School & Beyond (HS&B).

The NLSY79 is a representative, ongoing sample of persons born between 1957 and 1964 (Bureau of Labor Statistics; US Department of Labor, 2002). We collect education, earnings

and cognitive test scores for all men. We include members of the supplemental samples, but use weights to offset the oversampling of minorities. We use data from the Current Population Surveys (King, Ruggles, Alexander, Flood, Genadek, Schroeder, Trampe, and Vick, 2010) to impute the earnings of older workers. Appendices A and B provide additional details.

HS&B is published by the National Center for Educational Statistics (NCES). It covers 1980 high school sophomores. Participants were interviewed bi-annually until 1986. In 1992, postsecondary transcripts from all institutions attended since high school graduation were collected. We retain all men who report sufficient information to determine when they attended college and whether a degree was earned. HS&B also contains information on college tuition, financial resources, parental transfers, and student debt. Appendix C provides additional details.

## 3.1 Distributional Assumptions

Our distributional assumptions allow us to model substantial heterogeneity in assets, ability signals, and college costs in a parsimonious way. We set the number of types to $J = 120$. Each type has mass $1/J$. We assume that the $\left(\ln\left(\hat{k}_j\right),\ \hat{q}_j,\ \hat{m}_j\right)$ endowments of the $J$ types are drawn from a joint Normal distribution. The marginal distributions are: $\hat{k}_j \sim N\left(\mu_k, \sigma_k^2\right)$, $\hat{q}_j \sim N\left(\mu_q, \sigma_q^2\right)$, and $m \sim N(0,1)$. The log-Normal distribution of $\hat{k}_j$ enables the model to capture the fact that its empirical counterpart features a decreasing density with a large mass near zero.

We implement this by drawing three independent standard Normal random vectors of length $J$: $\varepsilon_k$, $\varepsilon_q$, and $\varepsilon_m$. Next, we set $\hat{k}_j = \mu_k + \sigma_k \varepsilon_{k,j}$, where $\varepsilon_{k,j}$ is the $j^{th}$ element of $\varepsilon_k$. We set $\hat{q}_j = \mu_q + \sigma_q \frac{\alpha_{q,k}\varepsilon_{k,j} + \varepsilon_{q,j}}{\left(\alpha_{q,k}^2+1\right)^{1/2}}$. Finally, we set $\hat{m}_j = \frac{\alpha_{m,k}\varepsilon_k + \alpha_{m,q}\varepsilon_{q,j} + \varepsilon_{m,j}}{\left(\alpha_{m,k}^2+\alpha_{m,q}^2+1\right)^{1/2}}$. The $\alpha$ parameters govern the correlations of the endowments. The numerators scale the distributions to match the desired standard deviations.

The ability grid $\hat{a}_i$ approximates a Normal distribution with mean $\bar{a}$ and variance 1. We set the number of grid points to $N_a = 9$. Each grid point has the same probability, $\Pr\left(\hat{a}_i\right) = 1/N_a$. We think of grid point $i$ as containing all continuous abilities in the set $\Omega_i = \left\{a : \frac{i-1}{N_a} \leq \Phi\left(a - \bar{a}\right) < \frac{i}{N_a}\right\}$ where $\Phi$ is the standard Normal cdf. We therefore set $\hat{a}_i = \mathbb{E}\left\{a | a \in \Omega_i\right\}$. We model the joint distribution of abilities and signals as a discrete

approximation of a joint Normal distribution given by

$$a = \bar{a} + \frac{\alpha_{a,m} m + \varepsilon_a}{\left(\alpha_{a,m}^2 + 1\right)^{1/2}}, \tag{14}$$

where $\varepsilon_a \sim N(0,1)$. The denominator ensures that the unconditional distribution of $a$ has a unit variance. We set $\Pr(\hat{a}_i|j) = \Pr\left(a \in \Omega_i | m = \hat{m}_j\right)$. For notational convenience, we normalize $\bar{a}$ such that $\hat{a}_i = 0$. For computational efficiency, we draw all Normal random variables using Halton quasi random numbers.

## 3.2 Mapping of Model and Data Objects

We discuss how we conceptually map model objects into data objects. Variables without observable counterparts include abilities, ability signals, consumption, and preference shocks. 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 year 2000 prices.

**Schooling.** We count a student as attending college if he attempts at least 9 non-vocational credits in a given year.[11] In NELS:88 data, 70% of community college entrants intend to attain a 4-year college degree (Bound, Lovenheim, and Turner, 2010). We therefore classify persons who ever attended college without attaining a 4-year degree as college dropouts. In our HS&B data, only 10% of these students earn a training certificate or an associate's degree.

**College credits.** We measure $n_t$ as the number of completed college credits by the start of college year $t$ divided by the number of credits taken, assuming a full course load, which is defined as the number of credits attempted by students who eventually graduate from college. In the data, college dropouts attempt fewer credits than college graduates. Since our model abstracts from variation in course loads, we treat taking less than a full course load as failing the courses that were not taken. This captures the fact that taking fewer courses slows a student's progress towards graduation, which is a key element of our model.

[11] Students attending vocational schools (e.g., police or beauty academies) are classified as high school graduates.

**Test scores.** For calibration purposes, it is helpful to utilize test scores to proxy for unobserved abilities. In the calibration, we divide agents into test score quartiles.

In NLSY data, we use the 1989 Armed Forces Qualification Test (AFQT) percentile rank. 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.[12]

Since HS&B lacks AFQT scores, we treat high school GPA quartiles as equivalent to AFQT quartiles. Borghans, Golsteyn, Heckman, and Humphries (2011) show that high school GPAs and AFQT scores are highly correlated. Sidestepping the question what cognitive test scores measure (see Flynn 2009), we use the term "test scores" in the text and the symbol $IQ$ in mathematical expressions.

In mapping test scores to the model, we assume that test scores are noisy measures of the ability signals observed by the agents. This implies that the agents know more about their abilities than we do. Specifically, we model test scores as signal plus Gaussian noise:

$$IQ = \frac{\alpha_{IQ,m} m + \varepsilon_{IQ}}{\left(\alpha_{IQ,m}^2 + 1\right)^{1/2}} \tag{15}$$

with $\varepsilon_{IQ} \sim N(0,1)$. If $m$ were continuous, the distribution of test scores would be standard Normal. Since $m$ is restricted to take on values on the grid $\hat{m}_j$, only the conditional distribution $IQ|m$ is Normal.

**Financial variables.** We interpret college costs $q$ as collecting all college related payments that are *conditional* on attending college. In HS&B data, we measure tuition and fees net of scholarships, grants, and labor earnings.[13] We set $q$ equal to the average of these values over the first two years in college plus $987 for other college expenditures, such

[12] Some persons take the AFQT after graduating from high school. This raises the concern that the AFQT partly measures skills learned in college. To address this concern, we experimented with removing age effects using a separate regression for each school group. This makes little difference.

[13] Appendix C.2 describes the financial data in detail.

as books, supplies, and transportation.[14] $q$ does not include room and board, which are included in consumption.

We interpret $k_1$ as collecting financial resources the student receives *regardless* of college attendance. In the data, we measure $k_1$ as the student's financial assets at high school graduation and any transfers received from his parents in the six years that follow. In the model we assume that $k_1$ is paid out as a lump sum at high school graduation.[15] As students move through college, $k_t$ may fall below zero, which we interpret as student debt.

## 3.3 Fixed Parameters

Table 1 summarizes the values of parameters that are fixed a priori.

1. The discount factor is $\beta = 0.98$.

2. Based on McGrattan and Prescott (2000), the gross interest rate is set to $R = 1.04$.

3. The scale of the preference shock that governs consumption choice is set to $\pi_c = 0.2$. This value is low enough that most students choose their preferred consumption level, but high enough that the objective function minimized by the calibration algorithm is smooth in the parameters.

4. Motivated by the fact that in our HS&B sample 95% of college graduates finish college by their 6th year (Bowen, Chingos, and McPherson 2009 report a similar finding), we set the maximum duration of college to $T_c = 6$. The number of credits needed to graduate is set to $n_{grad} = 20$. In each year, students attempt $n_c = 5$ credits. This number is set so that students who pass most of their courses graduate in 4 or 5 years, which accords with the data.[16]

[14] Since HS&B lacks information on these expenditures, we compute them as the average cost for 1992-93 undergraduate full-time students in the National Postsecondary Student Aid Study, conducted by the U.S. Department of Education. These costs are defined as the amount student reported spending on expenses directly related to attending classes, measured in year 2000 prices.

[15] If $k_1$ is paid out over time, the household problem gains a state variable, which is computationally costly.

[16] In the data, students typically complete around 130 credits by the time of college graduation. Increasing the number of model credits would increase the number of ability signals a student receives in each period, which may affect the rate of learning. It is, however, computationally costly.

Table 1: Fixed Model Parameters

| Parameter | Description | Value |
|---|---|---|
| Preferences | | |
| $\beta$ | Discount factor | 0.98 |
| $\pi_c$ | Scale of preference shocks at consumption choice | 0.20 |
| College | | |
| $T_c$ | Maximum duration of college | 6 |
| $n_{grad}$ | Number of credits required to graduate | 20 |
| $n_c$ | Number of credits attempted each year | 5 |
| $k_{min}$ | Borrowing limit | -$19,750 |
| Other | | |
| $R$ | Gross interest rate | 1.04 |

5. Borrowing limits are set to approximate Stafford loans, which are the predominant form of college debt for the NLSY79 cohorts (see Johnson 2010). Until 1986, students could borrow $2,500 in each year of college up to a total of $12,500 ($19,750 in year 2000 prices). We ignore the restriction that loan amounts cannot exceed college related expenditures and set $k_{min} = -\$19,750$.

## 3.4 Calibrated Parameters

The remaining model parameters are jointly calibrated to match the target data moments summarized in Table 2. We show the data moments in Section 3.5 where we compare our model with the calibration targets. Appendix D discusses the identification of key model parameters.

For each candidate set of parameters, the calibration algorithm simulates the life histories of 100,000 individuals. It constructs model counterparts of the target moments and searches for the parameter vector that minimizes a weighted sum of squared deviations between model and data moments.

Table 3 shows the values of the 20 calibrated parameters. We highlight parameters that are important for our findings.

Table 2: Calibration Targets

| Target | Value |
|---|---|
| Fraction in population, by (test score quartile, schooling) | Figure 1 |
| Lifetime earnings, by (test score quartile, schooling) | Figure 2 |
| Dropout rate, by (test score quartile, year in college) | Figure 3 |
| Fraction of credits passed, by graduation status and year | Table 6 |
| Mean and standard deviation of $k_1$ (HS and college) | Table 7 |
| Mean and standard deviation of $q$ (college) | Table 7 |
| Fraction of students in debt, by year in college | Table 8 |
| Mean student debt, by year in college | Table 8 |
| Average time to BA degree (years) | 4.4 |

Notes: Schooling and lifetime earnings targets are taken from NLSY79 data. The remaining targets are taken from HS&B data.

The first section of the table reports parameters governing the joint distribution of initial endowments. Table 4 shows the implied endowment correlations. Abilities and signals are highly correlated. Still, high school graduates face substantial uncertainty about their abilities. From (14), the standard deviation of $a$ conditional on $m$ is $\left(\alpha_{a,m}^2 + 1\right)^{-1/2} = 0.32$, compared with an unconditional standard deviation of 1. This feature helps the model account for the timing of college dropouts (see Appendix D).

High ability students not only enjoy larger wage gains from attending college, they also have more assets and face lower college costs. This allows the model to capture the empirical findings that average college costs among college students are very close to zero and negatively correlated with high school GPAs (see Table 7).[17]

The middle section of Table 3 reports the parameters that govern lifetime earnings. The effective dispersion of abilities is governed by $\phi_s$. A one standard deviation increase in ability raises lifetime earnings by 0.15 for high school graduates and by 0.19 college graduates. These values are similar to the ones estimated by Hendricks and Schoellman (2011). The sensitivity analysis in Appendix D shows that larger values of $\phi_s$ are associated with a

[17] This is consistent with Bowen, Chingos, and McPherson (2009) who report that average tuition payments for public 4-year colleges roughly equal average scholarships and grants.

Table 3: Calibrated Parameters

| Parameter | Description | Value |
|---|---|---|
| Endowments | | |
| $\mu_k, \sigma_k$ | Marginal distribution of $\ln(k_1)$ | $0.41, 1.17$ |
| $\mu_q, \sigma_q$ | Marginal distribution of $q$ | $3.01, 5.81$ |
| $\alpha_{m,k}, \alpha_{m,q}, \alpha_{q,k}, \alpha_{a,m}, \alpha_{IQ,m}$ | Endowment correlations | $0.23, -0.11, -0.44, 2.97, 1.78$ |
| Lifetime earnings | | |
| $\phi_{HS}, \phi_{CG}$ | Effect of ability on lifetime earnings | $0.153, 0.194$ |
| $y_{HS}, y_{CG}$ | Lifetime earnings factors | $3.90, 3.91$ |
| $\mu$ | Earnings gain for each college credit | 0.014 |
| Other parameters | | |
| $\pi$ | Scale of preference shocks | 0.767 |
| $U_{CD}, U_{CG}$ | Preference for job type $s$ | $-1.11, -2.98$ |
| $\gamma_1, \gamma_2, \gamma_{min}$ | Probability of passing a course | $0.68, 7.89, 0.42$ |

Table 4: Correlation of Endowments

| | $IQ$ | $a$ | $m$ | $q$ |
|---|---|---|---|---|
| $IQ$ | 1.00 | | | |
| $a$ | 0.77 | 1.00 | | |
| $m$ | 0.78 | 0.90 | 1.00 | |
| $q$ | -0.14 | -0.16 | -0.12 | 1.00 |
| $k_1$ | 0.15 | 0.17 | 0.19 | -0.34 |

larger contribution of ability selection to the measured college premium.

A college degree makes two contributions to lifetime earnings. (i) Students complete college credits. Completing 20 credits, which is required for graduating from college, increases lifetime earnings by 28 log points. (ii) Students enjoy ability-dependent "sheepskin effects" which increase log lifetime earnings by $y_{CG} - y_{HS} + (\phi_{CG} - \phi_{HS})a$. Since $\phi_{CG} > \phi_{HS}$, high ability students enjoy larger returns to college. Since we set $\hat{a}_1 = 0$, the fact that $y_{CG}$ is close to $y_{HS}$ implies that sheepskin effects are very small for students of the lowest abilities.

## 3.5 Model Fit

This Section assesses how closely the model attains each set of calibration targets.

**Schooling and lifetime earnings.** Table 5 shows that the model closely fits the observed fraction of persons attaining each school level and their mean log lifetime earnings. Key features of the data are: (i) 46% of those attempting college fail to attain a bachelor's degree. (ii) College graduates earn 45 log points more than high school graduates over their lifetimes. For college dropouts, the premium is only 8 log points.

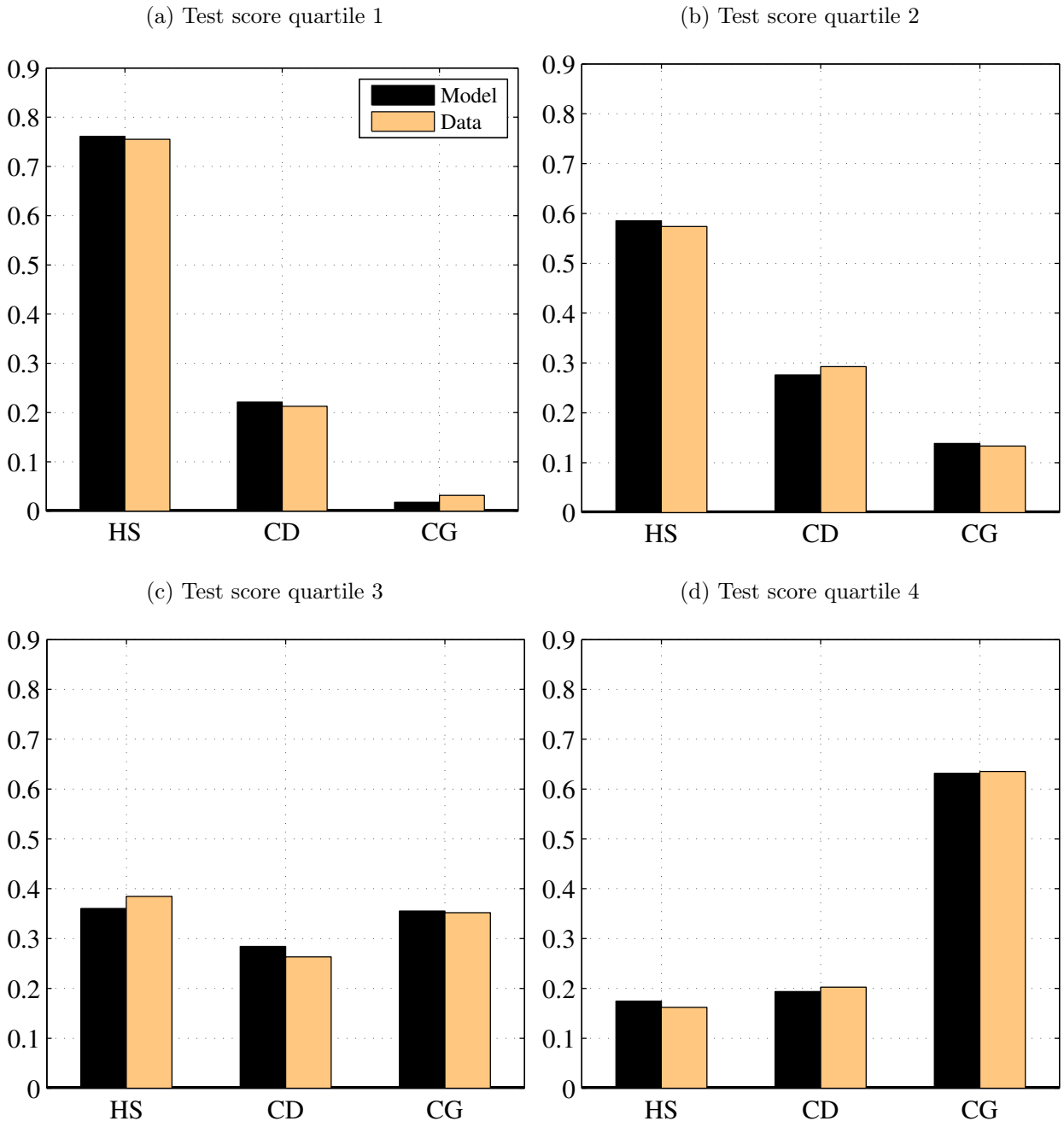
Figure 1 breaks down the schooling outcomes by test score quartiles. The model replicates the patterns observed in the data. Test scores are strong predictors of college entry and college completion. More than 80% of students in the top test score quartile attempt college and more than 60% earn college degrees. In the lowest test score quartile, only 20% of students enter college and fewer than 5% earn degrees.[18] One question our model answers is why these students attempt college, even though their graduation prospects are bleak.

Figure 2 shows mean log lifetime earnings by school group and test score quartile. Each panel displays one school group. The model broadly matches the data cells with large numbers of observations. The largest discrepancy occurs for college graduates in the lowest test score quartile, which are quite rare (22 observations).

**Dropout rates.** Figure 3 compares dropout rates between the model and High School & Beyond data. Dropout rates are defined as the number of persons dropping out at the

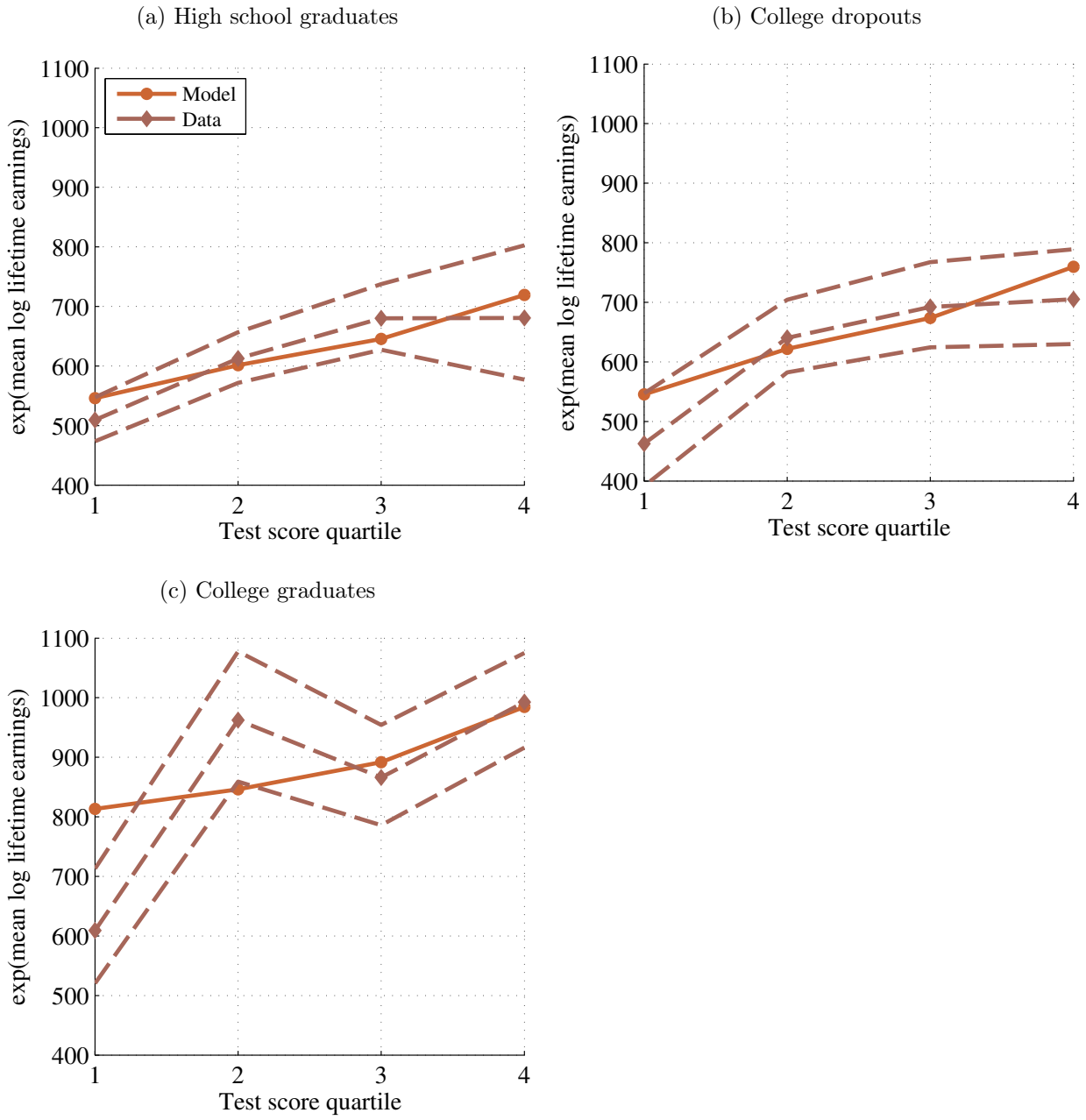
[18] Bound, Lovenheim, and Turner (2010)'s Figure 2 documents similar patterns in NLS72 and NELS:88 data.

Figure 1: Schooling and Test Scores




Notes: For each test score quartile, the figure shows the fraction of persons who attain each schooling level.

Figure 2: Lifetime Earnings




Notes: The figure shows the exponential of mean log lifetime earnings, discounted to model age 1 and expressed in thousands of year 2000 dollars, for each school group and test score quartile. Dashed lines show two standard error bands.

Table 5: Schooling and Lifetime Earnings

| | School group | | |
|---|---|---|---|
| | HS | CD | CG |
| Fraction | | | |
| Data | 46.9 | 24.3 | 28.8 |
| Model | 47.1 | 24.4 | 28.5 |
| Gap (pct) | 0.4 | 0.4 | -1.0 |
| Lifetime earnings | | | |
| Data | 600 | 643 | 944 |
| Model | 596 | 643 | 934 |
| Gap (pct) | -0.7 | -0.0 | -1.0 |

Note: The table shows the fraction of persons that chooses each school level and the exponential of their mean log lifetime earnings, discounted to age 1, in thousands of year 2000 dollars. "Gap" denotes the percentage gap between model and data values.

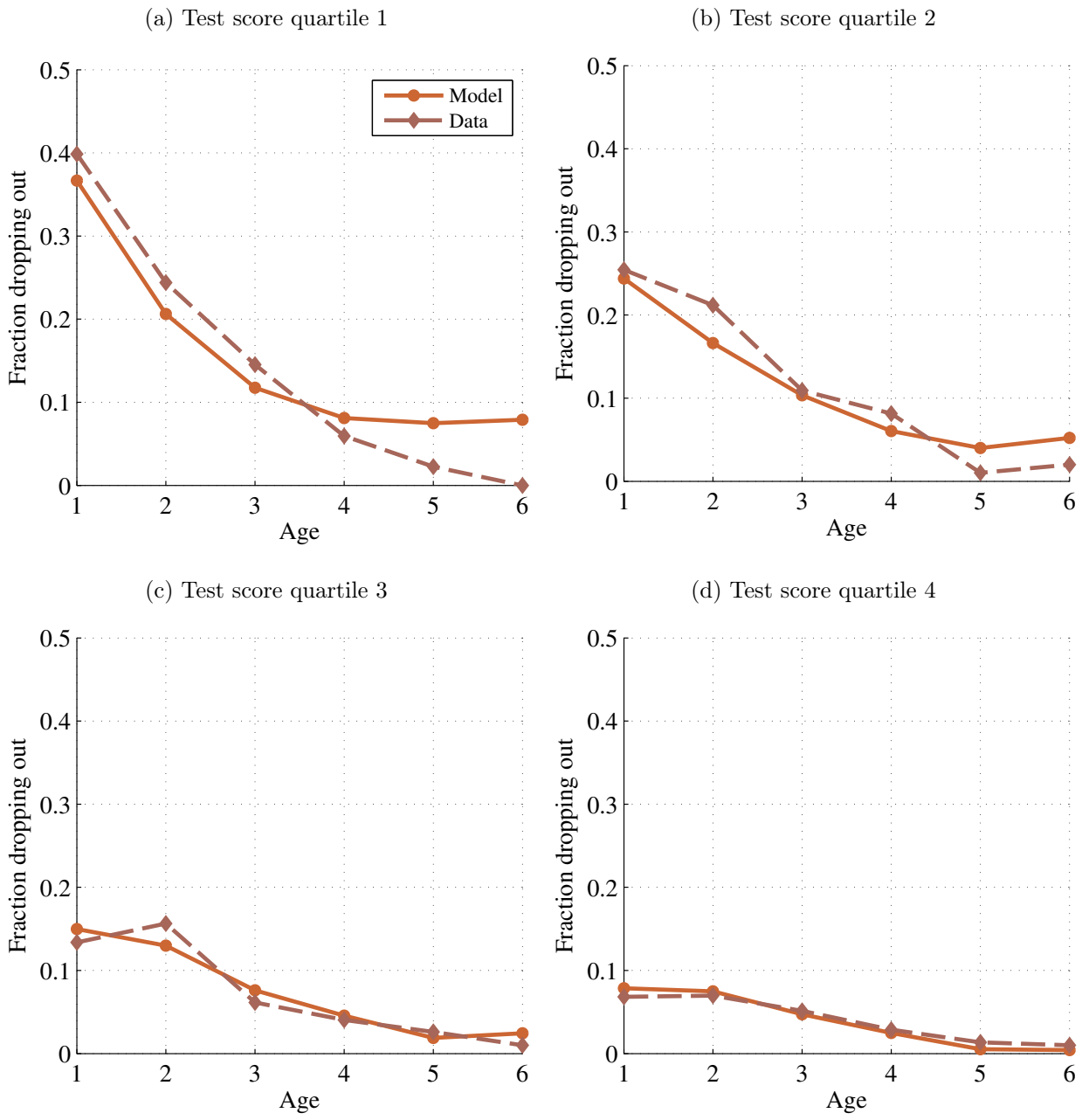
end of each year divided by the number of college entrants in year 1. Dropout rates decline strongly with test scores and with time spent in college.

**College credits.** Table 6 shows the credit passing rate for each year in college. In the model, the credit passing rate is defined as $n_{t+1}/(tn_c)$. In the data, it is defined as the number of completed credits divided by a full course load (see Section 3.2). Students are divided into two groups: those who eventually drop out and those who eventually earn a college degree.

While college graduates pass around 95% of the credits they attempt, college dropouts pass only around two-thirds. The gap in passing rates is roughly the same across years. It follows that college dropout freshmen would have to expect their passing rate to improve dramatically over time, if they wanted to graduate within five years. The model implies passing rates that are close to the data in all years.

**Financial resources.** Table 7 reveals that the model effectively matches the means and standard deviations of initial assets ($k_1$) for high school graduates and for college entrants.

Figure 3: Dropout Rates




Notes: The figure shows the fraction of persons initially enrolled in college who drop out at the end of each year in college.

Table 6: Credit Passing Rates

| | College dropouts | | College graduates | |
|---|---|---|---|---|
| Year | Model | Data | Model | Data |
| 1 | 66.7 | 67.7 | 95.9 | 98.7 |
| 2 | 67.9 | 71.8 | 95.8 | 96.3 |
| 3 | 64.7 | 66.9 | 95.8 | 95.7 |
| 4 | 57.7 | 63.8 | 95.8 | 94.9 |

Notes: The credit passing rate is the number of college credits completed at the end of each year divided by a full course load.

The distribution of college costs ($q$) is only observed for college students, where the mean of $q$ is slightly negative. Even though, in the population, higher ability (test score) students face lower college costs, the correlation is reversed among college students. This results from selection. Low ability students only enter college, if it is very cheap.

Table 8 shows student debt levels at the end of the first 4 years in college. The model roughly matches mean debt levels, conditional on being in debt. However, it understates the fraction of persons in debt early on, but overstates it in later years. One reason is that, in the model, all parental transfers are received at age 1 while, in the data, transfers are received in each year. The model therefore overstates measured assets during the early years in college. As a result, the fraction of students in debt is too small. At the same time, asset levels decline too fast over time because no new transfers are received, which leads the model to overstate the fraction of indebted students in year 4.

# 4 Results

## 4.1 Ability Selection

This section presents our main finding. Part of the lifetime earnings gap between college graduates and high school graduates represents ability differences between the two groups rather than returns to schooling. We use our calibrated model to measure this part.

In the model, the mean log lifetime earnings of school group $s$, discounted to age 1, are

Table 7: Financial Moments

(a) Entire population

| | Model | Data |
|---|---|---|
| Distribution of $k_1$, HS | | |
| mean | 16,770 | 16,630 |
| standard deviation | 22,867 | 23,266 |
| Distribution of $k_1$, college | | |
| mean | 38,011 | 37,390 |
| standard deviation | 37,329 | 38,475 |
| Distribution of $q$, college | | |
| mean | -740 | -584 |
| standard deviation | 4,928 | 5,787 |

(b) Test score quartiles

| | Mean $q$ | | Standard deviation | | |
|---|---|---|---|---|---|
| Test score quartile | Model | Data | Model | Data | $N$ |
| 1 | -2,934 | -2,266 (678) | 5,075 | 5,253 | 60 |
| 2 | -1,362 | -1,741 (454) | 4,896 | 6,121 | 182 |
| 3 | -560 | -509 (308) | 4,924 | 5,692 | 341 |
| 4 | -173 | -20 (253) | 4,764 | 5,704 | 510 |

Notes: Panel (a) shows the means and standard deviations of $k_1$ among high school graduates and college students. College costs $q$ are only observed for college students. Panel (b) breaks down the college costs by test score quartile. Estimated standard deviations of mean $q$ are shown in parentheses. $N$ is the number of observations in each test score quartile. In HS&B data, test scores are high school GPAs.

Table 8: Student Debt

| | Mean debt | | Fraction with debt | |
|---|---|---|---|---|
| Year | Model | Data | Model | Data |
| 1 | 5,827 | 3,549 | 15.3 | 27.7 |
| 2 | 6,740 | 6,060 | 27.6 | 36.0 |
| 3 | 7,907 | 8,045 | 49.8 | 42.5 |
| 4 | 11,000 | 9,740 | 72.1 | 48.0 |

Notes: The table shows the fraction of students with college debt ($k < 0$) at the end of each year in college. Mean debt is conditional on being in debt.

given by

$$\mathbb{E}[\phi_s a + \mu n_\tau + y_s + \ln(R^{-\tau})|s], \tag{16}$$

where $\tau = 1$ and $n_\tau = 0$ for high school graduates. The mean log lifetime earnings gap between school group $s$ and high school graduates may then be decomposed into four terms:

1. prices: $y_s - y_{HS} + (\phi_s - \phi_{HS})\mathbb{E}(a|s)$;
2. credits: $\mathbb{E}(\mu n_\tau|s)$;
3. delayed labor market entry: $\mathbb{E}\{\ln R^\tau|s\} - \ln R^{-1} = \mathbb{E}\{\ln R^{1-\tau}|s\}$;
4. ability selection: $\phi_{HS}[\mathbb{E}(a|s) - \mathbb{E}(a|HS)]$.

For a student of given ability, earning a college degree has three effects on lifetime earnings. (i) It changes the skill prices earned in the labor market. (ii) It requires a certain number of earned college credits. (iii) Earning these credits delays entry into the labor market, which reduces lifetime earnings. Taken together, these three effects represent the return to college graduation. As in much of the recent related literature, the (ex post) return to schooling varies across individuals (see Card 2001). The remaining gap between the mean log earnings of college graduates and high school graduates represents ability selection.

Table 9 shows the decomposition implied by the model. College graduates earn 45 log points more than high school graduates. Since postponing entry into the labor force reduces lifetime earnings by 18 log points, it follows that completing college increases lifetime earnings, discounted to age $\tau$, by 63 log points. Of this increase, 30 log points are due to

Table 9: Ability Selection

| Gap relative to HS | College dropouts | | College graduates | |
|---|---|---|---|---|
| (in log points) | Gap | Fraction | Gap | Fraction |
| Total gap | 8 | – | 45 | – |
| Delayed labor market entry | -9 | -124 | -18 | -39 |
| Prices: $y_s$ and $\phi_s$ | 0 | 0 | 11 | 24 |
| Credits | 11 | 143 | 30 | 67 |
| Ability selection | 6 | 81 | 22 | 48 |

Notes: Row 1 shows mean log lifetime earnings of college dropouts and college graduates relative to high school graduates. The remaining rows decompose these lifetime earnings gaps into the contributions of various factors defined in the text. "Fraction" denotes the fraction of the lifetime earnings gap due to each factor.

credit accumulation, 11 log points are due to prices ($y_s$ and $\phi_s$), and the remaining 22 log points (48% of the college lifetime earnings premium) are due to ability selection.

For college dropouts, the mean log earnings gap relative to high school graduates is much smaller (8 log points). By assumption, the effect of prices is zero. The effect of earned credits is only marginally larger than the cost of delayed labor market entry, implying that most of the earnings gap relative to high school graduates is due to ability selection.

The finding that ability selection accounts for roughly half of the college earnings premium is quite robust, as we show in Appendix E. One reason why ability sorting is strong is that it occurs at two levels: Selection at college entry accounts for 67% of the ability gap between college graduates and high school graduates.[19] Selection at college completion accounts for the remaining 33%.

Figure 4 illustrates both levels of selection. Panel (a) shows school outcomes for students in each signal decile. College entry is strongly related to ability signals. While only 22% of students in the lowest decile attempt college, 89% of students in the highest signal decile do. Panel (b) shows the same information, sorting students by ability level rather than signal. Since abilities and signals are strongly correlated, the two figures are similar. College entry rates range from 20% among low ability students to 89% among high ability students.

[19] $\mathbb{E}\{a|CD \vee CG\} - \mathbb{E}\{a|HS\} = 0.67\ (\mathbb{E}\{a|CG\} - \mathbb{E}\{a|HS\})$.

The second level of selection, college graduation, also depends strongly on abilities. The fraction of college entrants that graduates varies from near zero for the lowest ability level to 84% for the highest. Taken together the two levels of selection imply that the ability distributions for high school graduates and college graduates are strongly separated. Low ability students rarely attempt college and almost never graduate. High ability students typically attempt college and rarely drop out.[20]

One contribution of our analysis is to highlight how the two levels of selection interact. At college entry, low ability students recognize that their graduation prospects are poor. This deters them from attempting college. This interaction is absent in models that abstract from college completion risk. To quantify this interaction, we compute a version of the model where all agents face the same probability of passing courses. We set $\Pr_c(a)$ to the constant level that keeps the college entry rate the same as in the benchmark model. This modification cuts ability selection roughly in half. Even though the experiment affects college graduation prospects, most of the changes in selection occur at college entry.

## 4.2 Understanding College Entry

Figure 5 summarizes the two key considerations that determine an individual's college entry decision: lifetime earnings and graduation probabilities.

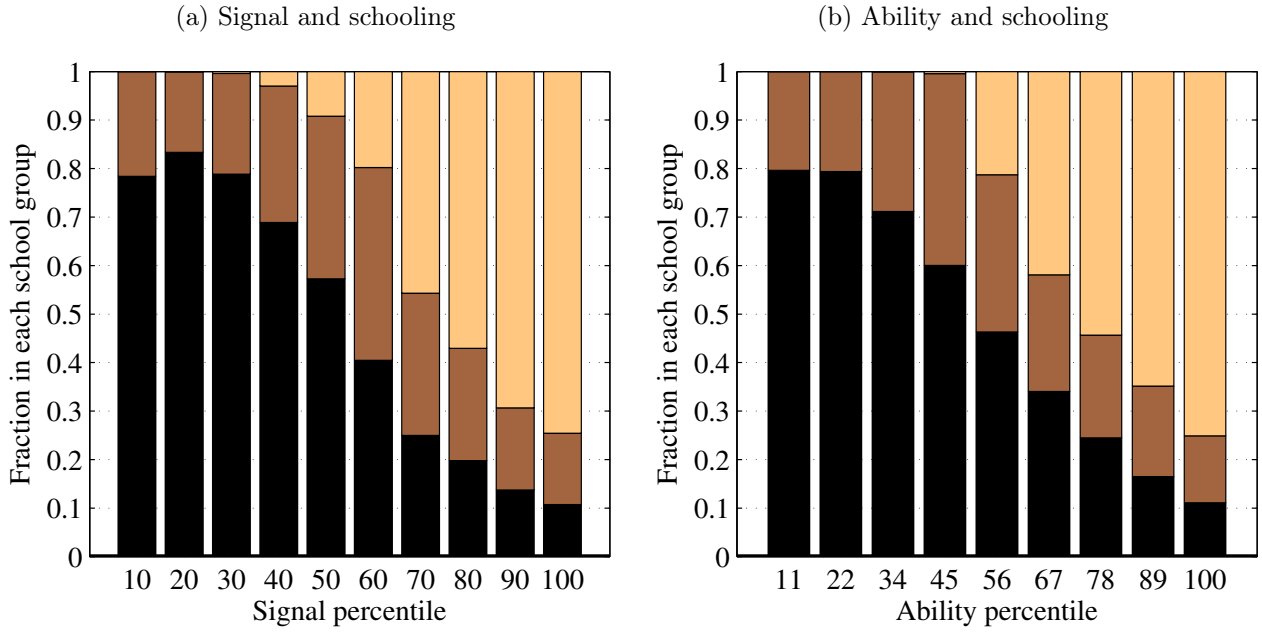
Panel (a) shows mean log lifetime earnings by school outcome and ability.[21] It summarizes the financial stakes that motivate entry and dropout decisions. Only high ability students can expect large gains from earning a college degree. The earnings gains from completing college increase from 16 log points for students of median abilities to 27 log points for students with the highest ability level. There are two reasons for this: (i) higher ability students graduate earlier, and (ii) the complementarity implied by $\phi_{CG} > \phi_{HS}$.

The gains from attending college without earning a degree are much smaller. For students of low abilities, dropping out of college reduces lifetime earnings as the costs of delayed

[20] School outcomes are also correlated with financial endowments. Students who face lower college costs or who have more assets are more likely to enter college and more likely to graduate, conditional on entry. These correlations are, in part, due to the correlation between abilities and financial endowments. To conserve space, we do not show the details.

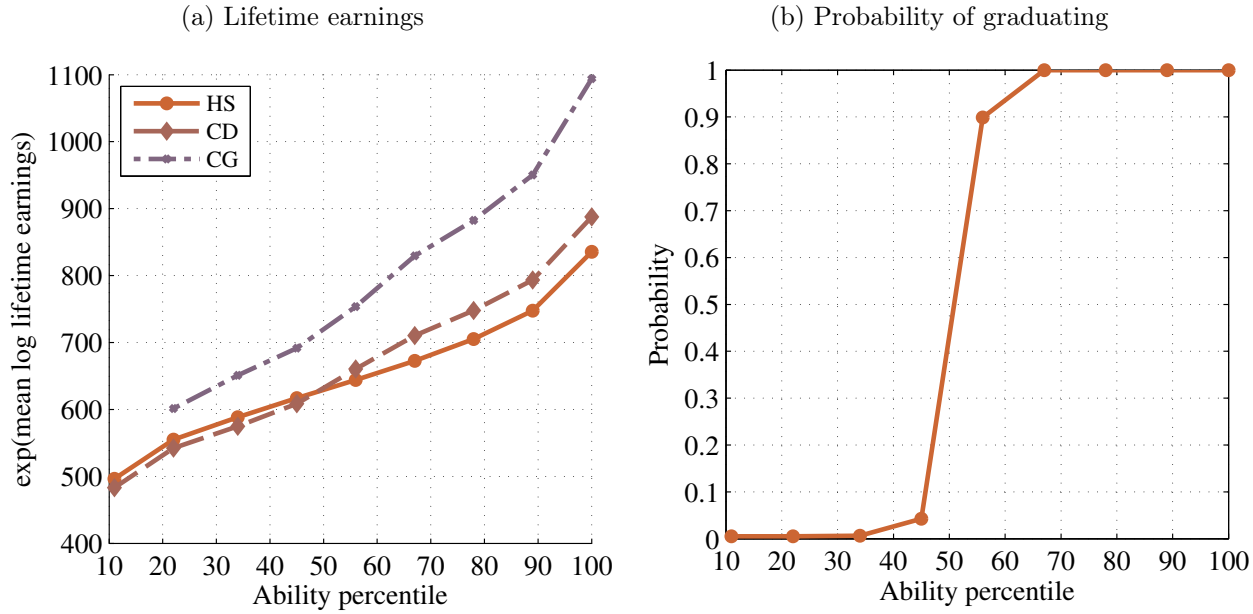
[21] Since the probability of graduating from college is very small for students with the lowest abilities, the model does not generate a lifetime earnings number for this group.

Figure 4: Schooling and Endowments




Notes: Each bar shows the fraction of persons attaining each school level (HS, CD, CG). Panel (a) sorts students into ability signal deciles. Panel (b) shows the outcomes for each of the $N_a = 9$ ability levels.

Figure 5: Abilities and Outcomes




Notes: Panel (a) shows the exponential of mean log lifetime earnings of students who attain each school level in thousands of year 2000 dollars. Panel (b) shows the probability of earning at least $n_{grad}$ credits in $T_c$ periods.

labor market entry outweigh the benefits of earning college credits. These small earnings gains could explain why college students spend little time studying while at the same time working for modest wages (Babcock and Marks, 2010).

In contrast to the commonly used Roy model, the large earnings gains that accrue to college graduates are not available to all agents. Only high ability students can expect to graduate from college. To illustrate this point, panel (b) shows the probability that a person of a given ability earns enough credits to graduate from college, if he remains in college for the maximum permitted number of periods.

The chances of graduating from college depend strongly on ability. Low ability students have essentially no chance of graduating. High ability students are virtually guaranteed to graduate. This is a robust feature of our model because the number of completed courses is drawn from a Binomial distribution. The probability of passing more than $n_{grad} = 20$ out of $n_c T_c = 30$ courses increases sharply in the probability of passing a single course.

This is a key feature of our model, which generates ability separation between college graduates and high school graduates. It implies that the payoff from attempting college

increases far more sharply with ability than lifetime earnings differences would suggest. This is important for understanding dropout behavior (see Section 4.3).

### 4.2.1 College entry incentives and ability signals

Since students do not observe their abilities, college entry depends on how earnings and graduation rates vary with the ability *signal*. This is summarized in Figure 6.

Panel (a) shows that college entry is strongly related to the ability signal and the associated chance of graduating from college (earning $n_{grad}$ credits in $T_c$ periods). The majority of students with graduation probabilities above 0.6 attempt college. However, the fraction of students that actually graduate is substantially lower than the fraction that could have earned the required number of credits. The gap is especially large for students with intermediate signals.

Panel (b) shows mean log lifetime earnings, discounted to age 1, received for each school outcome.[22] For all signal levels, completing college increases lifetime earnings by at least \$100,000.[23] However, dropping out of college yields much smaller, or even negative, earnings gains. One reason is that dropouts accumulate fewer credits per year in college. A second reason is that dropouts do not enjoy the price effects associated with graduation ($y_s$ and $\phi_s$).

The expected lifetime earnings gains due to attempting college increase sharply with the ability signal. Since students with low signals typically fail to graduate, attempting college reduces their lifetime earnings slightly. Since students with high signals typically graduate, they can expect to gain more than \$200,000 by attempting college.

One puzzling observation in NLSY data is that a sizeable fraction of students with low test scores and graduation rates attempts college (see Figure 1). Our model offers an explanation for this puzzle. As an example, consider students with ability signals in the median decile. Even though their gradation rate is only 22%, 43% of the students in this group attempt college. One reason is that the earnings losses from dropping out are quite

[22] Log lifetime earnings are averaged across simulated households in a given signal decile who choose a particular schooling level.

[23] The large earnings gains for low signal students are consistent with the small earnings gains for low ability students shown in Figure 5. The reason is selection. Conditional on graduating from college, even low signal students have high abilities.

small, while the gains from graduating are quite large. The asymmetry arises because the option of dropping out in response to poor academic performance limits the potential losses. A student can try college for one year, observe his credit passing rate, update his beliefs about his graduation prospects, and drop out if the news is bad. At least for students with low college costs, this option is almost costless because dropping out has little effect on expected lifetime earnings. In addition, some students with poor graduation prospects enter college because they receive large subsidies ($q$ is negative).

An important implication of the model is that low and high ability students respond to different incentives when deciding whether or not to enter college. High ability students typically attempt college in order to graduate and increase their lifetime earnings. Since college costs represent only a small fraction of lifetime earnings, these students are not sensitive to tuition changes. Low ability students, on the other hand, understand that their graduation prospects are poor. They only enter college if it is sufficiently cheap, and their entry decisions are highly sensitive to tuition costs. We return to this insight when we perform comparative statics experiments in Section 5.

## 4.3 Understanding College Dropouts

This section examines why nearly half of all students drop out of college. Our model offers three main reasons: money, luck, and preference shocks.

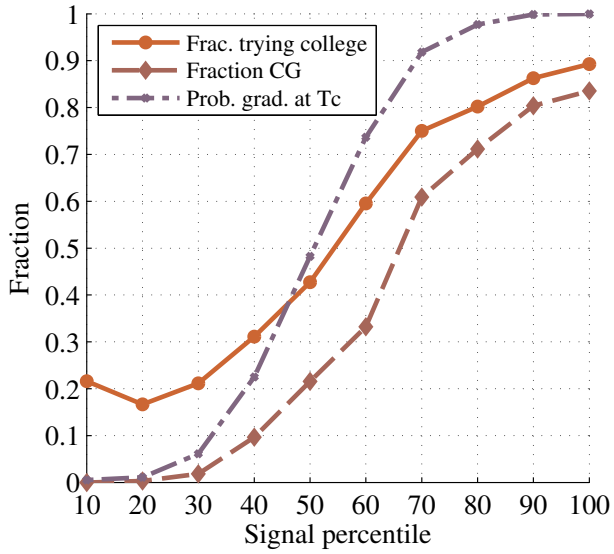
Some students in our model choose *ex ante* to drop out of college. They choose a consumption level that is so high that they run out of assets before it is feasible to graduate.[24] Panel (a) of Figure 7 shows the fraction of college students in each signal quintile who choose such high consumption levels. Among students with low ability signals, 40% make this choice. These students believe that they are of low ability, which renders college financially unattractive. Panel (b) reveals why these students attend college: their college costs are negative and their consumption in college is high (relative to that of higher signal dropouts).

Among students in the top signal quintile, more than 15% plan to drop out. These students would enjoy high earnings gains upon graduation. However, since either their college costs are high or their assets are low, these students would have to choose very low consumption

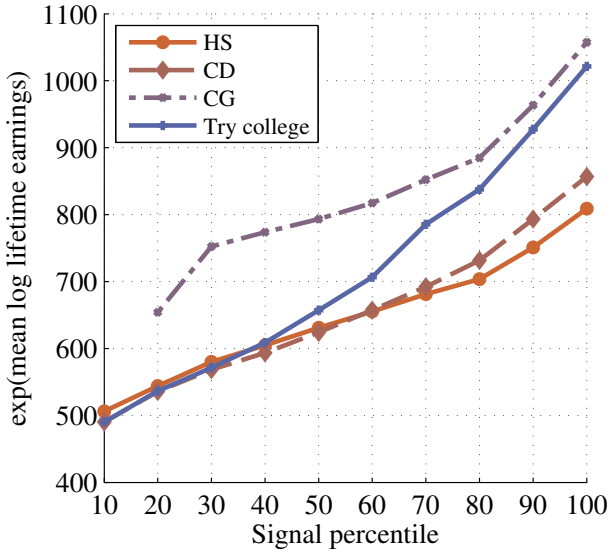
[24] Given that a student can earn at most $n_c$ credits per year, it is not possible to graduate in fewer than 4 years.

Figure 6: Signals and Outcomes

(a) Signals and graduation probabilities

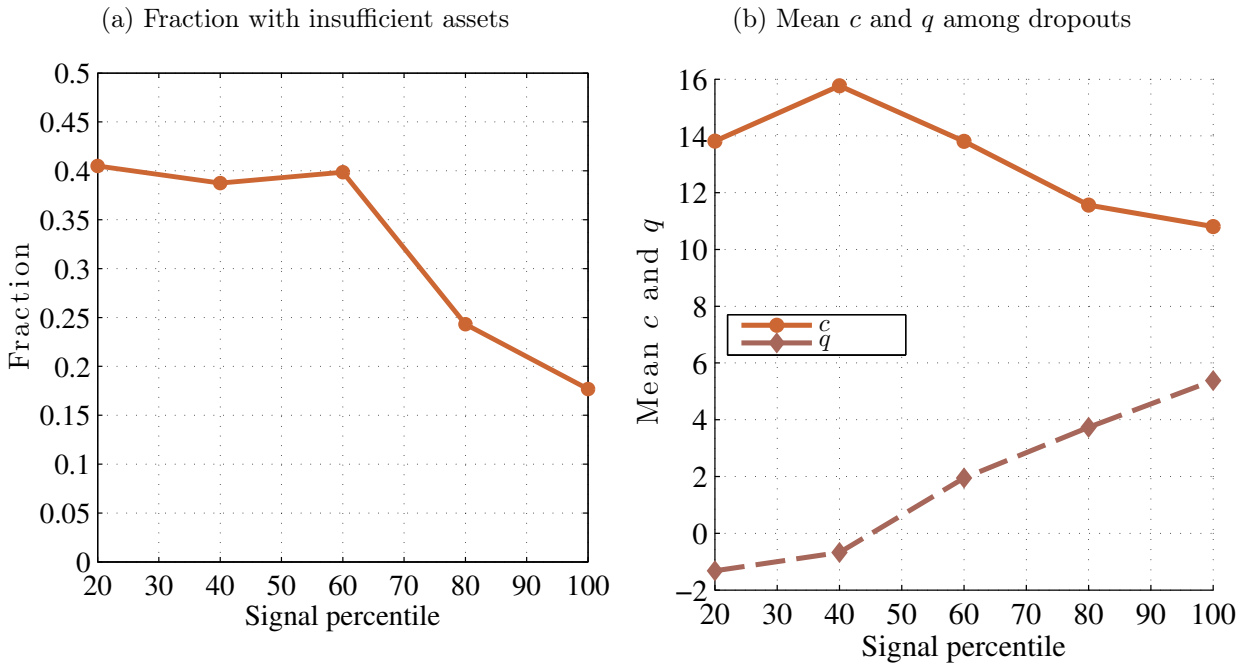



(b) Signals and lifetime earnings




Notes: For students in each ability signal decile, panel (a) shows the fraction of high school graduates that attempt college, the fraction of college entrants that graduates, and the probability of earning at least $n_{grad}$ credits in $T_c$ years. Panel (b) shows the exponential of mean log lifetime earnings for each school outcome and conditional on attempting college.

Figure 7: Understanding College Dropouts




Notes: Panel (a) shows the fraction of college entrants who choose consumption so high that they have insufficient assets to graduate. Panel (b) shows mean college costs and college consumption among dropouts in thousands of year 2000 dollars.

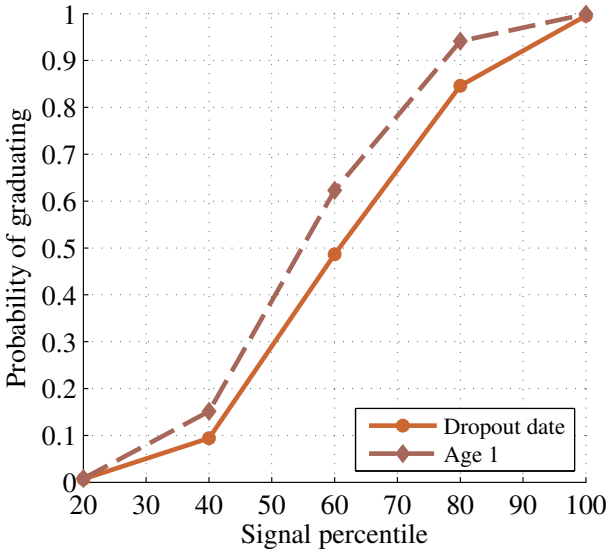
levels, if they wanted to stay in college for a long time. We show in Section 5.2 that these students respond strongly to increased borrowing opportunities.

The second reason for dropping out is bad *luck*. Consistent with the data, our model implies that college dropouts have low credit completion rates (see Table 6). In response, these students update their beliefs about their graduation prospects and some drop out.

For students in each signal decile, Figure 8 shows the probability of graduating from college conditional on staying in college for $T_c$ periods. The dashed line shows the students' beliefs before starting college. The solid line shows their beliefs at the time of dropping out. Dropouts of intermediate signals receive bad news during their college careers that lead to a substantial downward revision in their graduation probabilities. This model implication is consistent with the evidence of Stinebrickner and Stinebrickner (2012) who find that academic performance is strongly related to dropout decisions.

The last reason for dropping out is *preference shocks*. To isolate their effects, we recompute the model setting the realizations of preference shocks during college to zero. Students

Figure 8: Graduation Probabilities Among Dropouts




Notes: The figure shows the probability of earning $n_{grad}$ credits by the end of year $T_c$ among college dropouts. The probabilities are computed as of college entry (age 1) and at the time of dropping out of college.

follow the same decision rules as in the baseline model, so that their college entry decisions are not affected. This reduces the fraction of college entrants who drop out from 46% in the baseline model to 40%. Preference shocks mainly increase dropout rates among students of higher abilities.

# 5 Counterfactual Experiments

This section studies a number of counterfactual experiments. One objective is to learn more about how model agents respond to changed incentives. A second objective is to study the model's implications for policy questions.

## 5.1 Tuition Subsidies

The first pair of experiments illustrates a key feature of our model: High ability agents mainly view college as an investment, while low ability agents mainly view it as a consumption good. The two groups therefore respond very differently to changes in college costs

Table 10: Changing College Costs and Payoffs

| | School group | | |
|---|---|---|---|
| | HS | CD | CG |
| Fraction | | | |
| Baseline | 0.47 | 0.24 | 0.29 |
| Low tuition | 0.44 | 0.25 | 0.30 |
| High return | 0.44 | 0.21 | 0.35 |
| Mean log ability | | | |
| Baseline | -0.51 | -0.10 | 0.91 |
| Low tuition | -0.53 | -0.16 | 0.90 |
| High return | -0.57 | -0.29 | 0.88 |

and returns.

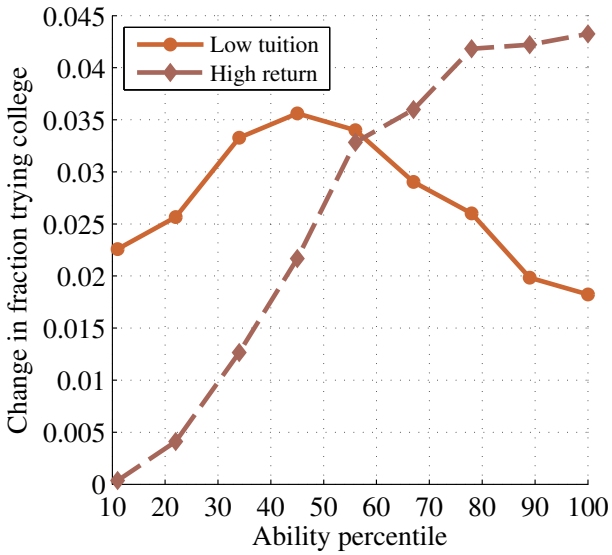
To illustrate this point, we study two experiments. The *low tuition* experiment reduces the mean of $q$ by \$1,000. This amount is chosen so that the model's implications can be compared with empirical estimates. The *high return* experiment increases $y_{CG}$ by 4 log points. This amount is chosen to yield roughly the same change in college enrollment as the low tuition experiment. For each case, we simulate individual life histories, holding all other parameters constant. Table 10 summarizes the changes in school attainment and $\mathbb{E}\{a|s\}$ for both experiments.

Consider first the low tuition experiment. College enrollment rises by 2.7 percentage points. The model's implications can be compared with a sizable empirical literature which 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 year 2000 prices) leads to a 3 to 4 percentage point increase in attendance. The model's implication is near the lower range of these estimates.

Figure 9 breaks down the change in college attendance by ability. Students of all abilities respond to tuition changes, with the largest responses occurring for median abilities. As a result, the fraction of college graduates rises by only 1.8 percentage points. Many of the new college entrants drop out.

The implications of the high return experiment are very different. Overall college enrollment rises by a very similar amount, 2.6 percentage points, but the fraction of college graduates

Figure 9: Changing College Costs and Payoffs




rises by 6.5 percentage points. The students that respond most to higher returns to college are drawn from the upper tail of the ability distribution (Figure 9). Most of these students graduate from college, so that the dropout rate declines.

From the perspective of the commonly used Roy model, it would seem surprising that college attendance responds so much to a change in tuition that represents a small fraction of lifetime earnings. On a per dollar basis, changing tuition has a much larger effect on college enrollment than changing lifetime earnings. A 4% increase in lifetime earnings of the average college graduate is worth about $40,000. Yet the implied changes in enrollment are similar to those implied by a $1,000 change in tuition, which is worth less than $5,000 for the typical college graduate who stays in college for less than 5 years. Dropout risk is key for understanding this result. While the tuition change affects the incentives for all students, the college premium is mainly relevant for high ability students who expect to graduate from college.

## 5.2 Relaxing Borrowing Limits

A large literature investigates whether borrowing constraints prevent a sizable number of students from attempting college. To address this question in our model, we recompute individual school choices when borrowing limits are increased 2-fold. All other model parameters remain unchanged. We find that, even though most college students never approach

Table 11: Increased Borrowing Limits

| | School group | | |
|---|---|---|---|
| | HS | CD | CG |
| Fraction | | | |
| Baseline | 0.47 | 0.24 | 0.29 |
| Double borrowing limits | 0.40 | 0.24 | 0.35 |
| Mean log ability | | | |
| Baseline | -0.51 | -0.10 | 0.91 |
| Double borrowing limits | -0.64 | -0.22 | 0.88 |

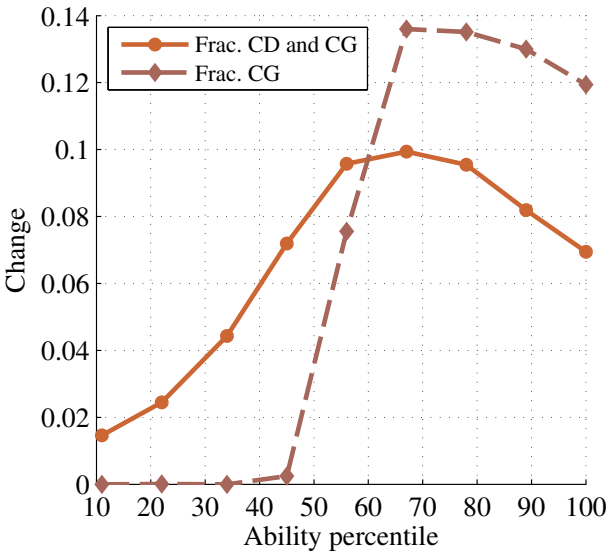
their borrowing limits (see Table 8), relaxing borrowing constraints has substantial effects on college entry and graduation rates.

Table 11 summarizes the resulting changes in schooling and abilities. The fraction of high school graduates who attempt college rises from 53% to 60%. Conditional on entering college, the graduation rate improves. Increased schooling reduces the mean log abilities of all school groups, but less so among college graduates, leading to a modest increase in the college lifetime earnings premium.

Figure 10 studies these changes in more detail. It shows that college entry rates increase mostly among median to high ability students. Borrowing constraints are less important for low ability students, who expect to drop out of college regardless of their financial assets. These students enter college only if college costs are low, in which case consumption can be financed without large debts.

The vast majority of the additional low ability students that attempt college fails to graduate. By contrast, graduation rates, conditional on college entry, increase among high ability students. The reason is that relaxed borrowing constraints allow students to stay in college longer, which reduces dropout rates during the early college years. This enables more high ability students to earn enough credits to graduate. Eventually, though, low ability students fail to earn enough credits, which forces them to drop out.

Figure 10: Increased Borrowing Limits




Notes: The figure shows the change in the fraction of persons who attempt college and who graduate from college for each ability level relative to the baseline model.

## 5.3 Dual Enrollment Programs

Our model suggests that the efficiency of school selection could be increased by providing students with information about their college aptitudes. In practice, this idea is implemented in the form of *dual enrollment programs* where high school students take courses at colleges or universities and receive both high school and college credit. In the 2010-11 school year, more than one million U.S. students participated in such programs (Stephanie Marken, Lewis, and Ralph, 2013).

We study the effect of such programs in our model by allowing each high school graduate to take 2 college courses before deciding whether to enter college. This amounts to 40% of an annual course load.

Figure 11 shows that the effects of such a program are small. Panel (a) shows the probability that a person of given ability enters college. The probability declines slightly for low ability students and rises slightly for high ability students, reflecting the higher precision of the students' beliefs at the time the college entry decision is made. The net effect on college enrollment is very small. Panel (b) shows that the fraction of students who earn a college degree increases slightly.

Figure 11: Dual Enrollment Program

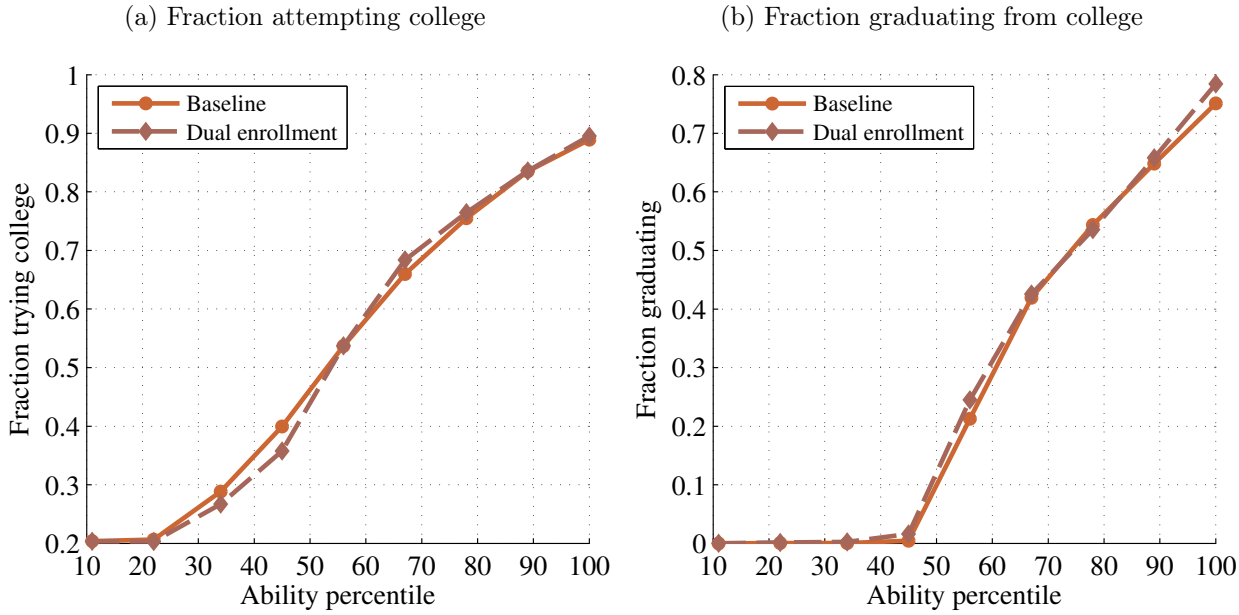
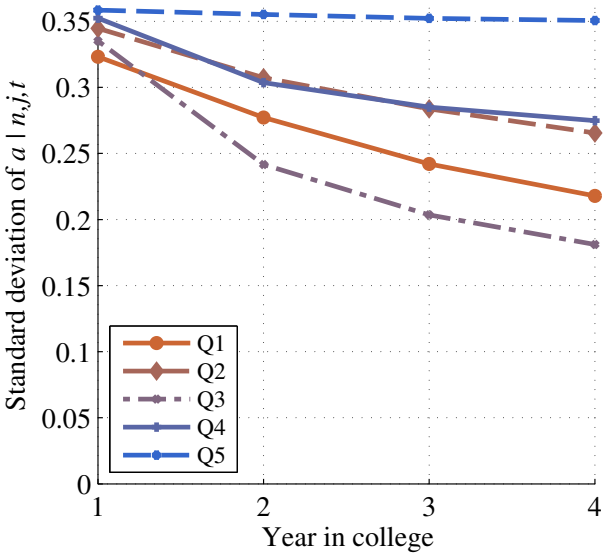



Figure 12 reveals why the changes are so small. It shows how uncertainty about student abilities evolves as students work their way through college. Each line represents a signal quintile. Each point shows the standard deviation of abilities, given the students' information sets $(n, j, t)$. This is averaged over students, using the mass of students in college by $(n, j, t)$ as weights.

Two main observations stand out. (i) At college entry, students face considerable uncertainty about their abilities. The standard deviation of abilities is above 0.3 for all signal quintiles, compared with an unconditional standard deviation of 1. (ii) For most students, the rate of learning is slow. At the start of the second year in college, the standard deviation is still above 0.25 for all but one of the quintiles. The rate of learning is particularly slow for students with very high abilities who pass nearly all of their courses. The additional information provided by 40% of one year's course load therefore has a small effect on most students' beliefs.

Another reason why this information does not alter college entry and exit decisions much is that these decisions often do not depend on the exact ability level. Students in the highest ability quartile know that they will graduate with near certainty, if they stay in college for $T_c$ periods. Students in the lowest ability quartile, on the other hand, know that they will almost surely not graduate (see Figure 5). For these students, information about their

Figure 12: Evolution of Ability Uncertainty




Notes: The figure shows the standard deviation of abilities conditional on the student's information set $(n, j, t)$. Each line represents an ability signal quintile.

abilities may have little value.

# 6 Conclusion

We conclude by considering potential avenues for future research. A key challenge is the identification of school sorting by ability and of human capital production in college. Modeling two additional decisions that can be observed for college students could help with this identification problem.

The first decision is the allocation of time between study effort, work, and leisure. In the data, college students spend little time attending classes and studying, while at the same time working for low wages (Babcock and Marks, 2010). This suggests a low marginal value of study effort and could be used to help identify human capital production in college, modeled here as the contribution of earned credits on lifetime earnings.

The second decision is the quality and cost of the college attended. Admission to better colleges may account for part of the higher returns to college enjoyed by high ability students (Dale and Krueger, 2002; Hoekstra, 2009). Observing how wages vary with test scores and

college qualities may help disentangle the effects of ability selection and human capital production. Allowing students to choose work hours and college costs could also soften the financial constraints that prevent some students in our model from earning a college degree.

Additional progress towards identification could be made by modeling the work phase in more detail. How wage dispersion changes with age contains information about the joint distribution of abilities and human capital endowments (see Huggett, Ventura, and Yaron 2006, 2011). This idea is pursued in Hendricks (2013) in the context of a stochastic Ben-Porath model that abstracts from college completion risk.

# References


Akyol, A., and K. Athreya (2005): "Risky higher education and subsidies," *Journal of Economic Dynamics and Control*, 29(6), 979 – 1023. 6, 12

Altonji, J. G. (1993): "The Demand for and Return to Education When Education Outcomes are Uncertain," *Journal of Labor Economics*, 11(1), pp. 48–83. 6

Arcidiacono, P. (2004): "Ability sorting and the returns to college major," *Journal of Econometrics*, 121(1), 343–375. 9

Arcidiacono, P., E. Aucejo, A. Maurel, and T. Ransom (2012): "College attrition and the dynamics of information revelation," Mimeo. Duke University. 3

Arcidiacono, P., P. Bayer, and A. Hizmo (2008): "Beyond signaling and human capital: Education and the revelation of ability," Discussion paper, National Bureau of Economic Research. 13

Arcidiacono, P., and P. B. Ellickson (2011): "Practical methods for estimation of dynamic discrete choice models," *Annu. Rev. Econ.*, 3(1), 363–394. 11

Babcock, P., and M. S. Marks (2010): "The Falling Time Cost of College: Evidence from Half a Century of Time Use Data," *The Review of Economics and Statistics*, (forthcoming). 32, 43

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. 13

Belzil, C., and J. Hansen (2002): "Unobserved Ability and the Return to Schooling," *Econometrica*, 70(5), 2075–2091. 3

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75(4), pp. 352–365. 9

Borghans, L., B. H. H. Golsteyn, J. Heckman, and J. E. Humphries (2011): "Identification problems in personality psychology," *Personality and Individual Differences*. 16

Bound, J., M. F. Lovenheim, and S. Turner (2010): "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources," *American Economic Journal: Applied Economics*, 2(3), 129–57. 2, 15, 21

Bound, J., and S. Turner (2011): "Dropouts and diplomas: The divergence in collegiate outcomes," *Handbook of the Economics of Education*, 4, 573–613. 3

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. 17, 19

Bureau of Labor Statistics; US Department of Labor (2002): "National Longitudinal Survey of Youth 1979 cohort," Produced and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH. 13

Cameron, S. V., and C. Taber (2004): "Estimation of Educational Borrowing Constraints Using Returns to Schooling," *The Journal of Political Economy*, 112(1), pp. 132–182. 13

Card, D. (1999): "The causal effect of education on earnings," vol. 3, Part A of *Handbook of Labor Economics*, chap. 30, pp. 1801 – 1863. Elsevier. 5

——— (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69(5), pp. 1127–1160. 28

Caucutt, E. M., and K. B. Kumar (2003): "Higher education subsidies and heterogeneity: a dynamic analysis," *Journal of Economic Dynamics and Control*, 27(8), 1459–1502. 6, 12

Chatterjee, S., and F. Ionescu (2010): "Insuring Student Loans Against the Risk of College Failure," Mimeo. Federal Reserve Bank of Philadelphia. 6

Cunha, F., J. J. Heckman, and S. Navarro (2005): "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers*, 57, 191–261. 4, 6

Dale, S. B., and A. B. Krueger (2002): "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *The Quarterly Journal of Economics*, 117(4), pp. 1491–1527. 43

Dynarski, S. M. (2003): "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion," *American Economic Review*, 93(1), 279–288. 38

Eckstein, Z., and K. I. Wolpin (1999): "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities," *Econometrica*, 67(6), 1295–1339. 6

Flynn, J. R. (2009): *What Is Intelligence? Beyond the Flynn Effect*. Cambridge University Press. 16

Garriga, C., and M. P. Keightley (2007): "A General Equilibrium Theory of College with Education Subsidies, In-School labor Supply, and Borrowing Constraints," Mimeo. Federal Reserve Bank of St Louis. 3, 6, 7, 12

Heckman, J. J., L. Lochner, and C. Taber (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1(1), 1–58. 4, 6

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. 13

Heckman, J. J., and S. Urzua (2008): "The Option Value of Educational Choices and the Rate The Option Value of Educational Choices and the Rate of Return to Educational Choices," Manuscript. University of Chicago. 6

Hendricks, L. (2013): "Accounting for the Evolution of U.S. Wage Inequality," Manuscript. University of North Carolina. 44

Hendricks, L., and T. Schoellman (2011): "Student Abilities During the Expansion of U.S. Education, 1950-2000," Mimeo. University of North Carolina. 13, 19, 58

Hoekstra, M. (2009): "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," *Review of Economics and Statistics*, 91(4), 717–724. 43

Huggett, M., G. Ventura, and A. Yaron (2006): "Human Capital and Earnings Distribution Dynamics," *Journal of Monetary Economics*, 53(2), 265–290. 44

——— (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101, 2923–54. 44

Jaeger, D. A., and M. E. Page (1996): "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education," *The Review of Economics and Statistics*, 78(4), 733–740. 2

Johnson, M. T. (2010): "Borrowing Constraints, College Enrollment, and Delayed Entry," Mimeo. Yale University. 18

Keane, M. P., and K. I. Wolpin (1997): "The career decisions of young men," *Journal of Political Economy*, 105(3), 473–522. 3, 6

——— (2001): "The effect of parental transfers and borrowing constraints on educational attainment," *International Economic Review*, 42(4), 1051–1103. 3

King, M., S. Ruggles, J. T. Alexander, S. Flood, K. Genadek, M. B. Schroeder, B. Trampe, and R. Vick (2010): "Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]," Minneapolis: University of Minnesota. 14, 50

MaCurdy, T., T. Mroz, and R. M. Gritz (1998): "An Evaluation of the National Longitudinal Survey on Youth," *The Journal of Human Resources*, 33(2), pp. 345–436. 51

Manski, C. F. (1989): "Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon," *Economics of Education Review*, 8(4), 305–312. 3, 12

Manski, C. F., and D. A. Wise (1983): *College choice in America*. Harvard University Press. 3

McGrattan, E. R., and E. C. Prescott (2000): "Is the Stock Market Overvalued?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 24(4). 17

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. 6

NLS User Services (1992): "NLS79 Profiles of American Youth. Addendum to Attachment 106," http://www.nlsinfo.org/ordering/display_db.php3. 16

Oreopoulos, P., and U. Petronijevic (2013): "Making College Worth It: A Review of Research on the Returns to Higher Education," Working Paper 19053, National Bureau of Economic Research. 2

Pryor, J. H., S. Hurtado, V. B. Saenz, and J. L. Santos (2007): *The American freshman: Forty year trends*. Los Angeles: Higher Education Research Institute. 3

Rust, J. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica: Journal of the Econometric Society*, pp. 999–1033. 11

Stange, K. M. (2012): "An Empirical Investigation of the Option Value of College Enrollment," *American Economic Journal: Applied Economics*, 4(1), 49–84. 3, 6

Stephanie Marken, L. G., L. Lewis, and J. Ralph (2013): "Dual Enrollment Programs and Courses for High School Students at Postsecondary Institutions: 2010-11," Discussion paper, National Center for Education Statistics. 41

Stinebrickner, R., and T. R. Stinebrickner (2003): "Understanding Educational Outcomes of Students from Low-Income Families: Evidence from a Liberal Arts College with a Full Tuition Subsidy Program," *The Journal of Human Resources*, 38(3), pp. 591–617. 3

Stinebrickner, T. R., and R. Stinebrickner (2012): "Learning about academic ability and the college drop-out decision," *Journal of Labor Economics*, 30(4), 707–748. 13, 36

Trachter, N. (2012): "Stepping Stone and Option Value in a Model of Postsecondary Education," Mimeo. Einaudi Institute for Economics and Finance. 6

United States Department of Education. National Center for Education Statistics (1988): "High School and Beyond, 1980: Sophomore and Senior Cohort Third Follow-up (1986)," . 55

Willis, R. J., and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87(5), pp. S7–S36. 5

# Online Appendix

# A CPS Data

## A.1 Sample

In our main source of wage data, the NLSY79, persons are observed only up to around age 45. We use data from the March Current Population Survey (King, Ruggles, Alexander, Flood, Genadek, Schroeder, Trampe, and Vick, 2010) to extend the NLSY79 wage profiles to older ages. Our sample contains men between the ages of 18 and 75 observed in the $1964 - 2010$ waves of the CPS. We drop persons who live in group quarters or who fail to report wage income.

## A.2 Schooling Variables

Schooling is inconsistently coded across surveys. Prior to 1992, we have information about completed years of schooling (variable higrade). During this period, we define high school graduates as those completing 12 years of schooling (higrade=150), college dropouts as those with less than four years of college (151,...,181), and college graduates as those with 16+ years of schooling (190 and above). Beginning in 1992, the CPS reports education according to the highest degree attained (educ99). For this period, we define high school graduates as those with a high school 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 a bachelors, masters, professional, or doctorate degree (14,...,17).

## A.3 Age Earnings Profiles

Our goal is to estimate the age profile of mean log earnings for each school group. This profile is used to fill in missing earnings observations in the NLSY79 sample and to estimate individual lifetime earnings.

First, we compute the fraction of persons earning more than $\$2,000$ in year 2000 prices for each age $t$ within school group $s$, $f(t|s)$. This is calculated by simple averaging across all

years. For the cohorts covered by the NLSY79, the fractions are similar to their NLSY79 counterparts.

Next, we estimate the age profile of mean log earnings for those earnings more than $\$2,000$ per year, which we assume to be the same for all cohorts, except for its intercept. To do so, we compute mean log earnings above $\$2,000$ for every [age, school group, year] cell. We then regress, separately for each school group, mean log earnings in each cell on age dummies, birth year dummies, and on the unemployment rate, which absorbs year effects. We retain the birth cohorts $1935-1980$. We use weighted least squares to account for the different number of observations in each cell.

Finally, we estimate the mean earnings at age $t$ for the 1960 birth cohort as:

$$g_{CPS}(t|s) = \exp\left(1960 \text{ cohort dummy} + \text{age dummy}(t) + \text{year effect}(1960+t)\right) f(t|s) \quad (17)$$

For years after 2010, we impose the average year effect. Figure 13 shows the fitted age profiles together with the actual age profiles for the 1960 birth cohorts calculated from the CPS and the NLSY79. We find substantially faster earnings growth in the NLSY79 data compared with the CPS data. The discrepancies are modest until around age 30 (year 1990), which is consistent with the validation study by MaCurdy, Mroz, and Gritz (1998). The reason for the discrepancies is not known to us.

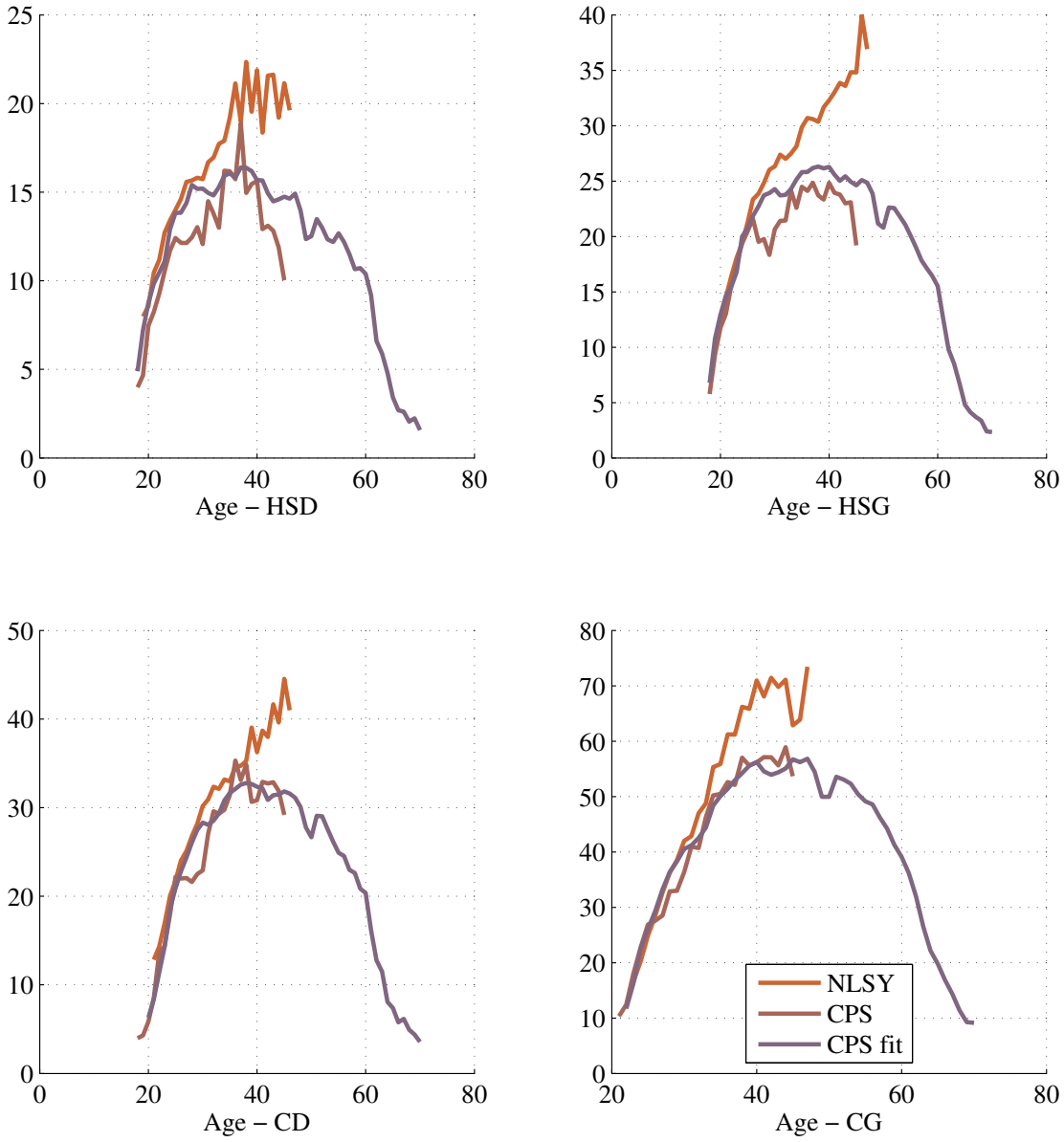
# B NLSY79 Data

The NSLY79 sample covers men born between 1957 and 1964 who earned at least a high school diploma. We use the $1979-2006$ waves. We drop persons who were not interviewed in 1988 or 1989 when retrospective schooling information was collected. We also drop persons who did not participate in the AFQT (about 6% of the sample). Table 12 shows summary statistics for this sample.

## B.1 Schooling Variables

For each person, we record all degrees and the dates they were earned. At each interview, persons report their school enrollments since the last interview. We use this information to determine whether a person attended school in each year and which grade was attended. For

Figure 13: Age-earnings Profiles




Notes: The figures show the exponential of mean log earnings by schooling and age in thousands of year 2000 dollars. Earnings are adjusted for the fraction of persons working at each age as described in the text.

Table 12: Summary Statistics for the NLSY79 Sample

| | HSG | CD | CG | All |
|---|---|---|---|---|
| Fraction | 46.6 | 25.3 | 28.1 | 100.0 |
| Avg. schooling | 12.1 | 14.1 | 17.0 | 14.0 |
| Range | 9 - 13 | 13 - 20 | 12 - 20 | 9 - 20 |
| AFQT percentile | 34.3 | 51.3 | 75.0 | 50.0 |
| $N$ | 1,447 | 800 | 675 | 2,922 |

Notes: For each school group, the table shows the fraction of persons achieving each school level, average years of schooling and the range of years of schooling, the mean AFQT percentile, and the number of observations.

persons who were not interviewed in consecutive years, it may not be possible to determine their enrollment status in certain years.

Visual inspection of individual enrollment histories suggests that the enrollment reports contain a significant number of errors. It is not uncommon for persons to report that the highest degree ever attended declined over time. A significant number of persons reports high school diplomas with only 9 or 10 years of schooling. We address these issues in a number of ways. We ignore the monthly enrollment histories, which appear very noisy. We drop single year enrollments observed after a person's last degree. We also correct a number of implausible reports where a person's enrollment history contains obvious outliers, such as single year jumps in the highest grade attained. We treat all reported degrees as valid, even if years of schooling appear low.

Many persons report schooling late in life after long spells without enrollment. Since our model does not permit individuals to return to school after starting to work, we ignore late school enrollments in the data. We define the start of work as the first 5-year spell without school enrollment. For persons who report their last of schooling before 1978, we treat 1978 as the first year of work. We assign each person the highest degree earned and the highest grade attended at the time he starts working. Persons who attended at least grade 13 but report no bachelor's degree are counted as college dropouts. Persons who report 13 years of schooling but fewer than 10 credit hours are counted as high school graduates. The resulting school fractions are close to those obtained from the High School & Beyond

sample.

## B.2 Lifetime Earnings

Lifetime earnings are defined as the present value of earnings up to age 70, discounted to age 19. Our measure of labor earnings consists of wage and salary income and 2/3 of business income. We assume that earnings are zero before age 19 for high school graduates, before age 21 for college dropouts, and before age 23 for college graduates.

Since we observe persons at most until age 48, we need to impute earnings later in life. For this purpose, we use the age earnings profiles we estimate from the CPS (see Appendix A). The present value of lifetime earnings for the average CPS person is given by $Y_{CPS}(s) = \sum_{t=19}^{70} g_{CPS}(t|s)R^{19-t}$. The fraction of lifetime earnings typically earned at age $t$ is given by $g_{CPS}(t|s)R^{19-t}/Y_{CPS}(s)$.

For each person in the NLSY79 we compute the present value of earnings received at all ages with valid earnings observations. We impute lifetime earnings by dividing this present value by the fraction of lifetime earnings earned at the observed ages according to the CPS age profile, $g_{CPS}(t|s)R^{19-t}/Y_{CPS}(s)$.

An example may help the reader understand this approach. Suppose we observe a high school graduate with complete earnings observations between the ages of 19 and 40. We compute the present value of these earnings reports, including years with zero earnings, $X$. According to our CPS estimates, 60% of lifetime earnings are received by age 40. Hence we impute lifetime earnings of $X/0.6$.

In order to limit measurement error, we drop individuals who report zero earnings for more than 30% of the observed years. We also drop persons with fewer than 5 earnings observations after age 35 or whose reported earnings account for less than 30% of lifetime earnings according to the CPS profile. Table 13 shows summary statistics for the persons for which we can estimate lifetime earnings. One concern is that the NLSY79 earnings histories are truncated around age 45, which leaves 20 to 30 years of earnings to be imputed. Fortunately, the fitted CPS age profiles imply that around 70% of lifetime earnings are earned before age 45.

Table 13: Lifetime Earnings

| | HSG | CD | CG |
|---|---|---|---|
| exp(mean log) | 600,061 | 643,153 | 944,269 |
| Standard deviation (log) | 0.51 | 0.55 | 0.50 |
| $N$ | 578 | 343 | 319 |

Notes: The table show exp(mean log lifetime earnings), the standard deviation of log lifetime earnings, and the number of observations in each school group.

# C NELS Data

We obtain data on the academic performance of college students and on their incomes and expenditures from data collected by the National Education Longitudinal Studies (NELS) program of the National Center for Education Statistics (NCES). The High School & Beyond (HS&B) survey covers the 1980 senior and sophomore classes (see United States Department of Education. National Center for Education Statistics 1988). Both cohorts were surveyed every two years through 1986. The 1980 sophomore class was also surveyed in 1992, at which point postsecondary transcripts from all institutions attended since high school graduation were collected under the initiative of the Postsecondary Education Transcript Study (PETS).[25] We restrict attention to white male sophomores surveyed at least through 1986, which leaves us with 14,825 student records and 17,363 transcripts collected from 4,079 institutions.

## C.1 Enrollment and Dropout Statistics

The sample contains 3,671 students who graduated from high school in 1982. We split these students into quartiles according to their high school GPA, which is available for 92% of our sample. For the remaining 8%, we impute high school GPA quartile with the quartile of their cognitive test score. This test was conducted in their senior year and was designed to measure quantitative and verbal abilities.

[25] PETS data files were obtained through a restricted license granted by the National Center for Education Statistics.

Table 14: School Attainment of College Entrants

| Quartile of HS GPA | | 1 | 2 | 3 | 4 | All | N |
|---|---|---|---|---|---|---|---|
| Fraction graduating | | 0.13 | 0.31 | 0.58 | 0.76 | 0.55 | 840 |
| Fraction dropping out, | year 1 | 0.40 | 0.25 | 0.13 | 0.07 | 0.16 | 200 |
| | year 2 | 0.24 | 0.21 | 0.16 | 0.07 | 0.14 | 191 |
| | year 3 | 0.15 | 0.11 | 0.06 | 0.05 | 0.08 | 108 |
| | year 4 | 0.06 | 0.08 | 0.04 | 0.03 | 0.05 | 52 |
| | year 5 | 0.02 | 0.01 | 0.03 | 0.01 | 0.02 | 30 |

Notes: The table shows the fraction of college entrants in each high school GPA quartile that drops out of college at the end of each year. $N$ is the number of observations.

Using PETS transcript data, we count the number of credits each student attempts and completes in each year in college. Credits are defined as follows. We count withdrawals that appear on transcripts as attempted but unearned credits. We drop transfer credits to avoid double counting. We drop credits earned at vocational schools, such as police academies or health occupation schools.

We truncate each student's college history when he earns his first bachelor's degree or within 5 years of his first year without college enrollment. We count a student as entering college if he attempts at least 9 credits in a given academic year. Using this definition, 53% of the cohort enters college immediately upon high school graduation. Another 3% of the cohort enter in the following year. 55% of immediate entrants obtain a bachelor's degree. Students that earn bachelor's degrees later than 5 years after their first break in enrollment are dropped from the sample.

For each high school GPA quartile, Table 14 shows the fraction of college entrants who graduate from college and who drop out at the end of each year. These statistics are computed from 1,436 college entrants with complete transcript histories. We refer to a college entrant as a year $x$ dropout if he/she enrolled continuously in years 1 through $x$, attempted fewer than 7 credits in year $x+1$, and failed to obtain a bachelor degree within 5 years. 86% of the college graduates in our sample are enrolled continuously until graduation.

Table 15: Financial Resources

| | year 1 | year 2 | year 3 | year 4 |
|---|---|---|---|---|
| q | -870 | -1,168 | 833 | 2,422 |
| Tuition | 4,428 | 9,276 | 16,897 | 24,632 |
| Grants | 1,308 | 2,619 | 4,119 | 5,682 |
| Parental transfers | 6,179 | 12,463 | 19,137 | 26,428 |
| Earnings | 5,013 | 9,505 | 15,132 | 20,247 |
| Loans | 983 | 2,184 | 3,416 | 4,670 |
| Fraction in debt | 0.28 | 0.36 | 0.42 | 0.48 |
| N | 1436 | 1236 | 1045 | 934 |

Notes: Dollar amounts are cumulative and in year 2000 prices. Average amounts include zeros.

## C.2 Financial Variables

In the second and third follow-up interviews (1984 and 1986), all students reported their education expenses, various sources of financial support, and their earnings. Table 15 shows the means of all financial variables for students who are enrolled in college in a given year.

We construct total parental transfers as the sum of school-related and direct transfers to the student. The school-related transfer refers to "payments on [the student's] behalf for tuition, fees, transportation, room and board, living expenses and other school-related expenses." It is available only for the first two academic years after high school graduation. For academic years 84/85 and 85/86, we proxy it with the 83/84 value, adjusted by the change in tuition net of grants and scholarships.

Direct transfers include in-kind support, such as room and board, use of car, medical expenses and insurance, clothing, and any other cash or gifts. We set the transfer values to the midpoints of the intervals they are reported in. For the highest interval, more than \$3,000 in current prices, we assign a value of \$4,000. We assume that half of the transfer is paid out in each semester of the calendar year for which the transfer is reported.

Tuition and fees and the value of grants are available for each academic year. Grants refer to the total dollar value of the amount received from scholarships, fellowships, grants, or other benefits (not loans) during the academic year.

Student earnings are available at calendar year frequencies. To convert these into academic

years for college students, we assume that the relative fractions of year $j$ earnings that are earned in academic years $ij$ and $jk$ are inversely related to the relative number of credits taken in years $ij$ and $jk$. This implies that the proportion of year $j$ earnings attributed to academic year $ij$ is given by $credits_{jk}/(credits_{ij} + credits_{jk})$. We attribute half of the 1982 earnings to the 82/83 academic year.

# D Identification

This section investigates the identification of selected model parameters. Since our model is computationally efficient, we are able to compute how the model fit changes as each parameter's value varies over a grid. For each grid point, we recalibrate all other model parameters.

We focus on parameters that we expect to be important for ability selection: the dispersion of abilities $\phi_{HS}$, the precision of ability signals governed by $\alpha_{a,m}$, the effect of college credits on earnings $\mu$, and the scale of preference shocks $\pi$. To conserve space, we summarize the results without presenting the details for each case.

**Ability dispersion.** The value of $\phi_{HS}$ is mainly identified by the relationship between lifetime earnings and test scores. Increasing $\phi_{HS}$ strengthens this relationship (the model implies too steep lines in Figure 2). The intuition is similar to Hendricks and Schoellman (2011): For given test score precision ($\alpha_{a,m}$ and $\alpha_{IQ,a}$), a higher values of $\phi_{HS}$ implies that a one standard deviation increase in test scores is associated with a larger increase in abilities and thus lifetime earnings. Fixing $\phi_{HS}$ below its calibrated value also implies a model college premium that is smaller than the calibration target. The difference between $\phi_{CG}$ and $\phi_{HS}$ allows the model to better match how lifetime earnings vary with test scores for all school groups.

Higher values of $\phi_{HS}$ increase the role of ability selection for the college premium. The reason is purely mechanical: for given school sorting, abilities account for a larger share in earnings variation across individuals.

**Signal noise.** The precision of the ability signal ($\alpha_{a,m}$) is mainly identified by the timing of college dropouts. Since the ability signal in the baseline model is quite high, we focus

on the implications of lower precision (more signal noise). The model then implies that too few students drop out of college, especially during the early years. The intuition is that the option of staying in college becomes more valuable. Students who perform poorly in their first year of college still have a chance to graduate and are less likely to drop out.

With large signal noise, the model also has trouble matching credit passing rates. The credit passing rates of dropouts and graduates are too similar at the start of college, but diverge too much later on. The reason is that large signal noise increases the role of luck for college completion. This reduces the gap in credit passing rates between graduates and dropouts early on. During later years in college, the gap opens up as students with persistently poor academic records update their beliefs about their graduation prospects and drop out. By contrast, in the data, the gap in the credit passing rates of college graduates and dropouts is roughly constant across years. With precise signals, the model replicates this fact. With imprecise signals it does not.

Since more signal noise prevents effective school sorting by abilities, it reduces the role of ability selection for the college premium.

**Effect of credits on earnings.** The value of $\mu$ is mainly identified by the relationship between lifetime earnings and test scores and by the timing of dropouts.

Holding other parameters constant, higher values of $\mu$ increase the relative earnings of college dropouts and college graduates. The calibration offsets this by reducing ability dispersion ($\theta_s$). The alternative would be to reduce $Y_{CG}$, but this does not lower the college dropout premium. It would also violate the constraint $Y_{CG} \geq Y_{HS} = Y_{CD}$ and imply that graduating from college reduces earnings for low ability students. The lower ability dispersion flattens the relationship between test scores and lifetime earnings (Figure 2). It also reduces the contribution of selection to the college premium.

Higher values of $\mu$ increase the incentives to stay in college longer, even for students who expect not to graduate. This leads college dropouts to stay in college longer than in the data (Figure 3).

**Preference shocks.** When preference shocks are smaller than in the baseline case, the model fails to account for the timing of dropout decisions. Too many low ability students stay in college until $T_c$. The intuition is that students with low $q$ and high $k_1$ have no reason

to drop out. They know from the outset that they will not graduate. For these students, college is mainly a consumption good. However, their dropout decisions are sensitive to shocks because the financial stakes are so small. Preference shocks prevent these students from staying in college too long.

Larger preference shocks weaken the association between college costs and college attendance. As a result, the model greatly overstates the rise in debt as students go through college. The baseline model avoids this because college students select more strongly on college costs. The scale of preference shocks is not important for ability selection.

# E Robustness

This section investigates the robustness of our main finding. Table 16 shows how the importance of ability selection varies with the value of selected parameters.

To illustrate how to read the Table, consider the first row. It varies $\phi_{HS}$ over a grid of values that range from 0.1 to 0.25, compared with a baseline value of 0.15. The model is recalibrated, fixing $\phi_{HS}$ at each grid point. The "selection" column reports the smallest and the largest fraction of the mean log lifetime earnings gap between college graduates and high school graduates that is due to ability selection. This is defined as in Section 4.1. The remaining rows of Table 16 vary the values of $\mu$, $\alpha_{a,m}$, and $\pi$ in similar ways.[26]

Across all parameter values covered in Table 16, we find that ability selection accounts for more than one-third of the college lifetime earnings premium. We have also experimented with alternative ways of constructing the calibration targets and with restricted or extended models. For example, we shut down the non-pecuniary schooling costs $U(s)$, and we allowed for a direct consumption utility of being in college. In all cases, we found our main result to be highly robust.

[26] We tried wider ranges for some of the parameters, but found that the model fit deteriorates dramatically.

Table 16: Robustness

| Variable | Base value | Range | Selection |
|---|---|---|---|
| $\phi_{HS}$ | 0.153 | 0.10–0.25 | 37.1–92.8 |
| $\mu$ | 0.014 | 0.005–0.020 | 66.4–36.6 |
| $\alpha_{a,m}$ | 2.971 | 1.00–3.00 | 51.0–50.4 |
| $\pi$ | 0.767 | 0.40–2.00 | 56.6–49.0 |

Notes: Each row varies one parameter over a grid of values. "Variable" indicates which parameter is varied. "Base value" shows the parameter's value in the baseline model. "Range" shows the range over which the parameter is varied. "Selection" shows the fraction of the mean log lifetime earnings gap between college graduates and high school graduates that is due to ability selection.