The Return to College: Selection Bias and Dropout Risk*

Lutz Hendricks† Oksana Leukhina‡

May 16, 2013

Preliminary.

Abstract

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

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

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

*For helpful comments we thank V. V. Chari, Patrick Kehoe, Christoph Winter as well as seminar participants at Indiana University, University of Washington, Ohio State, the Federal Reserve Bank of Minneapolis, the 2011 Midwest Macro Meetings, the 2011 SED Meetings, the 2011 Cologne Macro Workshop, and the 2011 CASEE Human Capital Conference. The usual disclaimer applies.
†University of North Carolina, Chapel Hill; lutz@lhendricks.org
‡University of Washington; oml@u.washington.edu
48% of the measured lifetime earnings gap between college graduates and high school graduates.


Key words: Education. College dropout risk.
1 Introduction

A large literature, surveyed by Card (1999), has investigated the causal effect of schooling on earnings. In U.S. data college graduates earn substantially more than high school graduates. However, part of this differential may be due to selection. 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 recognizes that not all students are able to complete the coursework required for graduation so that college completion is uncertain. We argue that this uncertainty affects the self-selection of college students and may help identify its contribution to the college premium.

Our approach is motivated by the data summarized in Table 1. We consider men observed in the NLSY79 starting at high school graduation around 1980 (Appendix B describes the data in detail). We divide persons into three groups according to their highest educational attainment: high school graduates (HSG), college dropouts (CD), and college graduates (CG). For each group we calculate average lifetime earnings in year 2000 prices and average cognitive test scores as measured by the AFQT. We highlight 4 observations:

1. 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 BA degree.\(^1\)

3. Even so, nearly half of those who start college fail to attain a BA degree, suggesting that dropout risk may be important for understanding the incentives for attending college.

4. College graduates have substantially higher cognitive test scores than do high school graduates, suggesting that ability differences may account for part of the college earnings premium.

Motivated by similar observations, a large literature has attempted to decompose the observed premium into selection and return to college. We discuss this literature in more detail in Section 1.1. Our contribution relative to this literature is to jointly model ability selection and college completion risk.

\(^1\)For evidence on sheepskin effects see Jaeger and Page (1996).
Table 1: Schooling and Lifetime Earnings

<table>
<thead>
<tr>
<th></th>
<th>HSG</th>
<th>CD</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean lifetime earnings</td>
<td>674</td>
<td>742</td>
<td>1068</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>316</td>
<td>408</td>
<td>541</td>
</tr>
<tr>
<td>Mean AFQT percentile</td>
<td>0.34</td>
<td>0.52</td>
<td>0.74</td>
</tr>
<tr>
<td>Fraction</td>
<td>0.45</td>
<td>0.26</td>
<td>0.29</td>
</tr>
<tr>
<td>Number of observations</td>
<td>578</td>
<td>343</td>
<td>319</td>
</tr>
</tbody>
</table>

Notes: Lifetime earnings are defined as the present value of earnings, discounted to age 19. AFQT stands for Armed Forces Qualification Test.

**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 (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 graduating 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 in the 1960 birth cohort. 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. In the main specification, 48% of the measured lifetime earnings gap between college graduates and high school graduates is due to ability selection.

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) 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.

Note that our notion of college completion risk differs from that of Keane and Wolpin (1997), Keane and Wolpin (2001), and others following their approach. 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.

Our interpretation of dropping out as failure is consistent with survey evidence. The students surveyed by Stinebrickner and Stinebrickner (2009) cite poor grades as their main reason for dropping out. 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.

We highlight a number of additional findings:

1. Our model accounts for the fact that a significant number of students with low cognitive test scores attempt college, even though their graduation prospects are poor. For these students, college is mainly a consumption good.

2. Students of intermediate abilities typically drop out in response to poor academic performance, which leads them to update their beliefs about their abilities as suggested by Manski (1989). These students would like to graduate from college, but find it too costly and time consuming.

\[\text{See Examples include Stange (2012) and Arcidiacono, Aucejo, Maurel, and Ransom (2012).}\]
3. College graduation prospects vary strongly with ability. Students in the lowest ability decile have a less than 20% chance of graduating from college, while student in the top decile graduate with near certainty. The inability to attain a BA degree is the main friction that prevents low ability students from entering college.

4. Except for the highest ability students, the financial returns to college completion are quite modest. Since students also have the option of dropping out in the future, many students are nearly indifferent between entering college and working as a high school graduate. This feature allows our model to account for the large effects of tuition changes on college enrollment estimated in the literature, even though most students are not borrowing constrained (see Section 4.2).

Plausibility: We highlight one feature of the data that contributes to the large role our model attributes to ability selection. Consider the following simple model. Individual log lifetime earnings are a function of the school specific skill price $z_s$ and ability $a$:

$$w_s = z_s + a$$

where $s \in \{HSG, CD, CG\}$. Abilities are observed with noise $\epsilon$ via cognitive test scores: $AFQT = a + \epsilon$. Denote the difference between college graduates and high school graduates by $\Delta$. For example, $\Delta z = z_{CG} - z_{HSG}$. The measured college lifetime earnings premium is then given by

$$z_{CG} - z_{HS} + \mathbb{E}(a|CG) - \mathbb{E}(a|HS) = \Delta z + \Delta \mathbb{E}(a)$$

$$= \Delta z + \Delta \mathbb{E}(a|AFQT) + \Delta \mathbb{E}(a) - \Delta \mathbb{E}(a|AFQT)$$

The first term is observable as the lifetime earnings premium after conditioning on $AFQT$: $\Delta z + \Delta \mathbb{E}(a|AFQT) = \Delta \mathbb{E}(w|AFQT)$. Since $\mathbb{E}(a|s, AFQT) = AFQT - \mathbb{E}(\epsilon|s, AFQT)$, we have $\Delta \mathbb{E}(a|AFQT) = -\Delta \mathbb{E}(\epsilon|AFQT)$ and the ability gap between college graduates and high school graduates is given by

$$\Delta \mathbb{E}(a) = \Delta \mathbb{E}(w) - \Delta \mathbb{E}(w|AFQT) - \Delta \mathbb{E}(\epsilon|AFQT)$$

The first two terms in (4) can be estimated as the mean log lifetime earnings gap between college graduates and high school graduates with $[\Delta \mathbb{E}(w|AFQT)]$ and without $[\Delta \mathbb{E}(w)]$.

---

3We thank Patrick Kehoe for suggesting this argument. Its spirit is similar to the monotone instrumental variables methods proposed by Manski and Pepper (2000).
conditioning on AFQT scores. We argue that it is plausible to think of the third term \([\Delta E(\epsilon|AFQT)]\) as positive. In particular, consider the case where schooling is increasing in ability, so that all individuals with ability above some cutoff \(a_{CG}\) graduate from college while all individuals with ability below \(a_{HSG}\) become high school graduates. If we observe a set of persons with identical AFQT scores but different school choices, we may infer that \(E(\epsilon|s, AFQT)\) is higher for the less educated persons. For example, a high school graduate with median AFQT score would have a positive \(\epsilon\) (AFQT is high relative to \(a\)), while a college graduate with median AFQT would have a high \(\epsilon\) (AFQT is low relative to \(a\)). A similar argument remains true when schooling is determined by ability plus noise.

If \(\Delta E(\epsilon|AFQT) > 0\), we can bound the contribution of abilities to the college premium by

\[
\Delta E(a) > \Delta E(w) - \Delta E(w|AFQT)
\]

We implement this bounding argument in Table 2. Regressing lifetime earnings on school dummies (standing in for skill prices \(z_s\)) yields \(E(w)\) as the difference between the college graduate and the high school graduate dummies. The estimated value is 0.64. Adding standard normal AFQT scores to the regression yields \(\Delta E(w|AFQT) = 0.29\). The contribution of ability selection to the lifetime earnings gap is then at least 0.35. We conclude that in any model where lifetime earnings are given by (1) and where AFQT errors are decreasing in schooling, at least 55% of the observed lifetime earnings gaps are due to ability selection.

### 1.1 Related Literature

This paper relates to a vast literature that estimates returns to schooling. For surveys, we refer the reader to Card (1999) and Heckman, Lochner, and Todd (2006). One strand of this literature uses econometric approaches, such as instrumental variables, to control for
selection bias in wage regressions. These efforts abstract from degrees and treat schooling as a continuous variable and are therefore silent about the college premium and completion risk.\footnote{A seminal contribution is Willis and Rosen (1979). Card (1999) surveys this literature and discusses how its findings may be interpreted.}

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 risk.\footnote{Examples include Heckman, Lochner, and Taber (1998), Cunha, Heckman, and Navarro (2005), and Navarro (2008).}

Models with college completion risk have, for the most part, abstracted from heterogeneity in abilities that affect earnings. Examples include Akyol and Athreya (2005), Garriga and Keightley (2007), Chatterjee and Ionescu (2010), and Stange (2012).\footnote{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.} These models cannot address the question how ability selection affects measured college wage premiums.

Two recent papers develop models with college dropouts and ability heterogeneity. Belzil and Hansen (2002) treat schooling as a continuous variable. The puzzle is then why so many students stop schooling after exactly 12 or 16 years. Belzil and Hansen solve this problem by assuming that the disutility of schooling spikes in years 13 and 17. Trachter (2011) presents a model of risky college completion with ability heterogeneity. Since his model features only two ability types, it is of limited value for studying ability selection. We also calibrate our model using a richer set of empirical observations, in particular regarding the relationship between measured abilities, college choices and earnings.

Keane and Wolpin (1997) and Keane and Wolpin (2001) propose alternative reasons why college students may drop out of college. In their models, any student can earn a BA degrees by attending college for 4 years. Some students choose to drop out of college earlier because they receive shocks to either the utility or to the financial payoff derived from work or schooling. We abstract from preference shocks and focus instead on the idea that students ability affects the time required to complete a BA degree.
2 The Model

2.1 Model Outline

We study a partial equilibrium model of school choice. We follow a single cohort, starting at the date of high school graduation, through college (if chosen), work, and retirement. Figure 1 summarizes the agent’s life-cycle. At the start of age \( t = 1 \), all agents graduate from high school. They are endowed with a type \( j \) which determines initial assets \( k_1 \), an ability signal \( m \), and a net price of attending college \( q \). Each person also draws an ability 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.

Agents choose whether to start working right away as high school graduates or to attempt college. Agents are not allowed to return to school after they start working. Working agents choose a consumption path to maximize lifetime utility subject to a lifetime budget constraint that equates the present value of income to the present value of consumption spending.

While in college, students accumulate college credits \( n \). Once a student reaches \( n_{grad} \) credits he graduates and works as a college graduate. The accumulation of credits is stochastic. More able students accumulate credits faster. This part of the model borrows from Garriga and Keightley (2007).

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

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

2.2 Endowments

Agents enter the model at age 1 and live until age \( T \). At age 1, a person is endowed with

1. \( n_1 = 0 \) completed college credits;
Figure 1: Model Timing

(a) Choices at HS graduation

<table>
<thead>
<tr>
<th>Draw endowments: type $j = (k_1, m, q)$ and $n_1 = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose college consumption: $c_F$</td>
</tr>
</tbody>
</table>

- Work as HS graduate
  $V_W(0, j, HS, 1)$

- Try college
  $V_C(0, j, 1)$

- Learn ability $a$
  $V(k_1, a, 0, HS, 1)$

(b) Timing and choices while in college at age $t$

<table>
<thead>
<tr>
<th>$V_C(n_t, j, t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consume $c_F$. $k_{t+1} = Rk_t - c_F - q$</td>
</tr>
<tr>
<td>Draw $n_{t+1}$. Update beliefs about $a$</td>
</tr>
</tbody>
</table>

- Graduate
  $V_W(n_{t+1}, j, CG, t + 1)$

- Drop out
  $V_W(n_{t+1}, j, CD, t + 1)$

- Study in $t + 1$
  $V_C(n_{t+1}, j, t + 1)$

- Learn ability
  $V(k_{t+1}, a, n_{t+1}, CG, t + 1)$

- Learn ability
  $V(k_{t+1}, a, n_{t+1}, CD, t + 1)$
2. learning ability $a$;
3. type $j \in 1, \ldots, J$.

$a$ determines the person’s productivity in school and at work. Abilities are not observed by the agents until they start working. A person of type $j$ is endowed with:

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, parental transfers that are conditional on attending college, and other costs or payoffs associated with attending college.
3. Initial assets $\hat{k}_j$. 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 Preferences

Individuals consume two goods: a market good $c_F$ and a non-market good $c_L$. Expected utility is given by

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

where $c_t = C(c_{Ft}, c_{Lt})$ is a consumption aggregator.

2.4 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. We summarize the value of working by the value function $V_W(k_\tau, n_\tau, j, s, \tau)$ where $\tau$ is the age at which work starts, $k_\tau$ denotes the level of assets, $n_\tau$ is the number of completed college credits, $s$ is the level of completed schooling. $s$ takes on the values HS for high school graduates, CD for college dropouts and CG for college graduates. Working as a high school graduate yields value $V_{HS}(j) = V_W(\hat{k}_j, 0, j, HS, 1)$. Section 2.5 describes how $V_W$ is determined.
Consumption in college is chosen from a discrete set of values that are indexed by $j_c = 1, ..., N_c$. We summarize the value of starting year $t$ in college with consumption level $c_{F,j_c,j}$ by $V_C(n_t, j_c, j, t)$. Section 2.6 describes how $V_C$ is determined and explains why $(n_t, j_c, j, t)$ is the correct state vector. Consumption choice maximizes lifetime utility subject to preference shocks $p_{jc}$ that are drawn from independent type I extreme value distributions with scale parameter $\pi_c > 0$:

$$j_c = \arg \max V_C(0, j_c, j, 1) + \pi_c p_{jc}$$

(7)

The implied choice probabilities are given by

$$\Pr (j_c | j) = \frac{\exp(V_C(0, j_c, j, 1) / \pi_c)}{\sum_j \exp(V_C(0, j, j, 1) / \pi_c)}$$

(8)

For computational reasons we assume that consumption remains fixed while a student attends college. In practice, we set the allowed consumption levels such that $c_{F,j_c,j}$ exactly exhausts a type $j$ student’s borrowing limits after $j_c$ periods. 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.

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

$$\max \{ V_C(0, j_c, j, 1) + \pi_E p_c, V_{HS}(j) + \pi_E p_w \}$$

(9)

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

$$\Pr (\text{college} | j_c, j) = \frac{\exp(V_C(0, j_c, j, 1) / \pi_E)}{\exp(V_C(0, j_c, j, 1) / \pi_E) + \exp(V_{HS}(j) / \pi_E)}$$

(10)

### 2.5 Work

Upon completing schooling, the worker learns his ability. The worker’s value after learning $a$ is determined according to

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

subject to the budget constraint

$$e^{\phi_s(a-\underline{a}) + \mu n_\tau} Y(s) + Rk_\tau = \sum_{t=\tau}^T c_{Ft} R^{\tau-t}$$

(12)
where $\beta > 0$ is a discount factor. In each period, the agent purchases market goods $c_{Ft}$ at a price that is normalized to 1. He receives a fixed amount of the non-market good, $\hat{c}_L$, for free. We discuss the role of the non-market good in Section 2.8.

The budget constraint equates the present value of consumption spending to lifetime earnings plus the value of assets owned at age $\tau$. $R > 1$ is the gross interest rate and $e^{\phi_s(a-a) + \mu n_{\tau}} Y(s)$ denotes the present value of lifetime earnings, discounted to age $\tau$. A worker with ability $a = \underline{a}$ and no completed college credits earns $Y(s)$ which depends on completed schooling $s$. Each completed college credit increases lifetime earnings by a constant factor $e^\mu$. This may reflect human capital accumulation. The effect of ability on lifetime earnings ($\phi_s$) may depend on schooling. One possible reason is on-the-job learning. Even though $Y(s)$ does not depend on $\tau$, staying in school longer reduces lifetime earnings discounted to age 1 as earnings are shifted back in time.

To ensure that graduation from college increases lifetime earnings for workers of all abilities, we impose $Y(CG) \geq Y(CD)$ and $\phi(CG) \geq \phi(CD)$, where $\underline{a}$ is set to an arbitrary value such that $a - \underline{a} \geq 0 \forall a$. We also impose $Y(CD) = Y(HS)$ and $\phi(CD) = \phi(HS)$, so that college dropouts differ from high school graduates only in the number of credits earned. This restriction ensures that attending college for a single period without earning credits does not increase earnings simply by placing a “college” label on the worker. All high school graduates share $\tau = 1$ and $n_{\tau} = 0$. However, there is variation in both $\tau$ and $n_{\tau}$ among college dropouts and college graduates.

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

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

(13)

where $Pr(\hat{a}_i|n_{\tau}, j, \tau)$ is the agent’s belief about her ability, which we derive in Section 2.7. $U_s$ captures the utility derived from job characteristics associated with school level $s$ that is common to all agents.

### 2.6 College

The value of being in college at age $t$, $V_C(n_t, j_c, j, t)$, is determined as follows. A student of type $(j_c, j)$ enters college with assets $\hat{k}_j$. In each period, the student earns gross interest
\( R \), pays tuition \( \hat{q}_j \) and consumes \( c_{F,j} \) so that next period’s assets are determined by the budget constraint:

\[
k_{j,c,t+1} = R k_{j,c,t} - c_{F,j} - \hat{q}_j
\]  

(14)

Flow utility is given by \( u(C[c_F, \bar{c}_L]) \), where \( \bar{c}_L \) denotes a fixed amount of non-market consumption given to the agent for free in each period. Borrowing is constrained by \( k_{t+1} \geq k_{\text{min}} \). Since assets (or debt) are a function of \((j,c,j,t)\), they are not a state variable.

In each period, the student attempts \( n_c \) courses and completes each with probability \( \Pr_c(a) \). More able students are more likely to pass a course: \( \Pr'_c(a) > 0 \). Based on the number of credits completed, \( n_{t+1} \), the student updates his beliefs about \( a \). Since \( n_{t+1} \) is drawn from the Binomial distribution \( B(n, \Pr_c(a)) \), it is a sufficient statistic for the student’s entire history of course outcomes. It follows that his beliefs about \( a \) are completely determined by \( n_{t+1} \) and \( j \).

Upon observing \( n_{t+1} \), the student decides whether to work or study in period \( t + 1 \). The option of studying next period is not available if

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

2. \( n_{t+1} < n_{\text{grad}} \) and \( t = T_c \): the student fails to earn enough credits in the last year of college. He must work as a college dropout with continuation value \( V_W(k_{j,c,t+1}, n_{t+1}, j, CD, t+1) \).

If neither of these conditions is satisfied, the student chooses to remain in college if the continuation value in college is greater than that of working as a college dropout. The Bellman equation is therefore given by

\[
V_C(n_t, j_c, j, t) = u(C[c_{F,j_c,j}, \bar{c}_L]) + \beta \sum_{n_{t+1}} \Pr(n_{t+1}|n_t, j, t)V_{EC}(n_{t+1}, j_c, j, t + 1) \tag{15}
\]

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

\[
V_{EC}(k, n, j_c, j, t) = \Pr(\text{study}|n, j_c, j, t)V_C(n, j_c, j, t) + [1 - \Pr(\text{study}|n, j_c, j, t)] V_W(k, n, j_c, j, CD, t) \tag{16}
\]
if the student can choose whether to work or study next period.

Assuming that the dropout decision is subject to type I extreme value preference shocks (scaled by \( \pi \)), the probability of staying in college is given by

\[
\Pr(\text{study} | n, j_c, j, t) = \frac{\exp \left( V_C (n, j_c, j, t) / \pi \right)}{\exp \left( V_C (n, j_c, j, t) / \pi \right) + \exp \left( V_W (k_{j_c,j,t}, n, j, CD, t) / \pi \right)}
\]  

(17)

2.7 Probabilities

We now derive the probabilities governing how students accumulate credits and form beliefs about their abilities. The following are given as primitives:

1. Each type has mass \( \Pr(j) = 1/J \).

2. \( \Pr(j|\hat{a}_i) \) is the probability of drawing type \( j \) conditional on ability \( \hat{a}_i \).

3. The probability of passing a course, \( \Pr_c(a) \), is an exogenous, increasing function of ability.

All other probabilities and beliefs are found using Bayes’ Rule. The probability of passing \( \Delta n \) courses in one year, \( \Pr_n(\Delta n|a) \), is given by the standard Binomial formula. Then

\[
\Pr(n_{t+1}|n_t, j, t) = \sum_{i=1}^{N_a} \Pr_n(n_{t+1} - n_t|\hat{a}_i) \Pr(\hat{a}_i|n_t, j, t)
\]  

(18)

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

\[
\Pr(\hat{a}_i|n_t, j, t) = \frac{\Pr(\hat{a}_i) \Pr(j|\hat{a}_i) \Pr(n_t|\hat{a}_i, t)}{\Pr(n_t, j|t)}
\]  

(19)

where \( \Pr(\hat{a}_i) \) is the unconditional probability of drawing ability level \( \hat{a}_i \). \( \Pr(n_t|\hat{a}_i, t) \) is the Binomial formula for \( n_t \) successes out of \( (t-1)n_c \) draws. Also from Bayes’ Rule, we have \( \Pr(n_t, j|t) = \sum_i \Pr(\hat{a}_i) \Pr(n_t, j|\hat{a}_i, t) \) and \( \Pr(n_t, j|\hat{a}_i, t) = \Pr(j|\hat{a}_i) \Pr(n_t|\hat{a}_i, t) \).

2.8 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 $\Pr_n$. Relative to the simpler alternative where dropping out is a shock (as in Akyol and Athreya 2005) our approach has the benefit that we can use data on the characteristics and the timing of dropouts in the calibration. 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.

Manski (1989) argues that *learning about ability* may explain why many students drop out of college. Stinebrickner and Stinebrickner (2009) 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. The sensitivity analysis examines how the findings change when we assume that abilities are perfectly known at the time of high school graduation.

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.

We assume that college has a *consumption value*, $\bar{c}_L$, in order to address two empirical observations. First, data on the financial resources available to college students, summarized in Section 3, suggest that students spend less on consumption than do working individuals at similar ages. Further, many students do not take advantage of available loans to smooth consumption between college and work periods (see Bowen, Chingos, and McPherson 2009, ch. 8). In our model, non-market consumption $c_L$ reduces the marginal utility of market consumption among college students. Even if individuals smooth the marginal utility of market consumption over time, consumption *spending* jumps upon graduation. In survey data, 88% of first year students believe that being in college is more enjoyable than not
being in college (Stinebrickner and Stinebrickner, 2009).

The second observation is that students with low measured abilities attempt college, even though their ex ante probability of graduating is low. We document this fact in Section 3. In our model, the consumption value of college is an important reason why low ability students attempt college.

Our concept of non-market consumption may remind the reader of the *psychic costs* commonly found in models of school choice (e.g., Cunha, Heckman, and Navarro (2005) and Navarro (2008)). However, the two concepts play very different roles. The psychic cost is an idiosyncratic utility or disutility of attending college. Its role is to account for the observed imperfect correlation between school choices and background variables, such as cognitive test scores and family income. Its standard deviation is typically large (e.g., it is $82,000 in Stange 2012) and a large fraction of school choices is determined by psychic costs.

In our model, the consumption value of college is the same for everyone. Its role is to generate reasonable levels of consumption and borrowing among college students. While our model features preference shocks, they account for only a small fraction of college entry and dropout decisions. Their main role is to ensure smoothness of the objective function minimized by the calibration algorithm. Since we agree with Heckman, Lochner, and Todd (2006) that “explanations [of school choice] based on psychic costs are intrinsically unsatisfactory” (p. 436), we view this as an important contribution.

3 Calibration

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

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. Appendix B provides additional detail.

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 detail.

### 3.1 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.

**Schooling.** We count a student as attending college if he attempts at least 9 non-vocational credits in a given year. In the NELS:88 sample, 70% of community college entrants intend to attain a BA degree (Bound, Lovenheim, and Turner, 2010). We therefore classify persons who ever attended college without attaining a BA degree as college dropouts.

**College credits.** We measure $n_t$ as the number of completed college credits by the start of year $t$ in college divided by the number of credits taken, assuming a full course load. A full course load 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.

**Test scores.** For calibration purposes, it is helpful to utilize test scores to proxy for unobserved abilities.

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.\textsuperscript{7}

Since HS&B lacks AFQT scores, we treat GPA quartiles as equivalent to AFQT quartiles. Borghans, Golsteyn, Heckman, and Humphries (2011) show that high school GPAs and AFQT scores are highly correlated.

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, labeled $IQ$ for lack of a better symbol, as signal plus Gaussian noise:

$$IQ = \frac{\alpha_{IQ,m} + \varepsilon_{IQ}}{(\alpha_{IQ,m}^2 + 1)^{1/2}} \sim N(0, 1)$$

(20)

with $\varepsilon_{IQ} \sim N(0, 1)$. In the calibration, we divide agents into test score quartiles.

Financial variables. Our interpretation of assets $k_1$ and college costs $q$ deserves a more detailed discussion.\textsuperscript{8} A college student’s budget constraint is given by $k_{t+1} = Rk_t - c_{F_t} - q$.

Consumption of college students, $c_{F_t}$, is not observed. $q$ collects all payments the student makes that are conditional on attending college. In the data, we observe tuition net of scholarships, grants, and labor earnings while in college. We label this sum $\bar{q}_t$. Room and board are contained in $c_{F_t}$ and therefore excluded from $\bar{q}_t$. We assume that all parental transfers are not conditional on college attendance and exclude them from $\bar{q}_t$. HS&B data allow us to measure $\bar{q}_t$ for the first 2 years in college. The model’s net cost of college $q$ corresponds to the average of $\bar{q}_t$ over the 2 years for which this quantity is observed plus $987 per year for books, supplies, transportation, and similar expenses.

$k_1$ collects financial resources the student receives regardless of college attendance. This includes the student’s financial assets and any transfers he receives from his parents that are not conditional on attending college. In the model we assume that $k_1$ is paid out as

\textsuperscript{7}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.

\textsuperscript{8}Appendix C.2 describes the financial data in detail.
a lump sum at high school graduation. Since we do not observe the financial resources a student could draw on to pay for college expenses, but chooses not to, we treat \( k_1 \) as unobservable.

As students move through college, they may choose to consume more than their total financial resources. In that case \( k_t \) falls below zero, which we interpret as student debt.

An important feature of our data is that the average cost of attending college (\( \bar{q}_t \)) is close to zero. This is consistent with Bowen, Chingos, and McPherson (2009) who find that, for public 4-year colleges, average tuition payments roughly equal average scholarships and grants. It implies that, for the majority of students, the only cost of college is foregone earnings. Given that college has a consumption value, it is an attractive choice for all students who can afford it, even if the chance of graduating is small. This feature enables our model to capture the fact that a sizable fraction of students with low test scores attempt college without earning a degree.

### 3.2 Distributional Assumptions

We model the joint distribution of assets, signals and college costs \((k_1, m, q)\) as a discrete approximation of a joint Normal distribution. This allows us to model substantial heterogeneity in a parsimonious way. Setting up the joint distribution involves the following steps.

Let \( \Sigma \) be the covariance matrix of \( (\tilde{k}_1, q, m) \) and let \( \Lambda \) be its Cholesky decomposition. We draw the endowments of each type from

\[
\begin{pmatrix}
\tilde{k}_1 \\
q \\
m
\end{pmatrix}
= \begin{pmatrix}
0 \\
\mu_q \\
0
\end{pmatrix} + \begin{pmatrix}
\varepsilon_k \\
\varepsilon_q \\
\varepsilon_m
\end{pmatrix}^T \Lambda
\]

(21)

where \( \varepsilon_k, \varepsilon_q, \varepsilon_m \) are independent standard Normal vectors of length \( J \).

In HS&B data, the density of parental transfers (the main component of \( k_1 \)) is decreasing with a mass point at 0. To capture this feature, we set initial assets to a monotone transformation of \( \tilde{k}_1 \) that is given by

\[
\hat{k}_i = \eta_1 \left[ \exp (\eta_2 (i - 1) / N_k) - 1 \right]
\]

(22)

\footnote{If \( k_1 \) is paid out over time, the household problem gains a state variable, which is computationally costly.}
Roughly speaking, $\eta_1$ determines mean assets while $\eta_2$ determines assets dispersion. High values of $\eta_2$ imply a large mass of agents with assets close to zero.

Abilities take on $N_a = 9$ discrete values with equal probabilities. The ability grid approximates a standard Normal distribution with the grid points set to the conditional means of $a$ in each interval. We model abilities as a discrete approximation of a joint Normal distribution given by

$$a = \frac{\alpha_{am}m + \varepsilon_a}{(\alpha_{am}^2 + 1)^{1/2}}$$  \hspace{1cm} (23)

We set $\Pr(\hat{a}_i|j)$ to the probability that $a$ lies in the interval that contains $\hat{a}_i$, given $j$.

### 3.3 Fixed Parameters

We impose the following functional forms:

1. Utility is logarithmic in the composite consumption good. The consumption aggregator is of the form $C(c_{Ft}, c_{Lt}) = [c_{Ft}^\rho + c_{Lt}^\rho]^{1/\rho}$.

2. The probability of passing a credit is determined by the logistic function

$$Pr_c(a) = \gamma_{min} + \frac{\gamma_{max} - \gamma_{min}}{1 + \gamma_1 e^{-\gamma_2a}}$$  \hspace{1cm} (24)

where we fix $\gamma_{max} = 0.98$.

We fix the following parameters:

1. The period is one year.

2. 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.

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

4. The elasticity of substitution between the two consumption goods is set to 2 ($\rho = 0.50$). While we lack evidence on this parameter, our results vary little with its value. The discount factor is $\beta = 0.98$. 

5. $\phi_{HS} = \phi_{CD}$: Relaxing this assumption would lead students to attend college for one period in order to increase their return to ability.

We set the following parameters based on outside evidence:

1. Prices: Based on McGrattan and Prescott (2000) we set the gross interest rate to $R = 1.04$.

2. College: 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 therefore 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.

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.

3. Borrowing constraints: For the NLSY79 birth cohorts, most loans taken out during college are Stafford loans (see Johnson 2010). Until 1986, students could borrow $2,500 in each year of college up to a total of $12,500. We ignore the restriction that loan amounts cannot exceed college related expenditures. We set $k_{min,t} = -12,500$ and convert dollar values into year 2000 prices.

Table 3 summarizes these parameter values.

### 3.4 Calibrated Parameters

We calibrate the model parameters to match the data moments summarized in Table 4. The Appendix describes our data construction in detail. We show the data moments in Section 3.5 where we compare our model with the calibration targets.

Table 5 shows the values of the 20 calibrated parameters. We highlight parameters that are important for our findings.
Table 3: Fixed parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferences</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>Discount factor</td>
<td>0.98</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Elasticity of substitution</td>
<td>0.50</td>
</tr>
<tr>
<td>$\pi_c$</td>
<td>Scale of preference shocks at consumption choice</td>
<td>0.200</td>
</tr>
<tr>
<td>College</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_C$</td>
<td>Maximum duration of college</td>
<td>6</td>
</tr>
<tr>
<td>$n_{grad}$</td>
<td>Number of credits to graduate</td>
<td>20</td>
</tr>
<tr>
<td>$n_c$</td>
<td>Number of credits attempted</td>
<td>5</td>
</tr>
<tr>
<td>$k_{min}$</td>
<td>Borrowing limit</td>
<td>$19,750$</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R$</td>
<td>Gross interest rate</td>
<td>1.04</td>
</tr>
</tbody>
</table>

Table 4: Calibration targets

<table>
<thead>
<tr>
<th>Target</th>
<th>Source</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lifetime earnings by (test score quartile, schooling)</td>
<td>NLSY</td>
<td>Figure 5</td>
</tr>
<tr>
<td>Fraction in population by (test score quartile, schooling)</td>
<td>NLSY</td>
<td>Figure 4</td>
</tr>
<tr>
<td>Dropout rate by (test score quartile, year in college)</td>
<td>HS&amp;B</td>
<td>Figure 6</td>
</tr>
<tr>
<td>Average time to BA degree (years)</td>
<td>HS&amp;B</td>
<td>4.4</td>
</tr>
<tr>
<td>Mean and standard deviation of initial assets (HS graduates)</td>
<td>HS&amp;B</td>
<td>$12,463 ; $23,266</td>
</tr>
<tr>
<td>Mean and standard deviation of college costs $q$</td>
<td>HS&amp;B</td>
<td>$-584 ; $5,787</td>
</tr>
<tr>
<td>Fraction of students in debt, by year in college</td>
<td>HS&amp;B</td>
<td>Table 10</td>
</tr>
<tr>
<td>Mean student debt, by year in college</td>
<td>HS&amp;B</td>
<td>Table 10</td>
</tr>
<tr>
<td>Fraction of credits passed, by graduation status and year</td>
<td>HS&amp;B</td>
<td>Table 8</td>
</tr>
</tbody>
</table>

Notes: Dropout rates are defined as number of dropouts / number of persons who started college in year 1. College costs include cumulative tuition and fees net of scholarships, grants, earnings, and parental transfers are measured over the first 2 years of college. As explained in Section 3.1, this corresponds to $2(q - k_1/T_c)$ in the model. Initial assets measure the parental transfers high school graduates receive during the first 2 years after graduation. As explained in Section 3.1, this corresponds to $2k_1/T_c$ in the model.
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. The latter value is larger than the one estimated by Hendricks and Schoellman (2011) who do not allow $\phi_s$ to vary by schooling. As we will show below, larger values of $\phi_s$ are associated with a larger contribution of ability selection to the measured college premium. We explore the implications of lower values in the sensitivity analysis.

The mean of $q$ is negative. The majority of students incurs no financial cost when attending college. This is consistent with our empirical finding that the mean of tuition net of scholarships, grants and earnings ($\bar{q}$) is negative among college students. $q$ subtracts parental transfers that are conditional on attending college from $\bar{q}$ and is therefore negative as well.

Figure 2 shows the distribution of agents’ beliefs over their abilities for selected values of the signal $m$. Agents face considerable uncertainty about their abilities. This allows the model to generate large numbers of college dropouts.

Table 6 shows the correlation coefficients of the individual endowments. The high negative correlation of $m$ and $q$ implies strong school sorting by $m$. Students who believe they can graduate from college typically face low college costs.

The chances of graduating from college depend strongly on ability. Figure 3 shows the probability of achieving at least $n_{\text{grad}}$ credits in $T_c$ years. Low ability students have essentially no chance of graduating. High ability students are virtually guaranteed to graduate. This is a key feature of our model, which generates ability separation between college graduates and high school graduates. Having experimented with numerous model specifications, we find this model property highly robust.

### 3.5 Model Fit

We compare how closely the model attains each set of calibration targets. Overall, the model matches all targets quite well.

**Schooling and lifetime earnings.** Table 7 shows that the model closely fits the observed fraction of persons attaining each school level and their mean log lifetime earnings. Two key features of the data are: (i) 46% of those attempting college fail to attain a bachelor’s degree; (ii) college graduates earn 45% more than high school graduates over their lifetimes.
Table 5: Calibrated parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Endowments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\eta_1, \eta_2$</td>
<td>Distribution of $k_1$</td>
<td>10.00, 4.17</td>
</tr>
<tr>
<td>$\mu_q$</td>
<td>Mean of $q$</td>
<td>3.27</td>
</tr>
<tr>
<td>$\sigma_q$</td>
<td>Standard deviation of $q$</td>
<td>6.06</td>
</tr>
<tr>
<td><strong>Endowment correlations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{MK}$</td>
<td>Governs correlation $m, k_1$</td>
<td>-1.97</td>
</tr>
<tr>
<td>$\alpha_{MQ}$</td>
<td>$m, q_1$</td>
<td>1.69</td>
</tr>
<tr>
<td>$\alpha_{QK}$</td>
<td>$k_1, q_1$</td>
<td>1.49</td>
</tr>
<tr>
<td>$\alpha_{AM}$</td>
<td>$a, m$</td>
<td>3.00</td>
</tr>
<tr>
<td>$\alpha_{AFQT,M}$</td>
<td>$AFQT, m$</td>
<td>2.07</td>
</tr>
<tr>
<td><strong>Lifetime earnings</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi_s$</td>
<td>Effect of ability on lifetime earnings</td>
<td>[0.150, 0.150, 0.189]</td>
</tr>
<tr>
<td>$Y(HS)$</td>
<td>Lifetime earnings factor</td>
<td>0.9</td>
</tr>
<tr>
<td>$Y(CG)$</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>Earnings gain for each completed credit</td>
<td>0.015</td>
</tr>
<tr>
<td><strong>Other parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi$</td>
<td>Scale of preference shocks</td>
<td>0.883</td>
</tr>
<tr>
<td>$\bar{c}_L$</td>
<td>Consumption in college</td>
<td>0.001</td>
</tr>
<tr>
<td>$U(s)$</td>
<td>Preference for job type $s$</td>
<td>[CD: -0.76, CG: -2.37]</td>
</tr>
<tr>
<td>$\gamma_{min}$</td>
<td>Min. probability of passing a course</td>
<td>0.57</td>
</tr>
<tr>
<td>$\gamma_1, \gamma_2$</td>
<td>Govern probability of passing a course</td>
<td>[2.11, 8.00]</td>
</tr>
</tbody>
</table>

Table 6: Correlation of Endowments

<table>
<thead>
<tr>
<th></th>
<th>AFQT</th>
<th>$a$</th>
<th>$m$</th>
<th>$q$</th>
<th>$k_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFQT</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a$</td>
<td>0.80</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m$</td>
<td>0.85</td>
<td>0.93</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q$</td>
<td>-0.25</td>
<td>-0.30</td>
<td>-0.28</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>$k_1$</td>
<td>0.07</td>
<td>0.08</td>
<td>0.06</td>
<td>-0.07</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2: Probability $a|m$

![Figure 2: Probability $a|m$](image-url)
Figure 3: Probability of graduating from college

Table 7: School Outcomes and Lifetime Earnings

<table>
<thead>
<tr>
<th>School group</th>
<th>HS</th>
<th>CD</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>46.9</td>
<td>24.3</td>
<td>28.8</td>
</tr>
<tr>
<td>Model</td>
<td>46.9</td>
<td>24.5</td>
<td>28.6</td>
</tr>
<tr>
<td>Gap (pct)</td>
<td>0.0</td>
<td>0.8</td>
<td>-0.8</td>
</tr>
<tr>
<td>Lifetime earnings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>600</td>
<td>643</td>
<td>944</td>
</tr>
<tr>
<td>Model</td>
<td>596</td>
<td>646</td>
<td>936</td>
</tr>
<tr>
<td>Gap (pct)</td>
<td>-0.7</td>
<td>0.5</td>
<td>-0.8</td>
</tr>
</tbody>
</table>

Note: The table shows the fraction of person that chooses each school level and the exponential of their mean log lifetime earnings, discounted to age 1. “Gap” denotes the percentage gap between model and data values.
Figure 4 breaks down the schooling outcomes by test score quartiles (Bound, Lovenheim, and Turner 2010’s Figure 2 documents similar patterns in NLS72 and NELS:88 data). The model replicates the patterns observed in the data.

The behavior of low ability persons poses a challenge. Around 20% of persons in the lowest AFQT quartile attempt college, while fewer than 5% achieve college degrees. Previous models invoked random psychic costs to account for the college entry decisions of these students. In our model, low AFQT students view college mainly as a consumption good.

Figure 5 shows average lifetime earnings by school group and test score quartile. Each panel displays one school group. The model closely matches the data cells with large numbers of
Table 8: College credits

<table>
<thead>
<tr>
<th>Year</th>
<th>College dropouts</th>
<th>College graduates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Data</td>
</tr>
<tr>
<td>1</td>
<td>74.2</td>
<td>67.7</td>
</tr>
<tr>
<td>2</td>
<td>72.9</td>
<td>71.8</td>
</tr>
<tr>
<td>3</td>
<td>70.0</td>
<td>66.9</td>
</tr>
<tr>
<td>4</td>
<td>64.0</td>
<td>63.8</td>
</tr>
</tbody>
</table>

Notes: The table shows the number of college credits completed at the end of each year in college divided by the number of credits attempted by students who eventually earn a college degree.

College dropouts. Figure 6 compares dropout rates between the model and High School & Beyond data. Dropout rates are defined as the number of persons dropping out at the 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.

Table 8 shows shows the fraction of completed college credits for each year in college. Students are divided into two groups: those who eventually drop out and those who eventually earn a college degree.

Financial resources. The final set of calibration targets consists of financial variables. Table 9 shows the means and standard deviations of initial assets ($k_1$) and college costs ($q$). Table 10 shows student debt levels at the end of the first 4 years in college.

4 Results

4.1 Ability Selection

Our main question is: What fraction of the college earnings premium represents selection by ability as opposed to returns to schooling?
Figure 5: Lifetime earnings by AFQT / school

Notes: Lifetime earnings are discounted to age 1 and expressed in thousands of year 2000 dollars.
Figure 6: Dropout rates by AFQT / year in college

Notes: The figure shows the fraction of persons initially enrolled in college who drop out at the end of each year in college.
Table 9: Financial moments

(a) Entire population

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution of $k_1$, HS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>15,322</td>
<td>12,463</td>
</tr>
<tr>
<td>standard deviation</td>
<td>20,949</td>
<td>23,266</td>
</tr>
<tr>
<td>Distribution of $q$, college</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>-766</td>
<td>-584</td>
</tr>
<tr>
<td>standard deviation</td>
<td>5,146</td>
<td>5,787</td>
</tr>
</tbody>
</table>

(b) High school GPA quartiles

<table>
<thead>
<tr>
<th>HS GPA quartile</th>
<th>Mean $q$</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Data</td>
</tr>
<tr>
<td>1</td>
<td>-2,521</td>
<td>-2,266 (678)</td>
</tr>
<tr>
<td>2</td>
<td>-1,810</td>
<td>-1,741 (454)</td>
</tr>
<tr>
<td>3</td>
<td>-454</td>
<td>-509 (308)</td>
</tr>
<tr>
<td>4</td>
<td>-232</td>
<td>-20 (253)</td>
</tr>
</tbody>
</table>

Notes: The table shows the mean and standard deviations of $k_1$ among high school graduates. $k_1$ is not observed for college students. The bottom panel shows the mean and standard deviation of $q$ for students in their second year of college.

Table 10: Student Debt

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean debt</th>
<th>Fraction with debt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Data</td>
</tr>
<tr>
<td>1</td>
<td>4,298</td>
<td>3,549</td>
</tr>
<tr>
<td>2</td>
<td>7,136</td>
<td>6,060</td>
</tr>
<tr>
<td>3</td>
<td>8,606</td>
<td>8,045</td>
</tr>
<tr>
<td>4</td>
<td>11,150</td>
<td>9,740</td>
</tr>
</tbody>
</table>

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.
Table 11: Ability Selection

<table>
<thead>
<tr>
<th></th>
<th>CD</th>
<th>CG</th>
<th>Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean log lifetime earnings</td>
<td>0.11</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>Schooling: ( Y(s) ) and ( \phi(s) )</td>
<td>0.00</td>
<td>0.05</td>
<td>10.5</td>
</tr>
<tr>
<td>Delayed work start</td>
<td>-0.08</td>
<td>-0.20</td>
<td>-43.7</td>
</tr>
<tr>
<td>Credits</td>
<td>0.13</td>
<td>0.33</td>
<td>73.1</td>
</tr>
<tr>
<td>Ability selection</td>
<td>0.06</td>
<td>0.21</td>
<td>47.8</td>
</tr>
</tbody>
</table>

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.

In the model, the mean log lifetime earnings of school group \( s \), discounted to age 1, is given by \( \mathbb{E}[\phi_s(a - \alpha) + \mu n_\tau + \log(R^{-\tau} Y(s))|s] \) where \( \tau = 1 \) and \( n_\tau = 0 \) for high school graduates. The lifetime earnings gap between college graduates and high school graduates may be decomposed into 4 terms:

1. a price effect: \( \log Y(s) - \log Y(HS) + (\phi_s - \phi_{HS})\mathbb{E}(a - \alpha|HS) \);
2. postponement of labor market entry: \( \mathbb{E}\log R^{-\tau} \)
3. credits earned: \( \mathbb{E}(\mu n_\tau|s) \)
4. ability selection: \( \phi_{HS}[\mathbb{E}(a - \alpha|s) - \mathbb{E}(a - \alpha|HS)] \).

Table 11 shows the decomposition implied by the model. It attributes 48% of the 0.45 log lifetime earnings gap between college graduates and high school graduates to ability selection. This result is quite robust, as we show in Section 4.6. One reason why ability sorting is strong is that it occurs at two levels: at college entry and at the dropout / college completion stage.

Figure 7 illustrates both levels of selection. Consider the first panel. For each ability level, it shows the fraction of persons who attain each schooling level. The second panel shows the same for each ability signal.

It is evident that selection into college attendance is strongly related to ability signals. Since abilities and signals are strongly correlated, this implies that college attendance and
abilities are also strongly related. Very few high ability students receive such poor signals that they refrain from attempting college.

The second level of selection, college graduation, is dominated by individual abilities. Even though around half of median ability persons attempt college, almost none manage to attain a degree. Taken together the two levels of selection imply that the ability distributions for high school graduates and college graduates are strongly separated.

To complete the characterization, panels (c) and (d) show how school outcomes vary with college costs and initial assets. College entry is strongly negatively related to $q$, while college completion is not. This partly reflects the negative correlation of $q$ and $m$. The correlation between initial assets and schooling is weak. This partly reflects the fact that borrowing constraints are not important for the cohort we study (see Section 4.4).

### 4.2 Understanding College Entry

Figure 8 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 percentile. It summarizes the financial stakes that motivate entry and dropout decisions. Due to the complementarity implied by $\phi_{CG} > \phi_{HS}$, only high ability students can expect large gains from earning a college degree. The gains from attending college without earning a degree are much smaller. 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.

For most students, the probability of graduating is either very small or very close to 1. This is a robust feature of our model that generates strong ability separation between college graduates and other workers. It happens because the number of completed courses is drawn from a Bernoulli distribution. The probability of passing more than $n_{\text{grad}} = 20$
Figure 7: Schooling and Individual Endowments

(a) Ability and Schooling

(b) Signal and Schooling

(c) Ability and Schooling

(d) Signal and Schooling

Notes: Each bar shows the fraction of persons attaining each school level (HS, CD, CG).
Notes: Panel (a) shows mean log lifetime earnings of students who attain each school level. Panel (b) shows the probability of graduating from college, conditional on staying in college for the maximum number of periods permitted.

out of $n_c T_c = 30$ courses increases sharply in the probability of passing a single course. The payoff from attempting college therefore increases far more sharply with ability that lifetime earnings differences would suggest. This is also key for understanding dropout behavior.

**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 9.

The top panel shows that college entry is strongly related to the ability signal and the associated chance of graduating from college (earning $n_g$ credits in $T_c$ periods). The majority of students with graduation probabilities above 0.5 attempt college. However, the fraction of students that graduate is substantially lower than the fraction that could have graduated, especially for intermediate signals.

The bottom panel shows lifetime earnings, discounted to age 1, received for each school outcome. Except for the lowest signal quartile, completing college increases lifetime earnings by at least $100,000$. However, attempting college without graduating yields only small earnings gains.
Only students who can expect to graduate with a probability of at least 0.5 increase their lifetime earnings substantially by attempting college. Still, the majority of students who graduate with probabilities of at least 0.5 attempt college. They do so for two reasons: (i) the option value of dropping out after gathering more information about their abilities and (ii) the consumption value of college. An important insight emerges: Since the financial cost of college is often negative, only students with low graduation probabilities or with high college costs \( q \) fail to try college.

**The value of college entry.** To illustrate selection into college, Figure 10 shows the value of attending college relative to starting work as a high school graduate. The value gap is expressed as an age 1 transfer. For example, a value gap of 100 means that the household would require a transfer of $100,000 in exchange for giving up the option of attempting college.

Students with better ability signals or lower college costs find it more profitable to try college. The ability signal matters more for the choice in the sense that all agents with high ability signal choose college regardless of college costs, while students with the lowest signals only attempt college if its cost is negative.

An interesting feature is the strong asymmetry in the value gain from attempting college. High ability agents reap large gains from attending college. Agents receiving the highest
signal are willing to pay more than $200,000 for the opportunity to attend college. This transfer is larger than the expected lifetime earnings gain because college has a consumption value and because students with low \( q \) realizations get paid to attend college. By contrast, agents receiving the lowest signals are nearly indifferent between attending college and working as high school graduates. Those with high college costs require payments, while those with low college costs would be willing to pay to attend college. Of course, some of these students would drop out of college after the first year. This is one reason why the required payment is so small.

This finding raises the concern that college attendance in the model is unreasonably sensitive to changes in tuition. To address this issue, we compute the effect of reducing tuition by $1,000. A sizable empirical literature estimates the effects of reducing tuition on college attendance. Dynarski (2003) summarizes this literature as well as her own estimates as follows: a $1,000 reduction in the cost of attending college (in 1998 prices) leads to a 3 to 4 percentage point increase in attendance. In our model, reducing the cost of college by $1,000 increases college attendance by 4 percentage points.

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.

A significant fraction of students enter college choosing not to graduate. Recall that students commit to a constant consumption level for their entire college career. If this level is chosen too high, students run out of assets before they have a chance to accumulate enough credits to graduate. Panel (a) of Figure 11 shows the fraction of college students in each signal quintile who choose such high consumption levels. The majority of students in the lowest signal quintile makes 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). Even in the top signal quintile more than 15% of students plan to drop out. These are students who face high college costs, which would make staying in college for a long time painful due to low consumption.

A 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 8). In response, these
Figure 10: College Choice At Age 1

Notes: The figure shows the one-time payment, in thousands of dollars, a student would require in order to reverse his college entry decision. The psychic costs $p_s$ are set to zero. Positive payments indicate that the student prefers to enter college. Negative payments indicate that the student prefers to work as a high school graduate.
students update their beliefs about their graduation prospects and some drop out.

For students in each signal decile, Figure 12 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 (2009) who find that academic performance is strongly related to dropout decisions.

A third reason for dropping out is preference shocks. To isolate their effects, we recompute the model setting $\pi = \pi_E = 0$. Figure 13 summarizes how this changes school sorting by ability and dropout rates. Panel (a) shows the fraction of persons attaining each school level in the baseline model. Panel (b) shows the model without preference shocks.

The fraction of persons attempting college does not change by a great deal, but ability sorting becomes stronger. Dropout rates fall substantially for all ability levels, cutting the overall dropout rate nearly by half. One reason why these changes are so large is that, for most students, the financial stakes are not very large. Lifetime earnings gains are modest, except for high ability students who rarely drop out (see Figure 7). Self-selection implies that the direct cost of college is close to zero (see Table 9). Relatively small preference
shocks are therefore sufficient to change the schooling decisions of many agents (see Figure 10).

4.4 How Important Are Borrowing Constraints?

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.

As shown in Table 12, the fraction of high school graduates who attempt college rises from 52% to 62%. The fraction of college students who drop out remains near 50%. Increased schooling reduces mean log abilities at all levels, but mostly among high school graduates, leading to a modest increase in the college lifetime earnings.
Figure 13: School outcomes without preference shocks

(a) Baseline model

(b) No preference shocks

Table 12: Increased borrowing limits

<table>
<thead>
<tr>
<th>School group</th>
<th>HS</th>
<th>CD</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction by schooling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>0.47</td>
<td>0.24</td>
<td>0.29</td>
</tr>
<tr>
<td>Relax borrowing constraints</td>
<td>0.39</td>
<td>0.25</td>
<td>0.36</td>
</tr>
<tr>
<td>Mean log ability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>-0.48</td>
<td>-0.12</td>
<td>0.95</td>
</tr>
<tr>
<td>Relax borrowing constraints</td>
<td>-0.66</td>
<td>-0.24</td>
<td>0.93</td>
</tr>
</tbody>
</table>
4.5 How Important Is Learning About Ability?

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. 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 80% of a semester course load.

Figure 14 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 15 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’ informa-
tion 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. 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. At the start of the 2nd year in college, the standard deviation is above 0.25 for all quintiles. For students with very high or low signals, the standard deviation has barely dropped due to learning.

### 4.6 Robustness

To be written.

### 5 Conclusion

To be written.
References


A Appendix: CPS Data

A.1 Sample

Our sample contains all men between the ages of 18 and 75 observed in the 1964-2010 waves of the March Current Population Survey (King, Ruggles, Alexander, Flood, Genadek, Schroeder, Trampe, and Vick, 2010). We drop persons who live in group quarters or who fail to report wage or business income.

A.2 Schooling Variables

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

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 within each school group. This is calculated by simple averaging across all years and birth cohorts. Denote this fraction by $f(t|s)$ where $t$ is age. Figure 16 shows the resulting profiles. For comparison, we also show the fraction of persons born between 1957 and 1964.
Notes: NLSY and CPS show the fraction of persons earning more than $2,000 in the 1957-1964 cohorts. CPS fit shows the estimated CPS profile $f(t)$.

earning more than $2,000. We estimate this from the CPS and from the NLSY79 sample described in Section B. Overall, the CPS profile $f(t|s)$ is close to the NLSY79 profile almost everywhere.

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 to 1980 and ages up to 70 years. We use weighted least squares to account for the different number of observations in each cell.

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

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

(25)
Figure 17: 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.

For years after 2010, we impose the average year dummy.

Figure 17 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 Appendix: NLSY79 Data

The NLSY79 sample covers men born between 1957 and 1964 who earned at least a high school diploma. We drop persons with missing information as detailed below. We observe schooling and earnings from 1978 to 2006. 94% of the sample participated in the AFQT in
Table 13: Summary statistics for the NLSY79 sample

<table>
<thead>
<tr>
<th>School class</th>
<th>HSD</th>
<th>HSG</th>
<th>CD</th>
<th>CG</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction</td>
<td>17.8</td>
<td>38.3</td>
<td>20.8</td>
<td>23.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Avg.school</td>
<td>10.5</td>
<td>12.1</td>
<td>14.1</td>
<td>17.0</td>
<td>13.4</td>
</tr>
<tr>
<td>Range</td>
<td>2 - 12</td>
<td>9 - 13</td>
<td>13 - 20</td>
<td>12 - 20</td>
<td>2 - 20</td>
</tr>
<tr>
<td>AFQT percentile</td>
<td>22.6</td>
<td>41.3</td>
<td>57.6</td>
<td>78.8</td>
<td>50.0</td>
</tr>
<tr>
<td>N</td>
<td>919</td>
<td>1447</td>
<td>800</td>
<td>675</td>
<td>3841</td>
</tr>
</tbody>
</table>

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.

1980. Table 13 shows summary statistics for this sample.

Since we classify individuals based on their schooling at the beginning of the first 5 year work spell, our sample contains fewer highly educated persons than other studies have found. For example, Hendricks and Schoellman (2011) find that roughly 2/3 of (white) men born around 1960 attempt college in the NLSY79 and in the U.S. Census. Both measures are based on educational attainment observed after age 35.

In each school group we observe a wide range of completed years of schooling. Notably, some high school graduates report only 9 or 10 years of schooling. This likely represents misreporting of either grades attended or degrees earned.

B.1 Schooling Variables

For each person, we record all degrees and the dates they were earned. Because degrees are not consistently reported before 1988, we drop all persons who are not interviewed in either 1988 or 1989.

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 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.

We assign each person a school group (high school dropout, high school graduate, college dropout, college graduate) based on the highest degree earned by the time work starts. 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.

Of the 5579 men in the sample we drop 1548 whose schooling variables are incomplete.

**B.2 Lifetime Earnings**

For each NLSY79 individual, we estimate the present value of lifetime earnings using the following method.

Our measure of labor earnings adds wage and salary income and \(2/3\) of business income. The consumer price index is used to convert earnings into year 2000 prices. Lifetime earnings are measured as the present value of earnings up to age 70, discounted to age 18. We assume that earnings are zero before age 18 for high school dropouts and high school graduates, before age 20 for college dropouts, and before age 22 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 Section A).
The present value of lifetime earnings for the average CPS person is given by 
\[ Y_{CPS}(s) = \sum_{t=18}^{70} g_{CPS}(t|s) R^{18-t}. \] 
The fraction of lifetime earnings typically earned at age \( t \) is given by 
\[ g_{CPS}(t|s) R^{18-t} / Y_{CPS}(s). \]

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, 
\[ g_{CPS}(t|s) R^{18-t} / Y_{CPS}(s). \]

An example may help the reader understand this approach. Suppose we observe a high school graduate with complete earnings observations between the ages of 18 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 (see Figure 18 below). 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 14 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, around 70% of lifetime earnings is earned before age 45. This is shown in Figure 18, which displays the cumulative fraction of lifetime earnings received by a given age. This is based on the fitted CPS profiles, \( g_{CPS}(t|s) \).

C High School and Beyond Longitudinal Survey and Postsecondary Education Transcript Study

The National Education Longitudinal Studies (NELS) program of the National Center for Education Statistics (NCES) was established to study the educational, vocational, and personal development of young people. Thus far, the NELS program consists of five major studies: the National Longitudinal Study of the High School Class of 1972, High School and Beyond (HS&B), the National Education Longitudinal Study of 1988, the Education Longitudinal Study of 2002, and the High School Longitudinal Study of 2009.
Table 14: Lifetime earnings data

<table>
<thead>
<tr>
<th></th>
<th>HSD</th>
<th>HSG</th>
<th>CD</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (thousands)</td>
<td>431</td>
<td>600</td>
<td>643</td>
<td>944</td>
</tr>
<tr>
<td>Min</td>
<td>41</td>
<td>30</td>
<td>82</td>
<td>159</td>
</tr>
<tr>
<td>Max</td>
<td>1627</td>
<td>2209</td>
<td>2404</td>
<td>2797</td>
</tr>
<tr>
<td>Standard deviation (log)</td>
<td>0.54</td>
<td>0.51</td>
<td>0.55</td>
<td>0.50</td>
</tr>
<tr>
<td>Fraction zero</td>
<td>6.2</td>
<td>2.9</td>
<td>3.5</td>
<td>2.3</td>
</tr>
<tr>
<td>Fraction &lt; 100,000</td>
<td>1.8</td>
<td>0.2</td>
<td>1.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Correlation with wage</td>
<td>0.56</td>
<td>0.61</td>
<td>0.74</td>
<td>0.78</td>
</tr>
<tr>
<td>N</td>
<td>278</td>
<td>578</td>
<td>343</td>
<td>319</td>
</tr>
</tbody>
</table>

Notes: Mean denotes exp(mean log lifetime earnings). Correlation with wage denotes the correlation between log lifetime earnings and log wages at age 40. N is the number of observations.

Figure 18: Cumulative fraction of lifetime earnings received by each age
The survey that corresponds most closely to the NLSY79 cohort, the cohort of focus in this paper, is the HS&B cohort. Precisely, the HS&B survey included two cohorts: the 1980 senior class and the 1980 sophomore class. Both cohorts were surveyed every two years through 1986, and the 1980 sophomore class was surveyed again in 1992. In 1992, i.e., 10 years after high school graduation, postsecondary transcripts from all institutions attended since high school graduation were collected for the sophomore cohort under the initiative of the Postsecondary Education Transcript Study (PETS). The transcript study for the senior cohort was conducted much earlier, in 1984, i.e. only 4 years after high school graduation. Because of the availability of postsecondary education transcript histories for 10 years after high school graduation, we choose to focus on the 1980 sophomore class.

The HS&B data files are available through ICPSR at the University of Michigan (ICPSR 8896). We supplement these with PETS, obtained through a restricted license granted by the National Center for Education Statistics. We restrict attention to sophomores surveyed at least through 1986, which leaves us with 14,825 student records and 17,363 transcripts collected from 4,079 institutions. Hence, while the demographic and financial information (e.g., parental transfers, earnings, school costs, grants) is available for the first four years after high school, college performance and enrollment information is available for ten years after high school.

C.1 Enrollment and Dropout Statistics

The sample is restricted to white males graduating from high school with their class 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.

We use the course-level data and institution-level data derived from postsecondary transcripts to compute, for each student, attempted and earned non-vocational undergraduate credits for each academic year. All postsecondary credits taken prior to the date a bachelor’s degree was obtained are considered undergraduate credits. For students that never earned a bachelor’s degree, all postsecondary credits are treated as undergraduate. Transfer credits are dropped to avoid double-counting. We count withdrawals that appear on transcripts as attempted but unearned credits. Vocational credits are identified as credits
taken at a vocational school (e.g., a police academy, a school of cosmetology or a health occupation school).

We say someone enrolls in college if they attempt at least 9 non-vocational credits. 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.

To obtain dropout statistics, we restrict attention to immediate 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 $x$, enrolled less than part time in year $x + 1$, and failed to obtain a bachelor degree within 5 years. The results reported in the paper are based on the 7 credit hours definition of part time enrollment. The results are not highly sensitive to varying the definition of "part time," as most college students enroll full time. Nearly all college graduates, precisely 86% of college graduates in our sample, exhibit continuous enrollment to graduation. In fact, it is common to treat those with a break in enrollment as permanent dropouts (e.g. Stange, 2010). It is true, however, that a few of those return to school and even graduate. This is why we classify anyone graduating within 5 years as college graduates. We also exclude from the sample those students that obtain a college degree but fall into our definition of dropouts (These are students with enrollment breaks returning to school later and graduating in 6 or more years.).

We compute fractions of college students dropping out in years 1 through 6, overall and by high school GPA quartile. These are used as calibration targets and shown in Table 15.

### C.2 Financial variables

In the second and third follow-up interviews (1984 and 1986), all students are questioned in detail regarding their education expenses, various sources of financial support, and own earnings. We translate all amounts into 2000 dollars using the consumer price index. Table 16 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 the school-related transfer and the direct transfer to the student in the form of inkind support and gifts. Precisely, 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.”
Table 15: School Attainment by High School GPA Quartile

<table>
<thead>
<tr>
<th>quartile of high school GPA</th>
<th>1.00</th>
<th>2.00</th>
<th>3.00</th>
<th>4.00</th>
<th>All</th>
<th>No. obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>fraction of entrants that graduate</td>
<td>0.13</td>
<td>0.31</td>
<td>0.58</td>
<td>0.76</td>
<td>0.55</td>
<td>840</td>
</tr>
<tr>
<td>fraction that drop out after y1</td>
<td>0.40</td>
<td>0.25</td>
<td>0.13</td>
<td>0.07</td>
<td>0.16</td>
<td>200</td>
</tr>
<tr>
<td>fraction that drop out after y2</td>
<td>0.24</td>
<td>0.21</td>
<td>0.16</td>
<td>0.07</td>
<td>0.14</td>
<td>191</td>
</tr>
<tr>
<td>fraction that drop out after y3</td>
<td>0.15</td>
<td>0.11</td>
<td>0.06</td>
<td>0.05</td>
<td>0.08</td>
<td>108</td>
</tr>
<tr>
<td>fraction that drop out after y4</td>
<td>0.06</td>
<td>0.08</td>
<td>0.04</td>
<td>0.03</td>
<td>0.05</td>
<td>52</td>
</tr>
<tr>
<td>fraction that drop out after y5</td>
<td>0.02</td>
<td>0.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.02</td>
<td>30</td>
</tr>
</tbody>
</table>

Notes: High school GPA quartiles are defined for the 3671 (white male) graduating seniors in 1982. Of those, 53% entered college, i.e. took at least 9 credits during the 82-83 academic year. Restricting the sample further to those entrants with complete transcript histories and cleaning the sample of inconsistencies leaves us with 1436 freshmen, for which the dropout rates are reported.

Table 16: Average Cumulative Financial Amounts for Currently Enrolled College Students, by Year

<table>
<thead>
<tr>
<th></th>
<th>82-83</th>
<th>83-84</th>
<th>84-85</th>
<th>85-86</th>
</tr>
</thead>
<tbody>
<tr>
<td>$qC$ cum</td>
<td>-8272</td>
<td>-16409</td>
<td>-23580</td>
<td>-31535</td>
</tr>
<tr>
<td>tuition cum</td>
<td>4428</td>
<td>9276</td>
<td>16897</td>
<td>24632</td>
</tr>
<tr>
<td>grants cum</td>
<td>1308</td>
<td>2619</td>
<td>4119</td>
<td>5682</td>
</tr>
<tr>
<td>school related parental transfer</td>
<td>3131</td>
<td>6558</td>
<td>9455</td>
<td>12652</td>
</tr>
<tr>
<td>direct parental transfer (inkind &amp; gifts)</td>
<td>2923</td>
<td>5856</td>
<td>9029</td>
<td>12187</td>
</tr>
<tr>
<td>earnings cum</td>
<td>5137</td>
<td>9728</td>
<td>15323</td>
<td>20377</td>
</tr>
<tr>
<td>loans cum</td>
<td>983</td>
<td>2184</td>
<td>3416</td>
<td>4670</td>
</tr>
<tr>
<td>fraction in debt</td>
<td>0.28</td>
<td>0.36</td>
<td>0.42</td>
<td>0.48</td>
</tr>
<tr>
<td>No. of observations</td>
<td>1436</td>
<td>1236</td>
<td>1045</td>
<td>934</td>
</tr>
</tbody>
</table>

Notes: All average amounts are unconditional, i.e. zero values are averaged in. Cumulative loans refer to the current debt amount. Fraction in debt refers to students with a positive amount of current debt.
The direct transfer to the student is the approximate dollar value of inkind support such as room and board (the coresidence benefit), use of car, medical expenses and insurance, clothing, and any other cash or gifts. The value of the direct transfer to the student is reported in terms of detailed intervals up to $3000 of current dollars. We proxy its value by using the midpoint value of the relevant interval and $4000 of current dollars if it falls into the last open interval. The direct transfer variable is available for calendar years only. We allocate it equally across the relevant academic years.

The school-related transfer is available only for the first two academic years after high school graduation. For academic years 8485 and 8586, we proxy it with the school-related transfer for 8384 adjusted by the change in tuition net of grants/scholarships.

We identify as high school graduates those with no postsecondary education history or those with strictly less than 7 nonvocational credits in each of the first four academic years since high school.\(^\text{10}\) Note this is consistent with our 7 credit part-time definition used above to measure dropout rates.

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 \(jk\). Simplifying obtains that the proportion of year \(j\) income attributed to academic year \(ij\) is given by \(\alpha_{ij} = \frac{cred_{jk}}{cred_{ij} + cred_{jk}}\). We attribute half of the 1982 earnings to the 8283 academic year.

\(^{10}\)We also worked with the cutoff value of 0. The resulting moments were affected very little.