Hendricks, Lutz; Leukhina, Oksana

Working Paper

How risky is college investment?

CFS Working Paper Series, No. 502

Provided in Cooperation with:
Center for Financial Studies (CFS), Goethe University Frankfurt

Suggested Citation: Hendricks, Lutz; Leukhina, Oksana (2015) : How risky is college investment?, CFS Working Paper Series, No. 502, Goethe University Frankfurt, Center for Financial Studies (CFS), Frankfurt a. M., https://nbn-resolving.de/urn:nbn:de:hebis:30:3-370162

This Version is available at:
https://hdl.handle.net/10419/106821

CENTER FOR FINANCIAL STUDIES

CFS Working Paper Series

No. 502

Lutz Hendricks, Oksana Leukhina

How Risky Is College Investment?

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

The CFS Working Paper Series

presents ongoing research on selected topics in the fields of money, banking and finance. The papers are circulated to encourage discussion and comment. Any opinions expressed in CFS Working Papers are those of the author(s) and not of the CFS.

The Center for Financial Studies, located in Goethe University Frankfurt's House of Finance, conducts independent and internationally oriented research in important areas of Finance. It serves as a forum for dialogue between academia, policy-making institutions and the financial industry. It offers a platform for top-level fundamental research as well as applied research relevant for the financial sector in Europe. CFS is funded by the non-profit-organization Gesellschaft für Kapitalmarktforschung e.V. (GfK). Established in 1967 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia. GfK members comprise major players in Germany's financial industry. The funding institutions do not give prior review to CFS publications, nor do they necessarily share the views expressed therein.

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

How Risky Is College Investment?*

Lutz Hendricks† Oksana Leukhina‡

January 30, 2015

Abstract

This paper is motivated by the fact that nearly half of U.S. college students drop out without earning a bachelor's degree. Its objective is to quantify how much uncertainty college entrants face about their graduation outcomes. To do so, we develop a quantitative model of college choice. The innovation is to model in detail how students progress towards a college degree. The model is calibrated using transcript and financial data. We find that more than half of college entrants can predict whether they will graduate with at least 80% probability. As a result, stylized policies that insure students against the financial risks associated with uncertain graduation have little value for the majority of college entrants.

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

Key words: Education. College dropout risk.

*For helpful comments we thank Philip Bond, V. V. Chari, Jonathan Heathcote, Larry Jones, José-Victor Rios-Rull, and seminar participants at the Federal Reserve Bank of Minneapolis and the University of Washington.

†University of North Carolina, Chapel Hill; lutz@lhendricks.org

‡University of Washington; oml@u.washington.edu

1 Introduction

A growing literature emphasizes that college is a risky investment.[1] This is motivated by the empirical finding that nearly half of all U.S. college students drop out before earning a bachelor's degree (Bound and Turner, 2011). A separate branch of the literature points out that college choices are strongly related to student characteristics, especially their cognitive skills (Belley and Lochner, 2007). These findings suggest that college dropouts may lack the abilities or preparation required for earning a college degree.

Whether students drop out of college mainly because of endowments (heterogeneity) or shocks (uncertainty) has important policy implications. If graduation outcomes are highly uncertain, policies that insure students against failure, such as income contingent loans (Chapman, 2006), may increase welfare. On the other hand, if dropping out is mainly due to students' inability to satisfy the requirements for a college degree, policies that improve student preparation, such as remedial coursework (Bettinger and Long, 2009), may be called for.

The main purpose of this paper is to quantify the relative importance of heterogeneity and uncertainty for college dropout decisions. Put differently, we ask: how predictable or risky is graduation from the point of view of college entrants?

Transcript data. To address this question, we develop a structural model of college choice. We depart from much of the literature by modeling in detail how students progress towards satisfying the requirements for a college degree. This allows us to introduce *transcript data*, which we argue are key for measuring the distribution of college preparation among freshmen and, therefore, their graduation probabilities. Transcripts reveal how rapidly students progress towards a college degree, which is directly related to their graduation chances. Moreover, transcripts provide *repeated* indicators of college preparation for the same individual. This enables the model to decompose the observed dispersion in college outcomes into the contributions of persistent heterogeneity and transitory shocks.

We obtain college transcripts for a representative sample of high school sophomores in 1980 from the Postsecondary Education Transcript Study (PETS, section 2). We focus on how rapidly students accumulate college credits. The data also contain information about high

[1] Recent work includes Arcidiacono et al. (2012), Athreya and Eberly (2013), Stange (2012), and Trachter (2014).

school GPAs, financial resources, college costs, and degrees earned. This information is used to calibrate the structural model.

The transcripts reveal large and persistent heterogeneity in individual credit accumulation rates. In their freshmen year, students in the 80^{th} percentile of the credit distribution earn twice as many credits as do students in the 20^{th} percentile. The correlation between credit accumulation rates across years is 0.43. Credit accumulation is strongly related to high school GPAs and to college graduation. College graduates earn about 50% more credits than do college dropouts. After 3 years in college, dropouts have earned fewer than half of the 125 credits that students have earned, on average, when they graduate. These observations suggest that students' inability to complete the requirements for a college degree may be an important reason for dropping out.

Structural model. To recover what the transcript data imply for the predictability of college graduation, we develop a structural model of college choice (section 3).[2] The model follows one cohort of high school graduates through their college and work careers. At high school graduation, students are endowed with heterogeneous abilities (college preparation) and financial resources. Following Manski (1989), we assume that high school graduates only observe a noisy signal of their abilities. They choose whether to attempt college or work as high school graduates. While in college, students make consumption-savings and work-leisure decisions subject to a borrowing constraint. Students face financial shocks that affect college costs and wages, as well as preference shocks.

Our main departure from the literature is to model students' progress through college in detail.[3] This allows us to map transcript data directly into model objects. We model credit accumulation as follows. In each period, a college student attempts a fixed number of courses. He passes each course with a probability that increases with his ability. At the end of each year, students who have earned a given number of courses graduate. The remaining students update their beliefs about their abilities based on the information contained in their course outcomes. Then they decide whether to drop out or continue their studies in the next period. Students must drop out if they lack the means to pay for college, or if

[2] The model extends Hendricks and Leukhina (2014), mainly by incorporating financial shocks while in college.

[3] In much of the literature, college is a black box. Exceptions include Arcidiacono et al. (2012), Garriga and Keightley (2007), and Stange (2012).

they fail to earn a degree after 6 years in college.

Results. We calibrate the model to match a rich set of observations that we construct from PETS, High School & Beyond (HS&B) and NLSY79 data (section 4). The calibrated model successfully accounts for students' credit accumulation, college entry and dropout decisions, both across years and across high school GPA quartiles, as well as for financial statistics.

The model allows us to compute the distribution of individual *graduation probabilities*, based on the information available to high school graduates at the time of college entry. Our main finding is that college graduation is highly predictable for a large fraction of college entrants (section 5). At the time of college entry, 30% of students face graduation probabilities above 80% while 24% face graduation probabilities below 20%.[4] Among high school graduates who do not enter college, graduation probabilities are far worse. 90% of these students face graduation probabilities below 20%.

We show that the high predictability of graduation outcomes has three causes.

1. The model implies substantial heterogeneity in students' *graduation prospects*. By this, we mean the probability of completing the required number of courses within the 6 years that students may be enrolled in college. Among college entrants, 61% face graduation prospects above 80% and 9% face graduation prospects below 20%.

2. The model implies that students' financial incentives for persisting in college depend strongly on their graduation prospects. For students with weak graduation prospects, staying in college until graduation (or for the permitted 6 years) reduces lifetime earnings relative to their optimal dropout decisions. As a result, many low ability students drop out of college, even though they could graduate. High ability students, on the other hand, can expect large earnings gains if they try to graduate. These students rarely drop out.

[4] In work in progress, we present econometric evidence supporting the notion that graduation is highly predictable. We estimate probit regressions to predict the graduation outcomes of college entrants in HS&B data. We estimate the regressions on one half of the sample and predict college graduation for the other half. Consistent with the structural model, we find that we can predict outcomes (graduation or dropout) with high accuracy (a probability of at least 80%) for about half of the students.

3. Students can accurately predict their abilities, and thus their graduation prospects, before they make college entry decisions. In other words, their ability signals are quite precise.

Financial heterogeneity plays a minor role for the predictability of college outcomes. One reason is that most model students are not close to their borrowing limits.[5] A second reason is that students can increase work hours to finance additional time in college.

We now discuss how the model recovers the distribution of graduation prospects, mainly from transcript data. The key data feature is the large dispersion of credit accumulation rates. According to our model, the number of courses passed is drawn from a Binomial distribution. Accounting for the observed dispersion in credit accumulation rates then requires substantial heterogeneity in the probabilities with which students pass courses. The Binomial distribution further implies that graduation prospects, the probabilities of passing enough courses for graduation, increase sharply in the course passing probability. Hence, the model implies that a large fraction of students face very high or very low graduation prospects (subsection 2.3).

Clearly, our main finding that graduation outcomes are highly predictable depends crucially on how we model credit accumulation. Unfortunately, we cannot draw on much prior research to motivate our modeling choices.[6] We view the finding that our model accounts for a broad range of college related observations, including the dispersion and persistence of credits, as providing indirect support.

We investigate why the model implies that students' ability signals are very precise. To do so, we recalibrate the model while fixing the precision of the ability signal at a lower level. Since we assume that high school GPAs are part of students' information sets, this also lowers the correlation between abilities and GPAs. The model then implies that high GPA students are too similar to low GPA students in terms of their credit accumulation rates, their college outcomes, and their lifetime earnings. Increasing signal noise also raises the option value of enrolling in college (the value of learning increases). As a result, students remain in college longer than in the data.

[5] This finding is in line with other research based on NLSY79 data (Carneiro and Heckman, 2002).

[6] Eckstein and Wolpin (1999) model credit accumulation in high school. Garriga and Keightley (2007) postulate a credit production function that is multiplicative in ability, registered credits, and effort. However, they lack the transcript data needed to validate the assumed functional form.

Policy implications. We take a first step towards exploring the policy implications of highly predictable college outcomes (section 6). We study the welfare gains generated by two stylized policy interventions. The first intervention provides insurance against the consumption risks associated with either dropping out of college or taking a long time to earn a degree. The second intervention provides students with information about their abilities and the associated graduation prospects before the college entry decision is made.[7] In both cases, we find that the welfare gains are small for a large fraction of college entrants, especially for those who face little uncertainty about their graduation outcomes.

1.1 Related Literature

This paper relates to a large literature that studies which individual characteristics predict whether a student will be successful in college. Most studies estimate probit or logit models, based on administrative data for a single college.[8] The focus of this literature is very different from ours. Its goal is to inform the design of college policies aimed at improving retention rates.

A growing literature studies structural models of college choice with dropout risk.[9] Our main departure from this literature is to model in more detail how students progress through college. In this respect our approach resembles Garriga and Keightley (2007) and Trachter (2014). Both study models where graduating from college requires a fixed number of earned credits. We add to their analyses by introducing transcript data, which we argue are central for measuring students' graduation chances (section 2). Compared with most previous studies, we also have access to more detailed data on college costs, scholarships, loans, parental transfers, and earnings in college.

Stange (2012) also uses transcript data to estimate a model of risky college completion. However, in his model, the transcript data play a very different role. College grades affect

[7] Actual policies that pursue these goals include income contingent student loans (Chapman, 2006) and dual enrollment programs (Stephanie Marken et al., 2013).

[8] Recent examples include Glynn et al. (2011) and Lin et al. (2012). Astin (1997) and Light and Strayer (2000) use data for many colleges. See Reason (2009) for a survey.

[9] Examples include Akyol and Athreya (2005), Altonji (1993), Athreya and Eberly (2013), Castex (2011), Caucutt and Kumar (2003), Chatterjee and Ionescu (2012), Johnson (2013), and Keane and Wolpin (1997). Since we are unable to determine how much college completion risk these papers imply, we cannot compare our findings with theirs.

the utility derived from studying, rather than the students' chances of completing the requirements for earning a degree. As in much of the literature, all students in Stange's model can earn a degree by staying in college for 4 years. By contrast, our model highlights that students differ in their abilities to satisfy the requirements for earnings a college degree. We use transcript data to measure student abilities that affect their progress through college and thus their chances of graduation.

2 Transcript Data and Graduation Prospects

2.1 Data Description

This section presents the transcript data used to measure how college students progress towards earning a bachelor's degree. We obtain college transcripts from the Postsecondary Education Transcript Study (PETS), which is part of the High School & Beyond dataset (HS&B; see United States Department of Education. National Center for Education Statistics 1988).[10] The data cover a representative sample of high school sophomores in 1980. Participants were interviewed bi-annually until 1986. In 1992, postsecondary transcripts from all institutions attended since high school graduation were collected.

We retain all students who report sufficient information to determine the number of college credits attempted and earned, the dates of college attendance, and whether a bachelor's degree was earned. HS&B also contains information on college tuition, financial resources, parental transfers, earnings in college, and student debt, which we use to calibrate the structural model presented in section 3. Appendix A provides additional details.

2.2 Credit Accumulation Rates

To measure students' progress towards earning a college degree, we focus on the number of completed college credits. Since the data we present are not commonly used in economics, we first present summary statistics.

Table 1 shows the distribution of earned credits at the end of each of the first four years in college. The median student earns around 30 credits in each year. Given that students

[10]The data construction follows Hendricks and Leukhina (2014).

Table 1: College Credits by Year

Group	Year 1	Year 2	Year 3	Year 4
20th percentile	17	41	68	100
50th percentile	28	57	87	119
80th percentile	33	66	98	130
College dropouts	21	43	60	77
College graduates	31	60	90	119

Notes: The table shows the number of credits earned at the end of each year in college by students in the 20^{th}, 50^{th}, and 80^{th} percentile of the credit distribution. The bottom panel shows the average number of credits earned by students who eventually drop out or graduate from college.

Source: High School & Beyond.

graduate, on average, with 125 credits, the median student is on track to graduate in 4 or 5 years.

A key feature of the data is substantial *heterogeneity* in credit accumulation rates. In the first year, students in the 80^{th} percentile earn nearly twice as many credits compared with students in the 20^{th} percentile. Over time, this ratio declines. One likely reason is that the least successful students drop out.

The bottom panel of Table 1 shows that credit accumulation rates are strongly associated with college outcomes. Students who eventually graduate from college earn around 50% more credits than do college dropouts. The gap is fairly stable over time. After three years in college, dropouts have earned fewer than half of the roughly 125 credits required for graduation. These findings suggest that students' inability to complete the requirements for a bachelor's degree may be an important reason for dropping out of college.[11]

Table 2 studies the distribution of college credits earned at the end of year 2 in more detail.

[11] Lacking detailed data on study effort, we are unable to distinguish whether students with low GPAs lack cognitive skills or whether they find studying unpleasant and therefore choose low study effort. For assessing the predictability of college graduation, the distinction is not important. The data collected by Babcock and Marks (2011, Table 6) suggest that differences in study time between students with high and low SAT scores are small.

Table 2: College Credits at the End of Year 2

	Credit distribution			Median credits		Fraction
GPA quartile	20th	50th	80th	CD	CG	graduating
1	21	38	57	32	57	10.7
2	36	50	62	46	57	24.9
3	37	55	64	44	58	50.8
4	50	61	68	45	62	73.6
All	41	57	66	44	60	52.5

Notes: The table shows the distribution of credits earned at the end of the second year in college. Students are divided into quartiles according to their high school GPAs. "Fraction graduating" is the fraction of college entrants that earns a bachelor's degree.

Source: High School & Beyond.

The data reveal that credit accumulation rates and college graduation rates are strongly related to high school GPAs. Students in the top GPA quartile earn 50% more credits than do students in the bottom quartile. They are also 7 times more likely to earn a bachelor's degree. This suggests that cognitive skills may be a strong predictor of college outcomes.

Even within GPA quartiles there is substantial heterogeneity in credit accumulation rates. In the lower quartiles, the 80^{th} percentile of credits earned is around double that of the 20^{th} percentile. This suggests either that GPAs are noisy measures of skills or preparation or that passing courses requires not only skills but also luck.

One observation that points towards an important role of skills rather than luck is the substantial *persistence* of individual credit accumulation rates over time. We construct two measures of persistence. First, the correlation between accumulation rates in consecutive years (computed for all students who are enrolled in both years, averaged over the first 3 years in college) is 0.43. Second, we construct transition matrices for quartiles of credits earned in t and $t+1$. The average of the second largest eigenvalues of these transition matrices is 0.47.

2.3 Graduation Prospects

We now discuss how data on credit accumulation help to estimate the distribution of students' graduation probabilities. The central idea of this paper is that credit accumulation rates contain information about students' abilities or college preparation. These in turn affect their chances of completing requirements for a degree. The structural model of section 3 formalizes this idea and quantifies its implications.

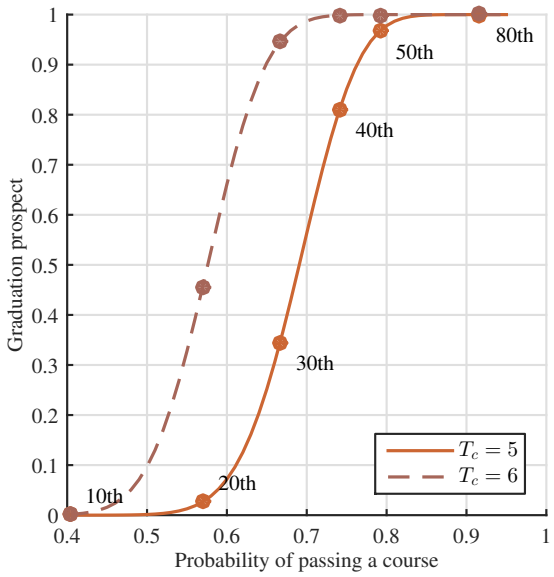
To gain intuition, we study the relationship between credit accumulation rates and graduation chances in the context of a simple model. Students enter college as freshmen with $n = 0$ earned courses. They attempt n_c courses per year but succeed in only a subset. The probability of passing each course p is determined by each student's ability, which is constant over time. Course outcomes are independent, so that the number of courses passed at the end of year t is given by a Binomial distribution with parameters $n_c t$ and p. Students graduate when they pass n_{grad} courses. Students who fail to graduate within T_c years must drop out of college.

Conditional on staying in college for the permitted T_c years, a student's graduation probability is given by $g(p) = \Pr(n_{T_c+1} \geq n_{grad}|p)$. We call $g(p)$ the student's *graduation prospect*.

In the structural model, we find that graduation prospects are the main determinant of individual graduation probabilities. It is therefore of interest to calculate the distribution of $g(p)$ in the simple model. To do so, we set the following parameter values. Since, in HS&B data, 95% of students exit college by year 6, we set $T_c = 6$. Assuming that each course yields 3 credits, it takes $n_{grad} = 42$ courses to earn the 125 credits for graduation. Finally, we set $n_c = 12$, so that students who pass all of their courses earn the number of credits earned by students in the 90^{th} percentile of the observed credit distribution.

Figure 1 summarizes the distribution of graduation prospects implied by the model. The probability of graduating (passing n_{grad} courses) within 5 or 6 years rises sharply with the course passing rate p. Students who pass only 54% of their courses, such as students in the lowest test score quartile or college dropouts in the data, have essentially no chance of graduating in 5 years. Even after 6 years in college, the probability of earning at least 125 credits is only 25%. By contrast, students who pass more than 80% of their courses, such as students in the highest test score quartile or college graduates in the data, are virtually guaranteed graduation in 5 years.

Figure 1: Graduation Prospects and Credit Accumulation Rates



Notes: The figure shows how graduation prospects vary with credit accumulation rates in the Binomial model described in the text. The circles indicate credit accumulation rates at various percentiles of the empirical distribution.

Figure 1 also indicates the graduation prospects for values of p that match select percentiles of the distribution of earned credits shown in Table 2.[12] The model implies large heterogeneity in graduation prospects across college entrants. While students above the median are virtually guaranteed to graduate if the persist for 6 years, students in the bottom decile of the distribution have essentially no chance of graduating. Students in bottom quartile may graduate, but cannot expect to do so in fewer than 6 years.

This heterogeneity in graduation prospects is of central importance in our analysis. It underlies the strong predictability of college graduation in our structural model (section 3). The simple model developed in this section clarifies how the distribution of graduation prospects is identified. Replicating the large heterogeneity in earned credits observed in the data, credit accumulation rates p must vary significantly across students. Given that

[12]Specifically, if a percentile group passes n credits by the end of year 2, the figure shows the graduation prospect associated with $p = n/(2n_c)$. Of course, part of the dispersion in credit accumulation rates is due to luck rather than ability heterogeneity. The structural model of section 3 accounts for this.

students can attempt a large number of courses before they must exit college, graduation prospects increase sharply with p. It follows that graduation prospects are very poor for students with in the bottom of the p distribution and very strong for students in the top of the p distribution. Moreover, replicating the large gaps in credit accumulation rates between college graduates and dropouts requires that the two groups differ significantly in their mean p's and therefore also in their graduation prospects.

3 The Model

To estimate the distribution of graduation probabilities among college entrants, we develop a structural model of college choice.[13]

3.1 Model Outline

We follow a single cohort, starting at the date of high school graduation ($t = 1$), through college (if chosen), and work until retirement. When entering the model, each high school graduate goes through the following steps:

1. The student draws an ability a that is not observed until he starts working. Ability captures all persistent characteristics that make students more successful at school and in the labor market.

2. The student draws a type j and a persistent financial shock ζ_1. These jointly determine his initial assets, college costs, parental transfers, and the student's beliefs about his ability.

3. The student chooses between attempting college or working as a high school graduate.

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

1. He decides how much to work v_t, consume c_t, and save k_{t+1}, subject to a borrowing constraint.

[13]The model and its parameterization extend those in Hendricks and Leukhina (2014) where we study ability selection and the return to college.

2. He attempts n_c college credits and succeeds in a random subset, which yields n_{t+1}. Based on the information contained in n_{t+1}, the student updates his beliefs about a.

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

An agent who enters the labor market in period t learns his ability a. He then chooses a consumption path to maximize lifetime utility, subject to a lifetime budget constraint that equates the present value of income to the present value of consumption spending. Agents are not allowed to return to school after they start working.

The details are described next. Our modeling choices are discussed in subsection 3.5.

3.2 Endowments

Agents enter the model at high school graduation (age $t = 1$) and live until age T. At age 1, a person is endowed with $n_1 = 0$ completed college credits and with random draws of ability a, type j, and financial shock ζ_1.

Learning abilities $a \in \{\hat{a}_1, ..., \hat{a}_{N_a}\}$ determine productivity in school and at work. We normalize $\hat{a}_1 = 0$ and order abilities such that $\hat{a}_{i+1} > \hat{a}_i$. The student's type $j \in \{1, ..., J\}$ determines $\left(\hat{m}_j, \hat{k}_j, \hat{q}_j, \hat{z}_j\right)$. $\hat{m}_j$ is a noisy signal of a. The agent knows the probability distribution of a given $\hat{m}_j$. $\hat{k}_j \geq 0$ denotes financial assets. $\hat{q}_j$ is the permanent component of college costs. We think of this as capturing tuition, scholarships, grants, and other costs or payoffs associated with attending college. $\hat{z}_j$ denotes parental transfers that are received during the first 6 periods after high school graduation, regardless of college attendance.

The financial shock $\zeta_t \in \{1, ..., N_f\}$ evolves according to a Markov chain with transition matrix Π. The student's college costs, parental transfers, and college earnings in each period are jointly determined by (ζ_t, j). The distribution of endowments is specified in section 4.

3.3 Work

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

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

subject to the budget constraint

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

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

Lifetime earnings are a function of ability a, schooling s and college credits n_τ. A worker with ability $a = \hat{a}_1 = 0$ and no credits earns $\exp\left(y_s\right)$. Each college credit increases lifetime earnings by $\mu > 0$ log points. This may reflect human capital accumulation. A unit increase in ability raises lifetime earnings by ϕ_s. If $\phi_{CG} > \phi_{HS}$, high ability students gain more from obtaining college degrees than do low ability students. This may be due to human capital accumulation in college or on the job, as suggested by Ben-Porath (1967). We impose $y_{CD} = y_{HS}$ and $\phi_{CD} = \phi_{HS}$ to ensure that attending college for a single period without earning credits does not increase earnings simply by placing a "college" label on the worker. The return to attending college without earning a degree is captured by μn_τ.

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

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

then given by

$$V_W(k_\tau, n_\tau, j, s, \tau) = \sum_{\iota=1}^{N_a} \Pr(\hat{a}_\iota | n_\tau, j, \tau) V(k_\tau + Z_{j,\tau}, n_\tau, \hat{a}_\iota, s, \tau), \tag{3}$$

where $Z_{j,\tau}$ denotes the present value of parental transfers received after the agent starts working. Our model of credit accumulation implies that the vector (n_τ, j, τ) is a sufficient statistic for the worker's beliefs about his ability, $\Pr(\hat{a}_\iota | n_\tau, j, \tau)$, which implies that $(k_\tau, n_\tau, j, s, \tau)$ is the correct state vector.

3.4 College

We now describe a student's progress through college. Consider an individual of type j who has decided to study in period t. He enters the period with assets k_t, financial shock ζ_t, and n_t college credits. In each period, the student attempts n_c credits and completes each with probability

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

We assume $\gamma_1, \gamma_2 > 0$, so that the probability of earning credits increases with ability.[14] Based on the number of completed credits, n_{t+1}, the student updates his beliefs about a. Since n_t is drawn from the Binomial distribution, it is a sufficient statistic for the student's entire history of course outcomes. It follows that his beliefs about a at the end of period t are completely determined by n_{t+1} and j.

While in college, students derive utility from consumption c_t and leisure $1 - v_t$. Their sources of funds are labor earnings $w_{coll} v_t$, interest on their assets (or debts) Rk_t, and parental transfers $\hat{z}_j$. They purchase consumption and college related items $q(\zeta_t, j)$ (including tuition). Hence, the budget constraint is given by

$$k_{t+1} + c_t + q(\zeta_t, j) = Rk_t + \hat{z}_j + w_{coll} v_t. \tag{5}$$

Work hours v are chosen from a discrete set $\Omega_\zeta \subset \{v_1, ..., v_{N_w}\}$ that depends on the financial shock ζ. The idea is that students receive either part time or full time work offers, which determine their feasible hours.[15] Financial shocks also affect college costs. The value of

[14] Allowing $p(a)$ to be an unrestricted function of a does not change any of our findings; see subsection 6.3.

[15] subsection 6.3 explores an alternative specification where financial shocks affects wages rather than feasible hours.

being in college at age t is then given by

$$V_C\left(n_t, k_t, \zeta_t, j, t\right) = \max_{c_t, v_t \in \Omega_\zeta, k_{t+1} \geq k_{min}} u\left(c_t, 1 - v_t\right) \tag{6}$$

$$+ \beta \sum_{n_{t+1}} \Pr\left(n_{t+1} | n_t, j, t\right) \sum_{\zeta_{t+1}} \Pi_{\zeta_t, \zeta_{t+1}} V_{EC}\left(n_{t+1}, k_{t+1}, \zeta_{t+1}, j, t+1\right) \tag{7}$$

subject to the budget constraint (5). k_{min} is a borrowing limit. $\Pr\left(n_{t+1} | n_t, j, t\right)$ denotes the probability of having earned n_{t+1} credits at the end of period t. This is computed using Bayes' rule from the students' beliefs about a. V_{EC} denotes the value of entering period t before the decision whether to work or study has been made. It is determined by the discrete choice problem

$$V_{EC}\left(n_t, k_t, \zeta_t, j, t\right) = \mathbb{E} \max\left\{V_C\left(n_t, k_t, j, t\right) - \pi\eta_c, V_W\left(k_t, n_t, \zeta_t, j, s\left(n_t\right), t\right) - \pi\eta_w\right\}, \tag{8}$$

where η_c and η_w are independent draws from a demeaned standard type I extreme value distribution with scale parameter $\pi > 0$.[16] $s\left(n\right)$ denotes the schooling level associated with n college credits (CG if $n \geq n_{grad}$ and CD otherwise). The implied choice probabilities and value functions have closed form solutions.[17]

In evaluating V_{EC} three cases can arise:

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

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

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

At high school graduation ($t = 1$), each student chooses whether to attempt college or work as a high school graduate. The agent solves

$$\max\left\{V_C(0, \hat{k}_j, j, 1) - \pi_E\eta_c,\ V_W(\hat{k}_j, 0, \zeta_1, j, HS, 1) - \pi_E\eta_w\right\}, \tag{9}$$

where η_c and η_w are two independent draws from a demeaned standard type I extreme value distribution with scale parameter $\pi_E > 0$.

[16] Allowing the preference shocks to be correlated with ability signals does not materially change the results; see subsection 6.3.

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

3.5 Discussion of Model Assumptions

Our model allows for the possibility that students do not perfectly observe their abilities before making college entry decisions. This captures the idea that *college is experimentation* (Manski, 1989). At high school graduation, students are uncertain about their aptitudes. This information is gradually revealed by the students' performance in college. Stinebrickner and Stinebrickner (2012) present survey evidence consistent with this idea. Abstracting from this feature would bias our results in favor of highly predictable college outcomes.

The discussion of section 2 suggests that the most important modeling choices relate to *credit accumulation*. We propose a simple model and show in subsection 4.4 that it accounts for a range of observations. The main obstacle to studying a richer model is tractability. For example, if the number of credits earned in period t depended either on study effort or on the number of credits earned in previous periods, students' beliefs about their abilities would depend on the entire history of course outcomes. In effect, we would have to abandon Manski's idea of college as experimentation, which would raise the question whether we biased our findings in favor of high predictability.

4 Setting Model Parameters

We calibrate the model parameters to match data moments for men born around 1960.[18] The model period is one year. Our main data sources are PETS and High School & Beyond (HS&B; described in subsection 2.1). We estimate lifetime earnings from National Longitudinal Surveys (NLSY79) data.

The NLSY79 is a representative, ongoing sample of persons born between 1957 and 1964 (Bureau of Labor Statistics; US Department of Labor, 2002). We collect education, earnings and cognitive test scores for all men. We include members of the supplemental samples, but use weights to offset the oversampling of minorities (see Appendix B for details). We use data from the Current Population Surveys (King et al., 2010) to impute the earnings of older workers (see Appendix C for details).

[18] The calibration strategy extends Hendricks and Leukhina (2014) to incorporate financial shocks.

4.1 Distributional Assumptions

Our distributional assumptions allow us to model substantial heterogeneity in assets, ability signals, and college costs in a parsimonious way. Here we summarize our modeling choices, relegating details to Appendix D.

We set the number of types to $J = 200$. Each type has mass $1/J$. We assume that the marginal distributions are given by $\hat{q}_j \sim N\left(\mu_q, \sigma_q^2\right)$, $\hat{z}_j \sim \max\left\{0, N\left(\mu_z, \sigma_z^2\right)\right\}$, $\hat{k}_j \sim \max\left\{0, N\left(\mu_k, \sigma_k^2\right)\right\}$, and $m \sim N(0,1)$. To capture the fact that transfers and assets are non-negative with a mass at 0, we set negative draws of $\hat{z}_j$ and $\hat{k}_j$ to 0. Aside from this truncation, we assume that these endowments are drawn from a joint Normal distribution.

The ability grid $\hat{a}_i$ approximates a Normal distribution with mean $\bar{a}$ and variance 1. Each of the $N_a = 9$ grid points has the same probability, $\Pr\left(\hat{a}_i\right) = 1/N_a$. For notational convenience, we normalize $\bar{a}$ such that $\hat{a}_1 = 0$. Abilities are correlated with ability signals. High school GPAs are noisy measures of students' ability signals: $IQ = \left(\alpha_{IQ,m} m + \varepsilon_{IQ}\right) / \left(\alpha_{IQ,m}^2 + 1\right)^{1/2}$ with $\varepsilon_{IQ} \sim N\left(0,1\right)$. This implies that the agents know more about their abilities than we do. All endowment correlations are calibrated.

Financial shocks are specified as follows. In each period, students receive either part time wage offers, which allow them to work at most 20 hours per week, or full time wage offers, in which case all work states are admissible. The transition matrix for shocks to the feasible set of work hours is given by $\begin{bmatrix} p_v & 1-p_v \\ 1-p_v & p_v \end{bmatrix}$. In addition, students receive i.i.d. college cost shocks that are drawn from two equally likely states: $q\left(\zeta, j\right) = \hat{q}_j \pm \Delta q$. Since parental transfers are highly persistent in the data, we abstract from shocks to z. This yields $N_f = 4$ states with obvious transition probabilities. We assume that all states have equal probability at age 1.[19]

4.2 Mapping of Model and Data Objects

We discuss how we conceptually map model objects into data objects. Variables without observable counterparts include abilities, ability signals, consumption, initial assets, and

[19] If we calibrate the persistence of cost shocks, the algorithm chooses no persistence. We also experimented with shocks to the wage rate rather than admissible work hours. This does not substantially change the findings (see subsection 6.3).

preference shocks. We use the Consumer Price Index (all wage earners, all items, U.S. city average) reported by the Bureau of Labor Statistics to convert dollar figures into year 2000 prices.

College. Students are classified as attending college if they attempt at least 9 non-vocational credits in a given year, either at 4-year colleges or at academic 2-year colleges. Students who earn 2-year college degrees are treated as dropouts, unless they transfer to 4-year colleges where they earn bachelor's degrees.[20] The returns to earning 2-year degrees are captured by the effect of credits on lifetime earnings, μn.[21] The fact that low test score students tend to enroll in 2-year colleges is reflected in their lower average tuition costs. Students attending vocational schools (e.g., police or beauty academies) are classified as high school graduates.

We measure $n_{t+1}/(tn_c)$ as the number of completed college credits by the end of college year t divided by the number of credits taken, assuming a full course load, which is defined as the 90th percentile of the number of credits earned (36 credits per year). 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. In the HS&B data, we divide students into quartiles according to their high school GPAs. In the NLSY79 data, we use their 1989 Armed Forces Qualification Test scores (NLSY79) instead. 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). Since Borghans et al. (2011) show that both measures are highly correlated, we treat GPA quartiles and AFQT quartiles as equivalent. Sidestepping the question of what test scores measure (see Flynn 2009), we use the term "test scores" in the text and the symbol IQ in mathematical expressions.

[20] Trachter (2014) studies the role of 2-year colleges as stepping stones towards bachelor's degrees.

[21] In HS&B data, only 12% of students who enter 2-year institutions earn a 2-year degree.

Financial variables. We interpret annual college costs q as collecting all college related payments that are *conditional* on attending college. In HS&B data, we measure tuition and fees net of scholarships and grants. We set q equal to the average of these values over the first two years in college plus $987 for other college expenditures, such as books, supplies, and transportation.[22] q does not include room and board, which are included in consumption.

Parental transfers z represent transfers received from parents in the 6 years that follow high school graduation. While HS&B does not report initial assets k_1, it does report student debt (negative values of k_t). $w_{coll}v$ represents labor earnings received at any time during the calendar year, including summer terms.

4.3 Model Parameters

Fixed parameters and functional forms.

1. The period utility function during work is given by $\bar{u}(c) = c^{(1-\theta)}/(1-\theta)$.
2. Utility in college is given by $u(c, 1-v) = \delta c^{(1-\theta)}/(1-\theta) + \rho \ln(1-v)$. The parameter $\delta \in (0, 1]$ reduces the marginal utility of consumption while in college. It is needed to account for the low consumption expenditures of college students implied by the financial data. The baseline model sets $\theta = 1$ (log utility). subsection 6.3 explores the implications of stronger risk aversion.
3. The discount factor is $\beta = 0.98$.
4. Based on McGrattan and Prescott (2000), the gross interest rate is set to $R = 1.04$.
5. Motivated by the fact that, in our HS&B sample, 95% of college graduates finish college by their 6th year, we set the maximum duration of college to $T_c = 6$.
6. Each model course represents 2 courses in the data. The number of courses needed to graduate is set to $n_{grad} = 21$ (125 data credits). In each year, students attempt $n_c = 6$ courses (36 data credits).

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

7. Work time: Students can choose from $N_w = 5$ discrete work hour levels in the set $\{0; 10; 20; 30; 40\}$. In setting the choice set for v, we start from an annual time endowment of 5824 hours (52 weeks with 16 hours of discretionary time per day). Based on Babcock and Marks 2011, we remove 35.6 hours of study time for 32 weeks, covering the fall and spring semesters, arriving at a time endowment net of study time of 90 hours per week. Given that v equals work time divided by time endowment, this implies $v \in \{0.00; 0.11; 0.22; 0.33; 0.44\}$.

8. Earnings in college: We set w_{coll} to the mean hourly wage earned by college students of \$7.60.

9. Assets: While in college, students can choose from $N_k = 12$ discrete asset levels. For each type j, the asset grid linearly spans the interval $\left[k_{min}, \hat{k}_j\right]$.

10. Borrowing limits are set to approximate those of Stafford loans, which are the predominant form of college debt for the cohort we study (see Johnson 2013). Until 1986, students could borrow \$2,500 in each year of college up to a total of \$12,500. Converting into year 2000 prices implies $k_{min} = -\$19,750$.

Calibrated parameters. The remaining 28 model parameters are jointly calibrated to match the target data moments summarized in Table 3. We show the data moments in subsection 4.4 where we compare our model with the calibration targets.

For each candidate set of parameters, the calibration algorithm simulates the life histories of $100,000$ individuals. It constructs model counterparts of the target moments and searches for the parameter vector that minimizes a weighted sum of squared deviations between model and data moments.[23]

Table 4 shows the values of the calibrated parameters. The first block of the table reports the parameters governing the joint distribution of endowments. Table 5 shows the implied endowment correlations. We highlight two features that are important for the predictability of college outcomes.

[23] Within each block of moments, such as the fraction of students who drop out of college by test score quartile and year in college, deviations are weighted by the inverse standard deviations of the data moments or, if this is not available, by the square root of the number of observations used to compute each data moment.

Table 3: Calibration Targets

Target	Value
Fraction in population, by (test score quartile, schooling)	Figure 2
Lifetime earnings, by (test score quartile, schooling)	Table 16
Dropout rate, by (test score quartile, t)	Figure 12
Average time to BA degree (years)	4.4
College credits	
Mean cumulative credits, by (graduation status, t)	Table 6
– by (test score quartile, t)	Table 6
Persistence of credits across years	Table 7
CDF of cumulative credits, by t	Figure 8
– by (graduation status, t)	Figure 9, Figure 10
– by (test score quartile, t)	Figure 11
Financial moments	
College costs q	Table 17
(mean by test score quartile, dispersion, autocorrelation)	
Parental transfers z	Table 17
(mean by test score quartile, dispersion)	
Earnings in college	Table 17
(mean by test score quartile, autocorrelation)	
Fraction of students in debt, by t	Table 18
Mean student debt, by t	Table 18

Notes: Lifetime earnings targets are based on NLSY79 data. The remaining targets are based on HS&B data.

First, ability signals are very precise. Thus, students can accurately predict their course accumulation rates and graduation prospects before entering college. Second, test scores are noisy measures of abilities. This implies that graduation rates and graduation prospects vary even more with ability than with test scores. Both features contribute to the predictability of college outcomes. We comment on the role played by other parameter values when we present the paper's findings in section 5.

4.4 Model Fit

This section assesses how closely the model attains selected calibration targets.[24] We view the fact that the model accounts for a broad range of observations, including the dispersion and persistence of earned credits, as providing support for our modeling choices.

College credits. Table 6 shows how the model fits the credit accumulation rates displayed earlier in section 2. The model replicates the large gaps in earned credits between college graduates and dropouts and across high school GPA quartiles. Table 7 shows that the model also replicates the observed persistence of credit accumulation rates, but not its decline over time. To conserve space, the distribution of credits is relegated to Appendix E.

Schooling and test scores. Figure 2 breaks down the schooling outcomes by test score quartiles. Test scores are strong predictors of college entry and college completion. 81% of students in the top test score quartile attempt college and 74% of them earn college degrees. In the lowest test score quartile, only 22% of students enter college and only 11% of them earn degrees.[25] One question our model answers is why low ability students attempt college, even though their graduation chances are small.

5 Results

This section examines what the calibrated model implies for the predictability of college outcomes. We summarize predictability by focusing on the fraction of students with graduation probabilities below 20% or above 80%. The idea is that these students can predict

[24] To conserve space, a more detailed assessment of the model's fit is relegated to Appendix E.

[25] Bound et al. (2010)'s Figure 2 documents similar patterns in NLS72 and NELS:88 data.

Table 4: Calibrated Parameters

Parameter	Description	Value
Endowments		
μ_k, σ_k	Marginal distribution of k_1	$36,620; 29,787$
μ_q, σ_q	Marginal distribution of q	5,331; 3,543
μ_z, σ_z	Marginal distribution of z	3,154; 5,542
$\alpha_{m,z}, \alpha_{m,q}, \alpha_{q,z}, \alpha_{a,m}$	Endowment correlations	0.46; -0.04; -0.12; 2.87
$\alpha_{k,m}$	Correlation k_1, m	-0.21
$\alpha_{IQ,m}$	Correlation IQ, m	1.20
Shocks		
Δ_q	q shock (\$)	1,684
p_v	Persistence of employment shock	0.51
π	Scale of preference shocks	1.197
π_E	Scale of preference shocks at entry	0.397
Lifetime earnings		
ϕ_{HS}, ϕ_{CG}	Effect of ability on lifetime earnings	0.155; 0.197
y_{HS}, y_{CG}	Lifetime earnings factors	3.91; 3.95
μ	Earnings gain for each college credit	0.010
Other parameters		
ρ	Weight on leisure	1.264
δ	Weight on consumption	0.612
U_{CD}, U_{CG}	Preference for job of type s	-1.08; -2.46
$\gamma_1, \gamma_2, \gamma_{min}$	Credit accumulation rate $p_(a)$	4.58; 2.10; 0.47

Table 5: Correlation of Endowments

	IQ	a	m	q	z
IQ	1.00				
a	0.67	1.00			
m	0.72	0.92	1.00		
q	-0.13	-0.15	-0.16	1.00	
z	0.27	0.35	0.37	-0.21	1.00
k_1	-0.19	-0.25	-0.27	0.04	-0.06

Table 6: Credit Accumulation Rates (%)

	Year 1		Year 2		Year 3		Year 4	
Group	Model	Data	Model	Data	Model	Data	Model	Data
Dropouts	59.0	57.1 (1.0)	58.8	59.6 (1.0)	58.1	55.6 (0.9)	56.2	53.6 (1.1)
Graduates	84.2	85.4 (0.6)	84.0	83.4 (0.5)	83.8	83.0 (0.4)	83.6	82.3 (0.4)
GPA quartile 1	53.6	48.1 (2.3)	54.7	53.7 (2.3)	55.8	58.1 (2.3)	58.0	62.3 (2.8)
GPA quartile 2	63.6	61.8 (1.6)	65.5	67.6 (1.4)	67.9	69.5 (1.4)	70.4	71.8 (1.5)
GPA quartile 3	71.6	71.0 (1.2)	73.5	71.5 (1.0)	75.4	72.4 (0.9)	77.1	75.3 (0.8)
GPA quartile 4	81.0	81.8 (0.9)	82.1	81.6 (0.7)	83.1	81.7 (0.6)	84.1	82.0 (0.5)

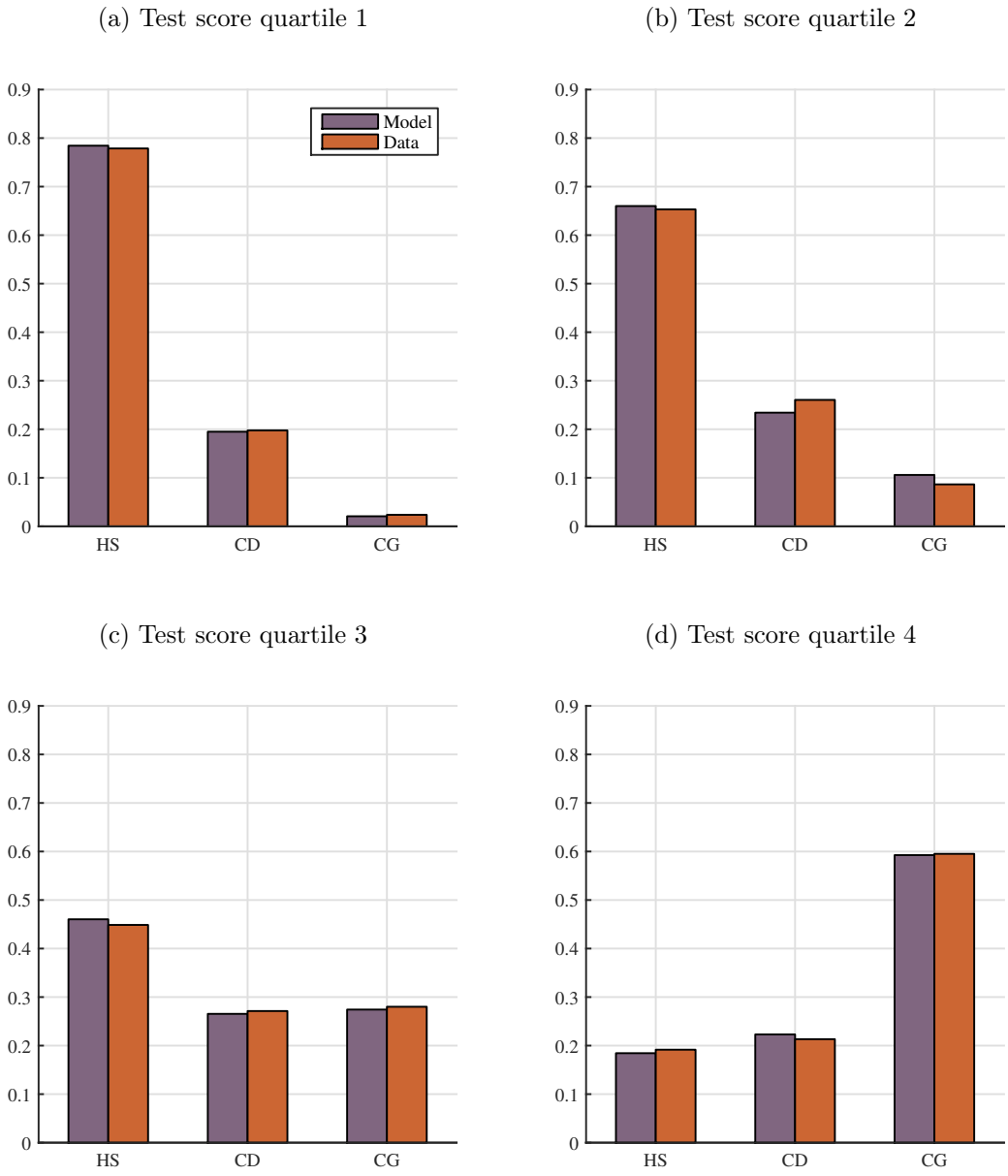
Notes: The credit accumulation rate is the number of college credits completed at the end of each year divided by a full course load. Standard errors are in parentheses.
Source: High School & Beyond.

Table 7: Persistence of Credit Accumulation Rates

	Year $1-2$	Year $2-3$	Year $3-4$
Correlations, model	0.46	0.45	0.43
data	0.48	0.42	0.39
Eigenvalues, model	0.47	0.47	0.47
data	0.51	0.47	0.41
N	1665	1378	1196

Notes: The table compares the persistence of the number of college credits earned implied by the model with the data. "Correlations" refers to the correlation coefficients of credits earned in t and $t+1$. "Eigenvalues" shows the second largest eigenvalues of transition matrices. Each transition matrix indicates the probability that a student in a given quartile of the distribution of credits earned in year t transits to another quartile in $t+1$.
Source: High School & Beyond.

Figure 2: Schooling and Test Scores



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

Source: High School & Beyond.

their graduation outcomes so well that the value of the remaining uncertainty is small. The policy experiments reported in section 6 confirm this intuition.

5.1 Distribution of Graduation Probabilities

Figure 3a shows the CDF of graduation probabilities for the population of college entrants.[26] Graduation is highly predictable for more than half of entrants. 24% of entrants face graduation probabilities below 20% while 30% face graduation probabilities above 80%. For comparison, the figure also shows the distribution for high school graduates who do not enter college. Their graduation probabilities, were these students to enter college, are very poor. 90% of them face graduation probabilities below 20%.

To understand why the model implies high predictability for many students, we study two counterfactual experiments. At high school graduation, students differ in two types of endowments that affect their graduation probabilities: ability signals and financial resources. We quantify the role of both endowments by recomputing the model while shutting down heterogeneity in one of them.

To isolate the role of financial resources, we shut down heterogeneity in ability signals. This involves the following steps. We assign each high school graduate the distribution of abilities $\Pr(a|j)$ that corresponds to the probability distribution of a among baseline college entrants. Financial resources still differ across types. We recompute students' decision rules and simulate their life histories, drawing abilities from the counterfactual $\Pr(a|j)$. We then compute the distribution of college graduation rates in the population of students who enter college in the baseline model.

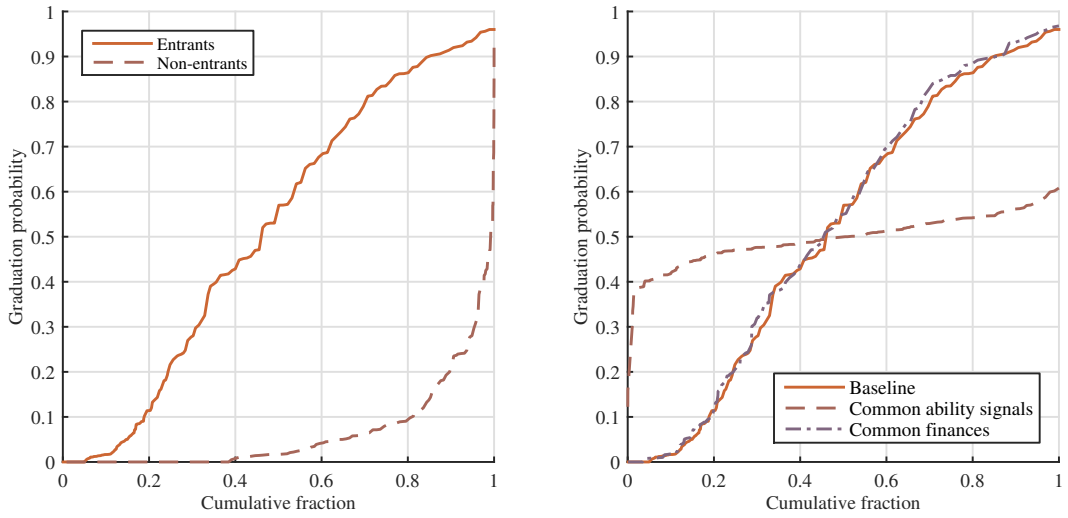
Figure 3b shows that the resulting distribution of graduation probabilities looks markedly different from the baseline model. It has almost no mass below 20% or above 80%. This shows that financial heterogeneity does not play an important role for predicting college outcomes. One reason is that very few students in the baseline model are close to their borrowing limits (Table 18) or close to the maximum number of hours they could work in college.

[26] This is based on the simulated life histories of college entrants. For each type j, we compute the fraction of college entrants that earns a bachelor's degree. The CDF is not smooth because there are only 200 types.

Figure 3: Distribution of Graduation Probabilities

(a) Baseline model

(b) Counterfactuals



Notes: The figure shows the CDF of graduation probabilities, conditional on student's information at the time of high school graduation. Panel(a) shows college entrants and high school graduates who do not enter college. Panel(b) shows two counterfactual experiments for students who enter college in the baseline model: (i) All students share common beliefs about their abilities. (ii) All students are endowed with the same financial endowments (k_1, q, z).

To isolate the role of heterogeneity in ability signals, we shut down financial heterogeneity. This involves the following steps. In the baseline model, we compute the mean financial endowments (k_1, q, z) among college entrants. In the counterfactual, we assign all students these mean financial endowments. The distribution of abilities, $\Pr(a|j)$, remains unchanged. We compute new decision rules and simulate students' life histories. Figure 3b shows that the resulting distribution of graduation probabilities is very similar to the baseline model. Shutting down financial heterogeneity alters graduation outcomes for only 3% of college entrants, compared with 33% in the case where heterogeneity in ability signals is shut down.[27]

These findings imply that heterogeneity in ability signals and therefore graduation prospects is the main reason why graduation outcomes are highly predictable for many students. Next, we explore the link between ability signals and college outcomes. We show that the high predictability of college graduation arises for the following reasons.

1. Graduation prospects differ greatly between high and low ability students. To account for the observed dispersion in credit accumulation rates, the model implies that course passing rates $p(a)$ increase sharply in a. The Binomial distribution of earned courses then implies that graduation prospects differ greatly between high and low ability students (subsection 5.2).

2. Graduation probabilities are closely related to graduation prospects. The incentives for remaining in college until graduation increase sharply with graduation prospects. As a result, high ability students rarely drop out of college. By contrast, many low ability students drop out, even though they could have graduated by staying in college longer (subsection 5.3).

3. Students' ability signals are very precise. Students can therefore accurately predict their course passing rates and graduation prospects (subsection 5.4).

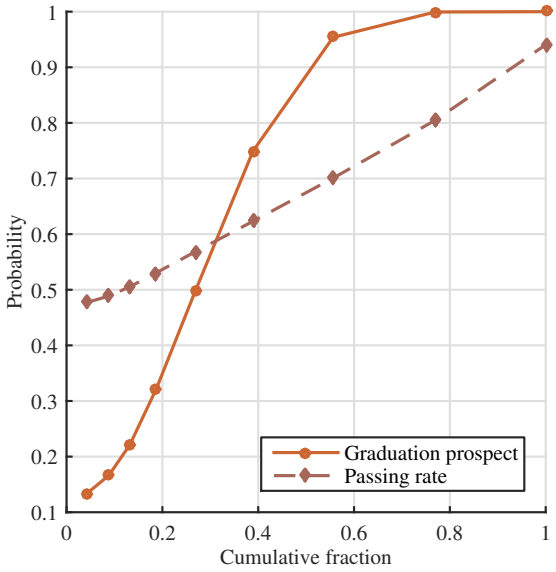
5.2 Distribution of Graduation Prospects

Figure 4a shows the distribution of graduation prospects (the probability of earning 21

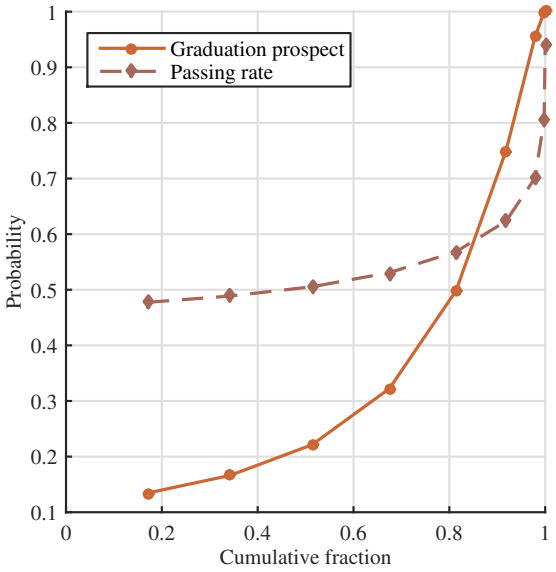
[27] These numbers are calculated from the simulated life histories of students who enter college in both the baseline model and in the counterfactual.

Figure 4: Distribution of Graduation Prospects

(a) College entrants



(b) College non-entrants



Notes: The figure shows the distribution of graduation prospects and course passing rates implied by the baseline model. The graduation prospect is the probability of earning at least n_{grad} courses in T_c years. Panel(a) shows college entrants. Panel(b) shows students who do not enter college.

courses in 6 years) among college entrants.[28] A large fraction of students (61%) faces graduation prospects above 80%. A much smaller fraction (9%) faces graduation prospects below 20%. The distribution of course passing rates is approximately uniform over the interval $[0.50, 0.90]$.

Figure 4b shows the same information for students who do not enter college. It indicates strong selection into college. Few students with strong graduation prospects fail to enter college. The median graduation prospect among non-entrants is only 0.22, suggesting that the inability to graduate is the main friction that prevents college entry.

To understand why the model implies that many students face either very high or very low graduation prospects, recall the discussion of subsection 2.3. This discussion made the following points. (i) Graduation prospects increase sharply with students' credit accumu-

[28] Graduation prospects are based on students' true abilities.

lation rates. (ii) The large dispersion in observed credit accumulation rates then implies that many students face either very high or very low graduation prospects.

A similar reasoning underlies the large dispersion in the graduation prospects generated by the structural model. To gain intuition, consider the distribution of credits earned at the end of the second year in college (see Table 2). Students at the 80^{th} percentiles of the distribution earn 66 credits, compared with only 41 credits for students at the 20^{th} percentile. The binomial distribution of earned courses implies that only a fraction of this dispersion can be accounted for by luck (random course outcomes).

To see this, consider the case where all students share the same course passing rate, given by the mean observed in the data ($p = 0.75$). The standard deviation of earned courses is then given by $[p\,(1-p)\,2n_c]^{0.5} = 2.1$ (6.4 credits). The 20^{th} and 80^{th} percentiles of the distribution are 48 and 60 credits, respectively. Clearly, some heterogeneity in p is needed to account for the observed dispersion of credits.

To get a sense of how much dispersion in p is needed, consider the case where p is uniformly distributed over the interval $[p_L, p_H]$. The values of p_L and p_H that match the observed number of credits earned at the 20^{th} and 80^{th} percentiles of the distribution are 0.50 and 1.00. The implied distribution of course passing rates is therefore quite close to what the calibrated model implies (see Figure 4a).

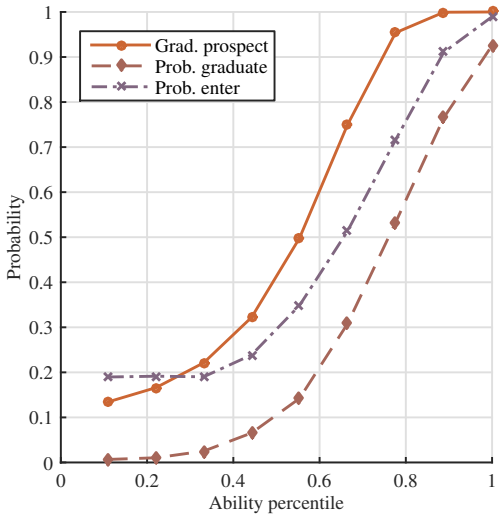
This example clarifies how the model identifies the distribution of graduation prospects. The binomial distribution limits the dispersion of credits that is due to luck. A large part of the dispersion must therefore be generated by heterogeneity in course passing rates and thus graduation prospects. The fact that the model is consistent with the observed persistence of credit accumulation rates over time provides indirect support for the assumed binomial distribution.

5.3 Graduation Prospects and College Outcomes

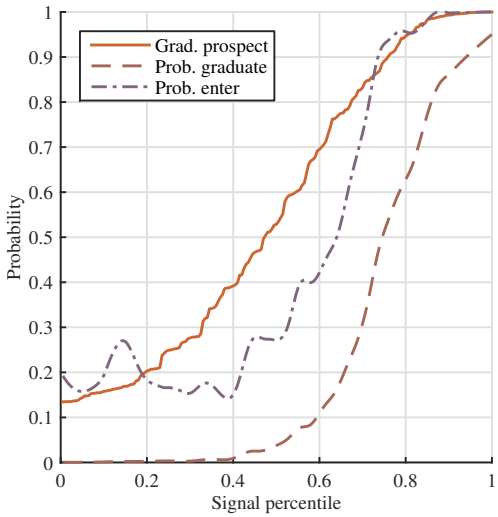
Figure 5a shows that college outcomes are closely related to graduation prospects. Both entry and graduation rates increase sharply with student abilities. While students in the lowest ability quintile have essentially no chance of graduating, students in the highest ability decile graduate with 93% probability. Only students with intermediate abilities face substantial uncertainty about their college outcomes. While 35% of students of median ability attempt college, only 14% of these entrants eventually graduate.

Figure 5: Graduation Prospects and College Outcomes

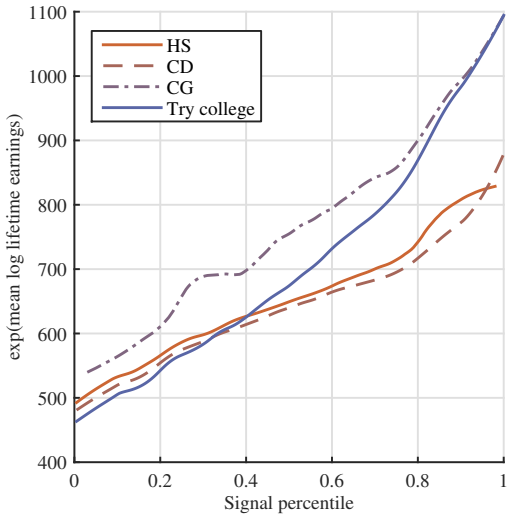
(a) College outcomes



(b) College outcomes and ability signals



(c) Lifetime earnings



Notes: Panel(a) shows the distribution graduation prospects, the fraction of students who enter college, and the probability of graduating conditional on entry. Panel(b) shows the same data conditional on information available to students at high school graduation. Panel(c) shows mean log lifetime earnings, discounted to age 1, by ability signal and schooling level. "Try college" shows the mean log lifetime earnings a student of given signal could expect to earn, if he remained in college as long as possible (until graduation or until year T_c). Lifetime earnings are based on simulated model histories. All lines in panels (b) and (c) are smoothed using quadratic local regressions.

Except for high ability students, graduation probabilities are far lower than graduation prospects. This raises the question why so many students of low to medium abilities drop out, even though many could graduate if they remained in college. To understand this, we consider the financial rewards associated with college attendance.

This is complicated by the fact that students are uncertain about their true abilities. Fortunately, the model implies that ability signals are very precise. The correlation between signals and abilities is 0.92. It follows that inaccurate student beliefs do not play a major role for college entry and dropout decisions. To show this, Figure 5b shows the same information as Figure 5a, except that all probabilities are conditional on information available at high school graduation rather than on true abilities. The similarity of both figures suggests that computing students incentives for college attendance conditional on information available at age 1 does not introduce major inaccuracies.[29]

Figure 5c summarizes the financial rewards associated with college attendance. For each signal, the figure shows the mean log lifetime earnings associated with each possible school outcome (computed from simulated life histories). It also shows how much students can expect to earn, if they stay in college until they either graduate or are forced to drop out (at the end of year T_c). For students with below average ability signals, this counterfactual reduces their lifetime earnings relative to those resulting from their optimal dropout decision. One reason is that these students are not likely to earn enough credits for graduation. A second reason is that sheepskin effects are small for low ability students (because y_{CG} is only slightly larger than y_{HS}).

It follows that, for students of low to medium ability signals, it makes little financial difference whether they remain in college a bit longer or shorter. As a result, these students are easily persuaded to drop out, either because they lack the financial resources to pay for college without working long hours, or because they are hit by shocks while in college. This reduces their graduation probabilities far below their graduation prospects. The situation is very different for students with high ability signals, who forego large earnings gains if they drop out prematurely.

Figure 5c also sheds light on why low ability students rarely enter college. For a given ability signal, working as a high school graduate or as a college dropout yields similar

[29] Among students with low ability signals, variation in college entry rates is largely due to financial endowments (and preference shocks), not to variation in graduation prospects. This is the reason why college entry rates among low signal students are not monotone in m.

lifetime earnings. The earnings gains associated with earning credits approximately offset the costs associated with postponing labor market entry (and tuition). Since low ability students rarely graduate, they have little financial incentive to enter college.

To summarize: In the data, we observe a wide dispersion in credit accumulation rates. Accounting for this requires that some model students pass so many of their courses that their graduation prospects are very strong. These students face strong financial incentives to remain in college until graduation. As a result, their graduation success is highly predictable. On the other hand, accounting for the lowest observed credit accumulation rates requires that some model students pass so few of their attempted courses that their graduation prospects are far from certain. For these students, remaining in college until graduation (or until year T_c) would reduce their lifetime earnings compared with dropping out earlier. Their financial incentives for remaining in college are weak. As a result, their failure to graduate is highly predictable.

5.4 Precision of Students' Ability Signals

The fact that students can accurately predict their graduation prospects plays an important role in the preceeding arguments. To understand why the model implies that signals are precise, we study the model's implications when we impose more signal noise. We do so by calibrating the model while fixing the correlation between signals and abilities at 0.44, compared with 0.92 in the baseline model (i.e., we set $\alpha_{a,m} = 0.5$). Examining which data features this alternative model fails to account for reveals how the precision of the signal is identified.

Table 8 displays selected changes relative to the baseline model. More signal noise reduces the correlation between test scores and abilities from 0.67 in the baseline model to 0.35. The mechanical reason is that test scores are themselves noisy measures of ability signals. As a result, college outcomes, such as credit accumulation rates or dropout rates, vary too little across test score quartiles (see rows 1 and 2 of Table 8).

The calibration algorithm attempts to offset these changes by increasing the correlation between test scores and ability signals. However, given that college entry decisions depend strongly on signals, this leads the model to overstate the association between test scores and college entry decisions (row 3). These findings suggest that the values of signal noise

Table 8: Implications of Higher Signal Noise

	Data	Baseline	Noisy signal
Differences between high/low test score students:			
- credit accumulation rate (year 2)	0.28	0.27	0.10
- college dropout rate	0.63	0.63	0.32
- college entry rate	0.59	0.60	0.76
College entry rate, ability above/below median		0.58	0.28
Log lifetime earnings gap, CD vs HS	0.07	0.06	-0.02
Fraction with low / high graduation probabilities		24 / 30	2 / 3

Notes: The table shows the implications of increasing signal noise relative to the baseline model. Rows 1 to 3 display differences between students in the top and bottom test score quartile. Row 4 shows the difference in the college entry rate between students with abilities above and below the median. Row 5 displays the log lifetime earnings gap between college dropouts and high school graduates. Row 6 shows the fraction of college entrants with graduation probabilities below 20% and above 80%.

and test score noise are mainly determined by data moments that characterize how college entry decisions and college outcomes vary across test score quartiles.

More signal noise also weakens the association between abilities and college entry rates (row 4). One mechanical reason is that students simply do not know their abilities at the time of college entry. A second reason is that the value of attempting college increases for students with low to medium ability signals. As a result, more high ability students work as high school graduates, while more low ability students enter college. This shrinks the lifetime earnings gap between college dropouts and high school graduates below its empirical value (row 5).

The predictability of college graduation declines substantially when signals are less precise (row 6). Virtually no college entrants face graduation probabilities below 20% or above 80%. In part, this is a mechanical consequence of students' uncertainty about their abilities. This is reinforced by a flatter relationship between course passing rates and abilities, $p(a)$. The latter allows more low ability students to graduate and thus prevents the dropout rate from rising too much above the target value.

Summary. The main result of this section is that graduation is highly predictable for a large fraction of college entrants. The primary reason is that many students face either very strong or very poor graduation prospects. Transcript data enable the model to identify the distribution of graduation prospects. Financial heterogeneity is not a major source of predictability. The model also implies that students are able to predict their graduation prospects with high precision before entering college. We demonstrate the robustness of these findings in subsection 6.3.

6 Policy Implications

This section explores policy implications. We ask how much (potential) college students value either insurance against college related risks or information about their graduation prospects (abilities).

Rather than study specific policies, we consider abstract interventions that provide *complete* insurance against college related earnings risks or *perfect* information about ability. These interventions quantify how much students value the uncertainty associated with college attendance. The main finding is that the welfare gains generated by these interventions are small for most students, especially for those facing little uncertainty about their graduation outcomes.[30]

6.1 Insuring Earnings Risk

If college is risky, welfare could potentially be raised by providing insurance against the financial risks associated with either dropping out of college or of requiring a long time to earn a degree. In this section, we ask how much potential college entrants value such insurance.

We consider idealized interventions that provide complete insurance against college related consumption risks without incurring any implementation costs or deadweight losses. In

[30] We expect the welfare gains that can be generated by implementable policies to be lower than those generated by the abstract interventions we study. The main reason is that implementable policies need to guard against moral hazard (students study less or drop out earlier), adverse selection (low ability students enter college, knowing they will drop out), or deadweight losses related to financing the interventions. We leave the study of such policies for future research.

particular, we assume that students do not change their college entry or continuation behavior in response to the intervention. We thereby abstract from the moral hazard problem associated with insurance provision: students have an incentive to drop out of college early, knowing that this has no effect on their earnings.

We study two insurance arrangements. In the first case, the policy maker provides complete consumption insurance during the work phase conditional on a student's endowments a and j. It assigns each worker the average age-consumption profile of workers endowed with the same (a, j) in the baseline model, starting at age $T_c + 1$ when all students have entered the labor market.

This policy offers insurance against uncertain credit realizations and the associated earnings and dropout risks. It also insures students against the consequences of ex-ante financial heterogeneity and financial or preference shocks that lead otherwise identical students to make different entry decisions.

The distribution of welfare gains is summarized in Table 9. Welfare gains are measured by the percentage change in consumption that makes each type indifferent between the baseline model and the insurance counterfactual. The median compensating differential for this intervention amounts to only 0.05% of consumption. This is a direct consequence of the high predictability of college graduation, which limits the uncertainty students face. The average standard deviation of log consumption within (a, j) groups is only 0.04, compared with 0.24 across all types. Limiting attention to baseline college entrants increases the median welfare gain to 0.27%. This happens because the average welfare gains for types with low entry probabilities are small.

The value of college related risks is closely related to graduation risk. To illustrate this point, Figure 6 plots the welfare gains for all (a, j) types against their graduation probabilities. The largest welfare gains occur for groups with intermediate graduation probabilities. Excluding those j types that face little uncertainty about college graduation (their graduation probabilities, conditional on entry, lie between 20% and 80%), increases the median welfare gain to 0.38%.

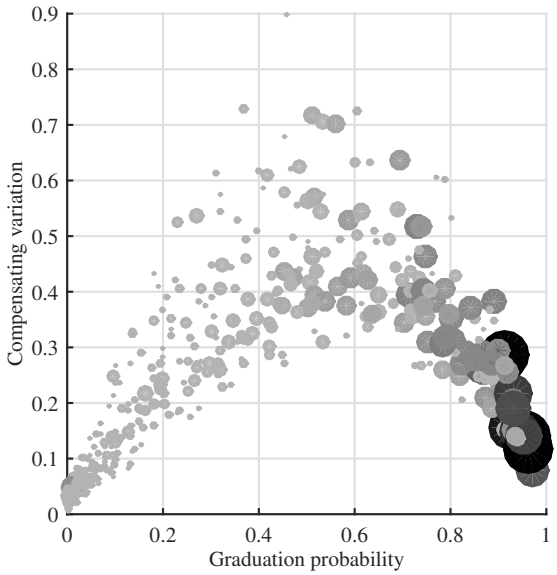
Larger welfare gains can be obtained by also insuring students against uncertainty about their true abilities. This is the second intervention shown in Table 9. For each type j, the intervention replaces the random consumption stream received in the baseline model with its type specific average (again starting at age $T_c + 1$). The median welfare gains

Table 9: Providing Insurance

Insurance within	(a,j) groups		j groups	
Welfare gains	median	mean	median	mean
All	0.05	0.15	0.22	0.40
Entrants	0.27	0.27	0.68	0.65
High risk entrants	0.38	0.38	0.89	0.90

Notes: The table shows the distribution of compensating differentials, expressed as percentages of baseline consumption, implied by two interventions. The interventions assign all college entrants within a type j or within an (a,j) group the group specific mean consumption.

Figure 6: Welfare Gains Due to Insurance Within (a,j) Groups



Notes: Each data point is one (a,j) type. The size of each marker is proportional to the mass of college entrants of each type.

range from 0.22% (all students) to 0.89% (college entrants with significant graduation risk) of baseline consumption. Of course, part of these welfare gains arise from redistribution between students with high and low ability endowments rather than from insurance against college related risks.

When interpreting these welfare gains, the reader should keep in mind that they arise from *complete* consumption insurance during the work phase. By comparison, Chatterjee and Ionescu (2012) find that forgiving the college costs of dropouts increases average welfare by 2.4% of consumption. This welfare gain is much larger, even though consumption insurance is far from complete.

6.2 Providing Information

Manski and Wise (1983) argue that students enter college in part to learn about their abilities. This experimentation is costly. Students forego labor earnings and pay tuition while in college. This argument suggests that welfare could be increased by providing high school graduates with information about their abilities before they make college entry decisions. This is one motivation behind dual enrollment programs that allow high school students take college level courses (Stephanie Marken et al., 2013).

We study the welfare implications of a stylized policy intervention that costlessly informs students about their abilities before they enter college. The details are as follows. For all combinations of (a, j), we solve for the decision rules of an agent who knows his ability. The associated value functions at high school graduations are labeled $V_{HS}^{a}(a, j)$. The expected value of type j before learning his ability is given by $\hat{V}_{HS}(j) = \mathbb{E}_a \{V_{HS}^{a}(a, j) | j\}$. The welfare gain of the intervention is the log change in annual consumption, starting at age 1, that makes each type indifferent between the baseline value $V_{HS}(j)$ and $\hat{V}_{HS}(j)$.

The average welfare gain across all high school graduates amounts to 0.14% of baseline consumption. However, for the majority of students, the gains are essentially zero. The reason is that students' ability signals are quite precise. Many high school graduates have sufficient information about their abilities to make ex post optimal college entry decisions with high probability. Large welfare gains only accrue to medium ability students who face substantial dropout risk.

Restricting attention to students who enter college in the baseline model excludes a large fraction of low ability students who do not value information about their abilities much.

This increases average welfare gains to 0.21% of baseline consumption. These findings highlight the importance of targeting the proper set of students for a policy intervention to generate significant gains.

6.3 Robustness

This section explores the robustness of our findings when selected model features are modified. Specifically, we examine the following modifications:

1. $p(a)$ is an unrestricted function of ability. Since there are only 9 ability levels, $p(a)$ can be specified as a vector of length 9.

2. Financial shocks affect wages rather than feasible work hours. There are 5 financial states, ordered from worst to best. The calibrated parameters are $\Delta q, \Delta w$ and p_{stay}. College costs take on the values $q(\zeta, j) = \hat{q}_j - (\zeta - 3)\Delta q$. Wages are given by $w_{coll} + (\zeta - 3)\Delta w$. The transition probabilities are $\Pr(\zeta' = \zeta|\zeta) = p_{stay}$, $\Pr(\zeta' = 2|\zeta = 1) = \Pr(\zeta' = 4|\zeta = 5) = 1 - p_{stay}$, $\Pr(\zeta' = \zeta \pm 1|\zeta) = 1 - p_{stay}/2$ for interior ζ, and $\Pr(\zeta'|\zeta) = 0$ otherwise.

3. Smaller preference shocks: $\pi = \pi_E = 0.1$.

4. Preference shocks that are correlated with ability signals: We assume that the means of the preference shocks experience in college are drawn from a joint Normal distribution together with the other student endowments (see subsection 4.1). Preferences shocks are correlated with m, but not correlated with other endowments, conditional on m.

5. Higher values of risk aversion (θ).

Table 10 summarizes how the predictability of college graduation and the welfare gains of policy interventions change for each model specification. The only significant changes occur when students are more risk averse than in the baseline model. This reduces the predictability of college graduation and increases the welfare gains of interventions. However, the model's ability to replicate the calibration targets deteriorates for higher values of risk aversion (which is why the baseline model features log utility).

Table 10: Robustness

	Fraction with graduation probability		Median welfare gain		
			Insurance		Known
Model	< 0.20	> 0.80	within j	within (a, j)	ability
Baseline	0.24	0.30	0.22	0.05	0.14
Unrestricted $p(a)$	0.26	0.30	0.21	0.05	0.17
Wage shocks	0.25	0.29	0.21	0.08	0.13
$\pi = \pi_E = 0.1$	0.30	0.31	0.18	0.01	0.05
(η_c, m) correlated	0.24	0.31	0.21	0.04	0.14
$\theta = 1.5$	0.22	0.19	0.44	0.15	0.10
$\theta = 2.0$	0.20	0.18	0.57	0.18	0.10
$\theta = 4.0$	0.19	0.17	1.64	0.84	0.80

Notes: The table explores the robustness of the main findings. Each row shows a model version. The columns show the fraction of college entrants with low or high graduation probabilities and the median welfare gains generated by the policy interventions described in the text.

7 Conclusion

We conclude by considering potential avenues for future research. Since our model captures the distribution of risks and returns associated with entering college, it provides a starting point for the study of college related policies, such as income contingent loans or dual enrollment programs. However, the model abstracts from two features that may be important for policy analysis.

The first feature is study effort. Insurance arrangements distort students' incentives to study. To capture these distortions, it is necessary to model how study effort affects college outcomes. The second feature is earnings risk during the work phase. One motivation of making college loans income contingent is to alleviate the tight budget constraints of young workers who may be borrowing constrained. Modeling earnings shocks would also imply that the model could be used to measure the predictability of lifetime earnings as of the age of high school graduation (Huggett et al., 2011).

References

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

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

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

_ **, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom**, "College attrition and the dynamics of information revelation," 2012. Mimeo. Duke University.

Astin, Alexander W, "How "good" is your institution's retention rate?," *Research in Higher Education*, 1997, *38* (6), 647–658.

Athreya, Kartik and Janice Eberly, "The College Premium, Risk, and Human Capital Investment," 2013. Mimeo, Federal Reserve Bank of Richmond.

Babcock, Philip and Mindy Marks, "The Falling Time Cost of College: Evidence from Half a Century of Time Use Data," *Review of Economics and Statistics*, May 2011, *93* (2), 468–478.

Belley, Philippe and Lance Lochner, "The Changing Role of Family Income and Ability in Determining Educational Achievement," *Journal of Human Capital*, 2007, *1* (1), 37–89.

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

Bettinger, Eric P and Bridget Terry Long, "Addressing the Needs of Underprepared Students in Higher Education Does College Remediation Work?," *Journal of Human Resources*, 2009, *44* (3), 736–771.

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

Bound, John and Sarah Turner, "Dropouts and diplomas: The divergence in collegiate outcomes," *Handbook of the Economics of Education*, 2011, *4*, 573–613.

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

Bowen, William G., Matthew M. Chingos, and Michael S. McPherson, *Crossing the Finish Line: Completing College at America's Public Universities*, Princeton University Press, 2009.

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

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

Castex, Gonzalo, "College Risk and Return," 2011. Mimeo. Central Bank of Chile.

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

Chapman, Bruce, "Income contingent loans for higher education: International reforms," *Handbook of the Economics of Education*, 2006, *2*, 1435–1503.

Chatterjee, Satyajit and Felicia Ionescu, "Insuring Student Loans Against the Financial Risk of Failing to Complete College," *Quantitative Economics*, 2012, *3* (3), 393–420.

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

Flynn, James R, *What Is Intelligence? Beyond the Flynn Effect*, Cambridge University Press, 2009.

Garriga, Carlos and Mark P Keightley, "A General Equilibrium Theory of College with Education Subsidies, In-School labor Supply, and Borrowing Constraints," 2007. Mimeo. Federal Reserve Bank of St Louis.

Glynn, Joseph G, Paul L Sauer, and Thomas E Miller, "A Logistic Regression Model For The Enhancement Of Student Retention: The Identification Of At-Risk Freshmen," *International Business & Economics Research Journal (IBER)*, 2011, *1* (8).

Hendricks, Lutz and Oksana Leukhina, "The Return to College: Selection and Dropout Risk," 2014. Mimeo. University of North Carolina.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *American Economic Review*, 2011, *101*, 2923–54.

Johnson, Matthew T., "Borrowing Constraints, College Enrollment, and Delayed Entry," *Journal of Labor Economics*, 2013, *31* (4), pp. 669–725.

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

King, Miriam, Steven Ruggles, J Trent Alexander, Sarah Flood, Katie Genadek, Matthew B Schroeder, Brandon Trampe, and Rebecca Vick, "Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]," 2010. Minneapolis: University of Minnesota.

Light, A and W Strayer, "Determinants of college completion: school quality or student ability?," *Journal of Human Resources*, 2000, pp. 299–332.

Lin, Tin Chun, William Wei Choun Yu, and Yi Chi Chen, "Determinants and probability prediction of college student retention: new evidence from the Probit model," *International Journal of Education Economics and Development*, 2012, *3* (3), 217.

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

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

_ **and David A Wise**, *College choice in America*, Harvard University Press, 1983.

Marken, Lucinda Gray Stephanie, Laurie Lewis, and John Ralph, "Dual Enrollment Programs and Courses for High School Students at Postsecondary Institutions: 2010-11," Technical Report, National Center for Education Statistics 2013.

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

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

Reason, Robert D, "Student variables that predict retention: Recent research and new developments," *Naspa Journal*, 2009, *46* (3).

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

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

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

Trachter, Nicholas, "Stepping Stone and Option Value in a Model of Postsecondary Education," *Quantitative Economics*, 2014, *forthcoming.*

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

CENTER FOR FINANCIAL STUDIES

CFS Working Paper Series

Recent Issues

All CFS Working Papers are available at *www.ifk-cfs.de*.

No.	Authors	Title
501	Christiane Baumeister, Lutz Kilian	*Understanding the Decline in the Price of Oil Since June 2014*
500	Christiane Baumeister, Lutz Kilian, Thomas K. Lee	*Inside the Crystal Ball: New Approaches to Predicting the Gasoline Price at the Pump*
499	Lutz Kilian	*The Impact of the Shale Oil Revolution on U.S. Oil and Gasoline Prices*
498	Pablo Guerron-Quintana, Atsushi Inoue, Lutz Kilian	*Impulse Response Matching Estimators for DSGE Models*
497	Winfried Koeniger, Julien Prat	*Human Capital and Optimal Redistribution*
496	Maria Bigoni, Gabriele Camera, Marco Casari	*Money is More than Memory*
495	Ronald Bosman, Roman Kräussl, Thomas van Galen	*Emotions-at-Risk: An Experimental Investigation into Emotions, Option Prices and Risk Perception*
494	Roman Kräussl	*Art as an Alternative Asset Class: Risk and Return Characteristics of the Middle Eastern & Northern African Art Markets*
493	Roman Kräussl, Thorsten Lehnert, Nicolas Martelin	*Is There a Bubble in the Art Market?*

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de