On the Distribution of College Dropouts: Wealth and Uninsurable Idiosyncratic Risk*

Ali K. Ozdagli
Federal Reserve Bank-Boston
Ali.Ozdagli@bos.frb.org

Nicholas Trachter
EIEF
nicholas.trachter@eief.it

October 14, 2011

Abstract

We present a dynamic model of the decision to pursue a college degree in which students face uncertainty about their future income stream after graduation due to unobserved heterogeneity in their innate scholastic ability. After matriculating and taking some exams, students reevaluate their expectations about succeeding in college and may decide to drop out and start working. The model shows that, in accordance with the data, poorer students are less likely to graduate and are likely to drop out sooner than wealthier students. Our model generates these results without introducing explicit credit constraints.

*We thank Christopher Foote, Alex Levkov, Andrea Pozzi, and Rob Shimer. We also benefited from comments by seminar participants at the Federal Reserve Bank of Boston, and the 2011 Midwest Macroeconomics Meetings. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Boston or the Federal Reserve System. Sarojini Rao provided excellent research assistance. All mistakes are our own.

1 Introduction

A large fraction of every cohort that enrolls in four-year U.S. colleges drops out before graduating and there is a higher concentration of dropouts among the students from lower-income families. We also observe that students from low-income families tend to drop out earlier than students from high-income families. Given the high return to graduation, the skewed distribution of dropouts generates a channel that perpetuates and exacerbates income inequality.

We present a dynamic model of educational choice to explore the relationship between household wealth and college dropout behavior. In the model, the differences in the students' unobserved innate scholastic ability and the families' initial wealth levels are the driving force behind the high and skewed dropout rate among low-income students. At every period in college, risk-averse students take an exam, the outcome of which provides both human capital and information that can be used to update students' beliefs about their ability level. Given the outcome of the exams and their income level, the students decide optimally if and when to drop out. Therefore, a student's optimal dropout behavior is characterized by the distance between her belief about her ability and a belief threshold at which the student drops out. We show that this threshold depends endogenously on the wealth level of the student's family, therefore providing the link between household wealth and dropout behavior.

Our model incorporates the ideas that investing in a college education is risky because the outcome is uncertain,[1] and that wealthier students are less risk averse.[2] This framework, combined with the learning mechanism, generates the result that poor students are less willing to take the risk associated with the uncertain outcome of college education and that they do not want to continue their education for as long as the rich students in order to learn their ability. Therefore, poor students are less likely to graduate and tend to drop out earlier compared to wealthier students.

[1] Chen (2008) finds that college investment is indeed risky after correcting for selection bias and accounting for permanent and transitory earnings risks.

[2] In our benchmark model, wealthier students are less risk averse because all students have the same constant relative risk aversion (CRRA) preferences. We also generalize our results to hyperbolic risk aversion (HARA) preferences, which include CRRA as a special case.

In order to model the evolution of beliefs about innate ability and the decision to drop out, we take the Miao and Wang (2007) framework of entrepreneurial learning and survival, and extend it with realistic features that are important for the dropout decision faced by college students. In particular, we allow the workers' lifetime wage profile to depend on their experience, measured by the time spent in the labor market, and on their tenure in college, as in Mincer (1974). Unlike Mincer, however, we also let the lifetime wage profile depend on whether the worker has graduated from college and allow the experience premium to interact with the individual's graduation status. Our results are robust to different specifications of the lifetime wage profile for students with different college tenures.

In a typical Mincerian framework, a student would choose the optimal number of years of college education by comparing the marginal gain from an extra year in school with the marginal gain from joining the workforce immediately. Within this framework, a nonlinear relationship between schooling and returns to education can potentially explain why many students decide to drop out even though the returns to earning a four-year degree are high. Still, this framework does not fully explain the data. First, the Mincerian model is silent about the relationship between wealth and educational profiles. Second, when college students are confronted with questions regarding their expectations about postsecondary educational outcomes, almost all of them respond that they intend to obtain four years of college education.[3] The failure of this basic Mincerian model suggests that it is necessary to have a story where information unfolds as time elapses, in order to explain the dropout behavior.

The literature provides three different mechanisms whereby information is revealed over time. The first one involves binding credit constraints. However, credit constraints do not seem to be the only determinant of dropout behavior because the dropout rate for students from the richest households is around 29 percent, which is still very high.[4] This seems to conflict with arguments in favor of credit constraints because if credit constraints were the main reason for dropping out we would expect that almost all the students at the top of the income distribution to graduate, given the high returns to education. Moreover, using NLSY 1997 data, Lovenheim

[3]See the data manual of *National Longitudinal Survey of Youth 1979*.

[4]Own calculations from NLSY79 using top 10 percent of income distribution.

and Reynolds (2010) find that rising house prices lead to higher graduation rates, especially among low-income families. The result implies that the effect of wealth on dropout behavior has not vanished despite the fact that there have been many improvements in financial markets and credit availability to poor students. The second potential mechanism is learning about scholastic taste, that is the student's attitude towards schooling, as in Stange (2011) or Heckman and Urzua (2008). However, there is no clear channel that relates scholastic taste with the wealth level of the family.

The third mechanism, used in this paper, is learning about unobserved ability through college attendance. Our modeling choice follows from the empirical work of Stinebrickner and Stinebrickner (2008) and Stinebrickner and Stinebrickner (2009), who construct a panel study in order to understand the dropout decisions of students in a particular four-year institution, Berea College. Stinebrickner and Stinebrickner (2008) calculate a lower bound on the percentage of attrition that would remain at Berea College even if low-income students were given access to loans, and find that this bound is very high, thus concluding that credit constraints cannot explain the dropout decision for the majority of students. Stinebrickner and Stinebrickner (2009) find that academic performance, a proxy for learning about one's scholastic ability, is a good predictor of dropout behavior, a result that is detrimental to explanations relying on scholastic taste.[5]

We do not claim that the alternative mechanisms discussed above play no role in dropout decisions but rather propose our model as a complementary explanation. For example, although our model does not include explicit borrowing constraints, it implies a natural borrowing limit in the spirit of Aiyagari (1994). Since a student does not know her ability ex-ante when attending college, she faces the risk of ending up in a low-paying job, an event that will cause her to have low wealth, after adjusting for discounted lifetime wage income and the cost of college. Due to Inada conditions, as the student's wealth level approaches zero her consumption as a

[5]Another mechanism that can explain the skewed distribution of dropouts is that richer students choose a longer duration of education when education is a normal consumption good. However, this cannot explain why academic performance matters so much for dropout decisions, as shown in Stinebrickner and Stinebrickner (2009), unless we believe that academic performance affects the marginal utility of education and that its effect on rich and poor students is different.

low-wage worker becomes infinitesimally small, and therefore her marginal utility of consumption post-education goes to infinity. This causes a lower willingness by the student to borrow against her future labor income in order to finance her college education. In Aiyagari's words: "Thus, a borrowing constraint is necessarily implied by non-negative consumption," where the non-negativity of consumption is guaranteed by the Inada conditions in our model.[6]

The next section uses data from the *National Longitudinal Study of the High School Class of 1972,* the *National Longitudinal Survey of Youth 1979* and the *National Longitudinal Survey of Youth 1997* to compare the dropout behavior of poor and rich students. We find that poor students are at least 27 percent more likely to drop out and they do so before rich students, controlling for measures of unobserved ability.

2 Evidence

To motivate our model, this section presents some statistics regarding dropout behavior based on the *National Longitudinal Study of the High School Class of 1972*, NLS-72 hereafter, the *National Longitudinal Survey of Youth 1979*, NLSY79 hereafter, and the *National Longitudinal Survey of the Youth 1997*, NLSY97 hereafter. For the NLS-72, we focus on individuals who enrolled in a four-year college during 1972 with no discontinuities in their education spells after starting college. For the NLSY79, we focus on individuals who enrolled in a four-year college during or after 1979 with no discontinuities in their education spells after college enrollment. For the NLSY97, we focus on individuals who enrolled in a four-year college during or after 1997 with no discontinuities in their education spells after starting college.[7]

[6] The intuition regarding a natural borrowing limit also holds for HARA preferences because the marginal utility of consumption goes to infinity as consumption approaches to a positive limit under these preferences.

[7] We also discarded students that attended two-year community colleges. The separation of community colleges from four-year colleges is important because the salary profile of graduates from both types of institutions is quite different. Within the context of our model, community colleges may serve as a stepping stone to four-year colleges by giving students more information about their skills so that high school graduates that are not optimistic enough to go to a four-year college initially

Both in our data analysis here and theoretical model in the next section we abstract from enrollment decision because we are interested in the dropout behavior of a student who is enrolled in college rather than the counterfactual question of a student's dropout behavior if she were to attend college. For the latter question, our results regarding the differences between poor and rich students can be considered as a lower bound because poor high school graduates are less likely to attend college and we can consider non-attendees as having dropped out before the first day of college. This view is also in accordance with Manski and Wise (1983) which shows that if people with a low probability of attending college were to attend, they would have a high probability of dropping out.

Our analysis follows from comparing dropout profiles of students from rich and poor families. For NLS-72 we use the socioeconomic status of the respondent's family at the moment of high-school graduation as a measure of wealth. For the NLSY79 and NLSY97 such a variable is not available, so we construct it ourselves. In particular, we first rank the respondents' families according to their income at the time the student graduated from high school. Then we match the fraction of families with different socioeconomic status in NLS-72 so that both datasets become comparable.[8]

Table 1 presents some aggregate summary statistics regarding dropout behavior for both rich and poor students. Here we withold students from middle income households to make the point clearer. In the three datasets, students from poor families have higher attrition rates. Furthermore, students from poor households tend to drop out before students with rich families. As shown in the table, they tend to drop a half year to a year sooner than rich students depending on the dataset.

To further explore the skewed distribution of dropouts with respect to wealth, we compare the dropout rates of rich and poor students in different years of college. If poor students tend to drop out earlier than rich students, as we suggest, then a larger proportion of dropouts among poor students should occur in the earlier

may still enroll community colleges. Trachter (2010) formalizes this idea to study the transition between community and four-year colleges.

[8]For students enrolling in college, our sample from NLS-72 has 16.3 percent of families classified as low socioeconomic status, 41.7 percent as average socioeconomic status, and the rest as wealthy.

Table 1 Dropout rates and mean time before dropping out by socioeconomic status of family

	Socio. status[a]	% that drop	Mean tenure in college[b]	st. dev. of tenure
NLSY97	Low	28.98	2.31	1.65
	High	13.40	3.14	1.73
NLSY79	Low	62.5	2.78	1.6
	High	26.96	3.94	1.9
NLS-72	Low	65.6	2.02	1.29
	High	52.86	2.69	1.63

[a] For the NLSY79 and NLSY97 we constructed the measure of socioeconomic status thorugh the income level of the family prior to the respondent's enrollment in college. We choose the deciles so as to match the distribution of socioeconomic status of the NLS-72. [b] We only have the length of the tenure in college for a sub-sample of the population.

college years, whereas a larger proportion of dropouts among rich students should occur in later college years. Table 2 provides a parsimonious way of checking this argument in the data by reporting the ratio of proportion of dropouts among poor students to the proportion of dropouts among rich students in different years of college. This statistic decreases from a number greater than 1 to a number less than 1 as we go from the earlier to later college years, providing support for our claim.

Table 2 Dropout rates of low- vs. high-income students

	tenure between						
	0 and 1 years	1 and 2 years	2 and 3 years	3 and 4 years	4 and 5 years	5 and 6 years	6 and 7 years
NLSY97	1.76	1.65	1.57	0.91	0.37	0.69	0.37
NLSY79	3.28	1.31	1.61	0.8	0.35	0.23	0.4
NLS-72	1.49	1.17	0.94	0.38	0.52	0.19	0.47

Each number in the table represents the dropout rate of low-income students as a share of the total dropout rate of low-income students divided by the yearly dropout rate of high-income students as a share of the total dropout rate of high-income students.

Although these summary statistics provide a useful overview they do not tell us if the difference between poor and rich students are statistically significant. Therefore, we also extend our analysis by controlling for available proxies of ability that do not seem to be strongly colinear with the household's socioeconomic status.[9] If students form their beliefs rationally, these proxies are also positively correlated with students' initial beliefs about their ability levels because the belief distribution of higher ability students should first-order stochastically dominate the belief distribution of lower ability students.

Table 3 presents the marginal effect and percentage effect of having high or low socio-economic status relative to medium socioeconomic status on the college dropout rates for the NLSY79, NLSY97, and NLS-72 datasets. This table provides the results of logit regressions, where the dependent variable is equal to one if the student drops out and zero otherwise. For the three datasets rich students are less likely to drop out than are middle-income students, with the marginal effect ranging from -5.88 percent to -13.8 percent depending on the dataset. Low income students are more likely to drop out than middle income students, with marginal effect ranging from 2.56 percent to 15.76 percent depending on the dataset. Given the fraction of college students that drop out, the probability of dropping out increases by at least 27 percent for students from poor households relative to those from rich households. These results confirm the results shown in previous tables.

Our evidence regarding the negative relationship between wealth and dropout rates is also supported by Lovenheim and Reynolds (2010) who find that house price appreciation leads to higher graduation rates, especially among low-income families.

We also look at the distribution of dropout times across students with different family income. To explore this connection we run an ordered logit regression of the dropout time for those students who dropped out. We choose ordered logit because it is a simple way to solve the interval censoring problem caused by the discrete measurements of dropout times in our data, and the right-censoring problem inherent in duration data.[10] The results of the regressions can be found in Table 8 for the

[9] The educational attainment of the student's father and mother are strongly co-linear with the household's wealth and therefore these measures are not used.

[10] See also Han and Hausman (1990) which shows that the proportional hazards specification for a duration model leads to a likelihood function of an ordered logit form in the absence of time varying covariates. We have also run standard OLS regressions of dropout times on socio-economic status

Table 3 Marginal and percentage effect of socioeconomic status on dropout probability

		dF/dx	std. error	% effect	p	N
NLSY97	Low SES	0.0265	0.0266	12.78	0.02	1948
	High SES	-0.0901	0.0187	-43.48		
NLSY79	Low SES	0.1576	0.063	41.38	0.03	635
	High SES	-0.0588	0.0468	-15.44		
NLS-72	Low SES	0.0256	0.0298	4.27	0.00	2705
	High SES	-0.138	0.0211	-23.03		

The results are obtained from running logit regressions on the dropout probability based on the socioeconomic status of the student's family and a set of controls. The complete regression results and description of control variables can be found in Table 5, Table 6, and Table 7 in Appendix E. N is the number of observations and p is the p-value of the χ^2 test that compares the effect of high and low socio-economic status on dropout probability.

NLSY97, Table 9 for the NLSY79, and Table 10 for NLS-72 in Appendix E. These table show that the dropout times of poor and rich students are different from each other at 1% significance level.

Using the estimation results in these tables, we produce the predicted probability of dropping out in a given year as a function of the socioeconomic status of the student's family. Figure 1 to Figure 3 show the estimated dropout probabilities of students with different socio-economic status in different years of college evaluated at the mean values of other control variables. The common pattern in all these figures is that poor students tend to drop out in earlier years of college whereas richer students tend to drop out in later years even after controlling for other characteristics.

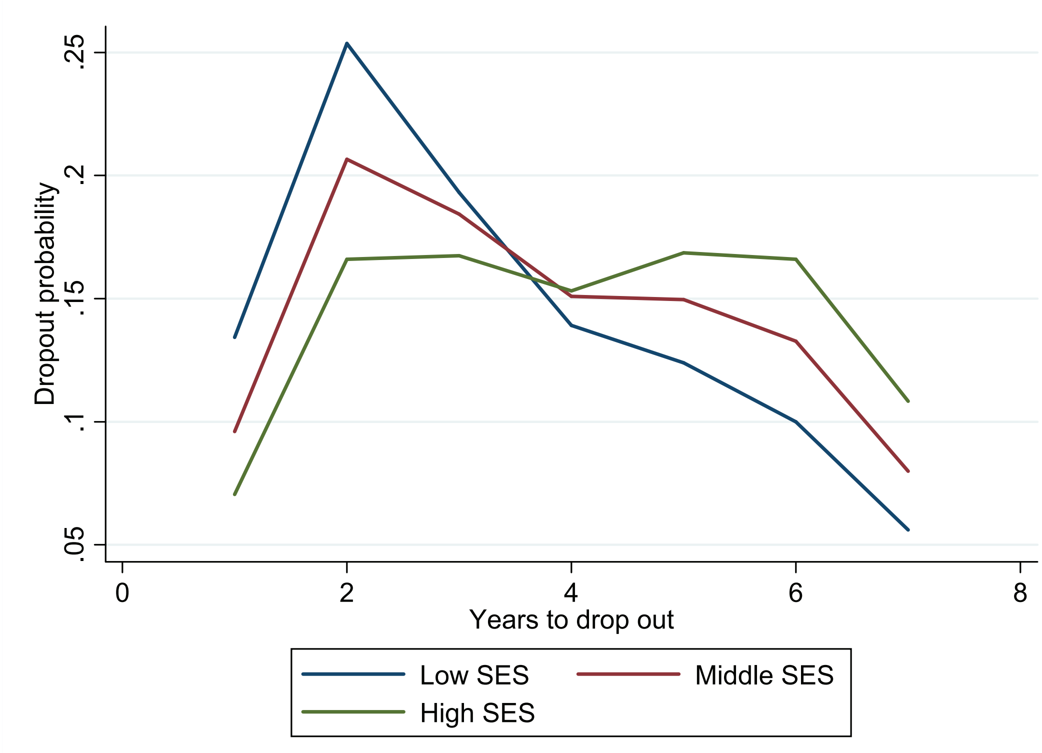
To summarize, a college student's family income not only affects the probability of dropping out but also the timing of attrition.

3 Model

At $t = 0$ students are enrolled in college, endowed with wealth level x_0. Students differ in their ability to acquire human capital at college; this ability can be low or

and other control variables and have found qualitatively similar results.

Figure 1 NLSY97: Predicted time to drop out



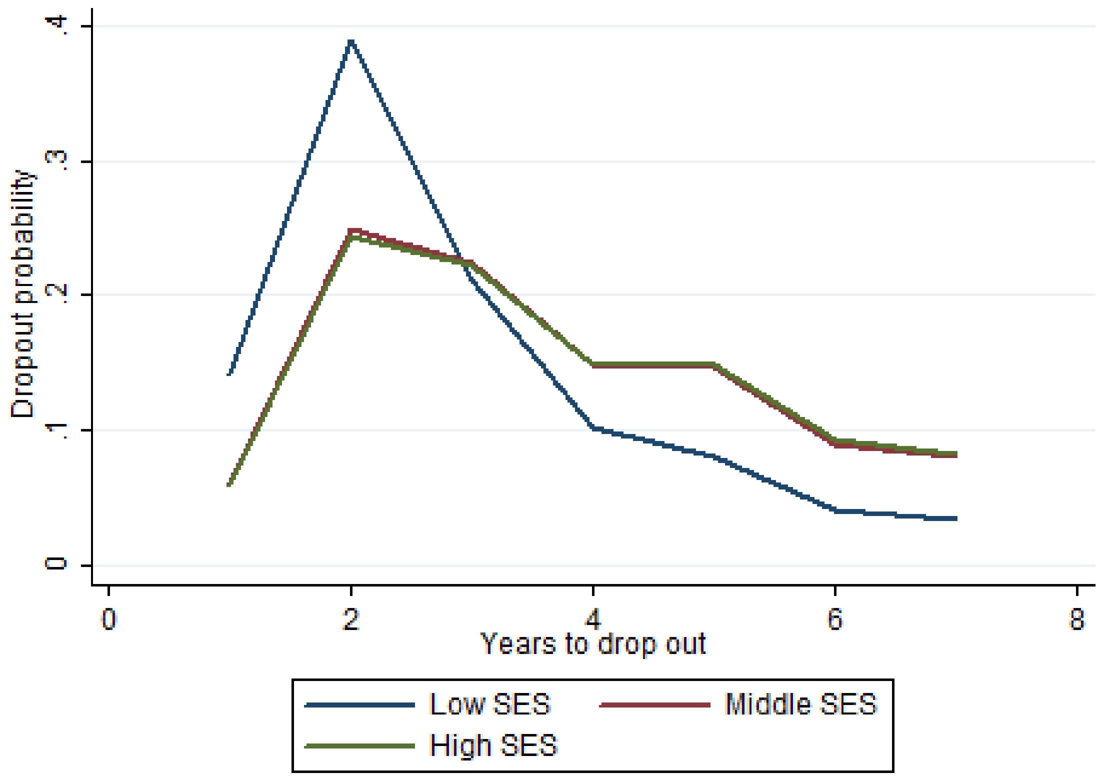
The figure plots the predicted probabilities of dropping out in a given time interval for students with different socioeconomic statuses, conditioning on the average characteristics of students who dropped out. To compute the probabilities we run an ordered logit regression that involves the probability of dropping out in a given year as a function of the socioeconomic status variable plus a set of controls. The complete regression results can be found in Table 8 in Appendix E.

high. Let $\mu \in \{0, 1\}$ denote the ability level, where $\mu = 0$ denotes low ability. The ability level is not observable at $t = 0$. Instead, individuals inherit a signal about their true type $p(0) = \Pr(\mu = 1)$.

At any point in time an agent can either be enrolled as a full-time student or working in low- or high-skilled sectors.[11] The high-skilled sector only hires high

[11]Post-secondary education is a combination of both two-year and four-year colleges with dy-

Figure 2 NLSY79: Predicted time to drop out



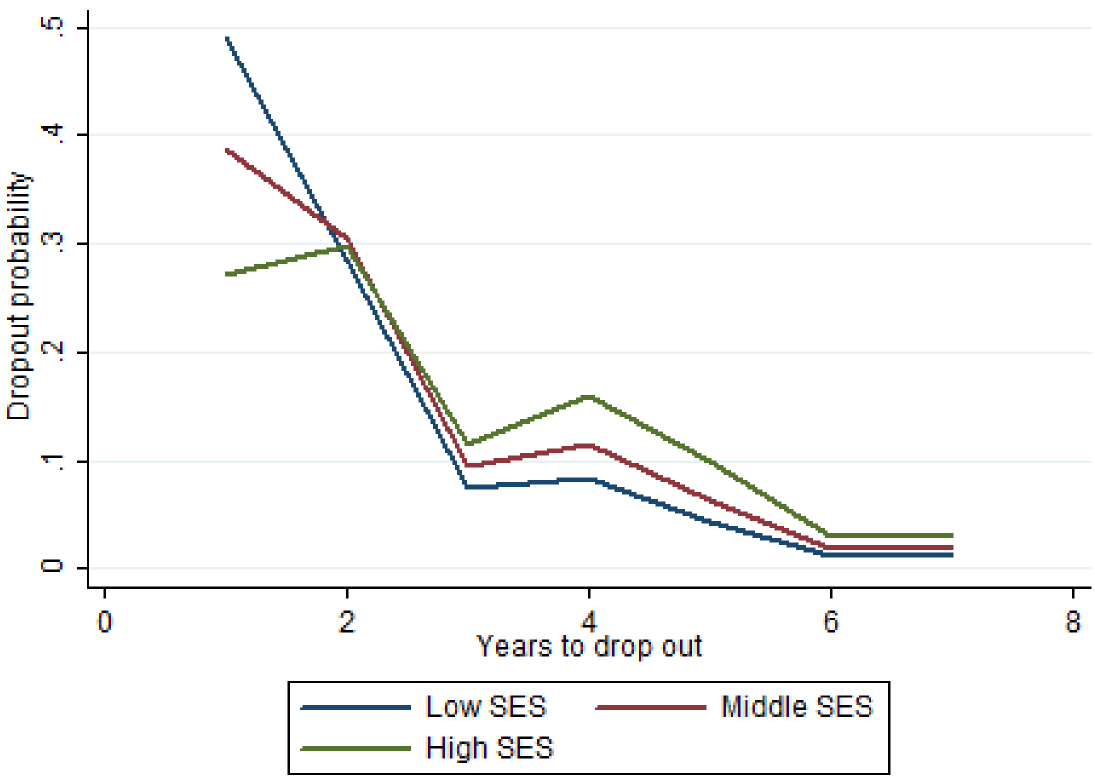
The figure plots the predicted probabilities of dropping out in a given time interval for students with different socioeconomic statuses, conditioning on the average characteristics of students who dropped out. To compute the probabilities we run an ordered logit regression that involves the probability of dropping out in a given year as a function of the socioeconomic status variable plus a set of controls. The complete regression results can be found in Table 9 in Appendix E.

ability workers with college degrees.[12] Work is assumed to be an absorbing state with a constant wage function $\tilde{w}(\mu, \tau)$ where $\tau \equiv T - t$ accounts for the amount

namic patterns that involve dropouts and transfers across types of schools. Using data from the National Longitudinal Survey of 1972, Trachter (2010) shows that students enrolled in four-year colleges either drop out or remain at the current type of institution until graduation.

[12]The interest of this paper is to understand dropout behavior. Dropouts usually leave school and join low-skilled sectors.

Figure 3 NLS-72: Predicted time to drop out



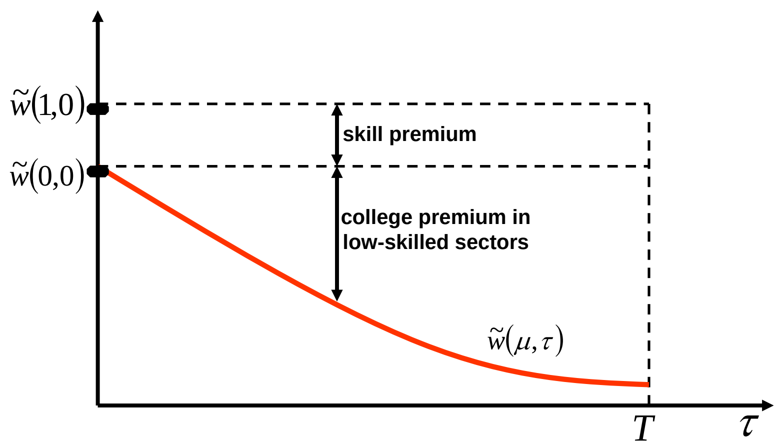
The figure plots the predicted probabilities of dropping out in a given time interval for students with different socioeconomic statuses, conditioning on the average characteristics of students who dropped out. To compute the probabilities we run an ordered logit regression that involves the probability of dropping out in a given year as a function of the socioeconomic status variable plus a set of controls. The complete regression results can be found in Table 9 in Appendix E.

of time left prior to graduation and T is the duration of college. We specify the function $\tilde{w}(\mu, \tau)$ as follows:

$$\tilde{w}(\mu, \tau) = \begin{cases} w(\tau) & \text{if } \tau > 0, \\ w(0) & \text{if } \tau = 0 \text{ and } \mu = 0, \\ w_1 & \text{if } \tau = 0 \text{ and } \mu = 1, \end{cases}$$

with $w(\tau_0) > w(\tau_1)$ if $\tau_0 < \tau_1$ and $w_1 > w(0)$. Therefore, the wage is increasing with the time spent in college and high-ability graduates enjoy higher wages. The function $\frac{w(\tau)}{w(T)}$ reflects the college premium in low-skilled sectors. A graphical representation of the wage function is depicted in Figure 4.

Figure 4 Skill and College Premium



The evolution of the wealth level x is given by

$$\frac{dx}{dt} = \begin{cases} rx + \tilde{w}(\mu,\tau) - c & \text{if working,} \\ rx - a - c & \text{if enrolled in college,} \end{cases}$$

where a denotes the per period cost of attending college.

At every period of length dt in college, students take an exam and get three possible grades: excellent, pass, or fail. High-ability students (i.e. $\mu = 1$) can get either a grade of pass or excellent for an exam. We let λ_1 denote the probability per unit of time that a high-ability student gets an excellent grade. Low-ability students (i.e. $\mu = 0$) can get either a passing or failing grade for an exam. We let λ_0 denote the probability per unit of time that a low-ability student gets a failing grade. Receiving a failing grade reveals that the student has low ability whereas receiving an excellent grade reveals that she has high ability. When the student receives a pass she updates her belief according to Bayes' rule. Table 4 presents the probability of the outcome of a given exam conditional on the student's true type, which is unobservable by the student.

We choose to model the student's learning process about her true type using discrete signals rather than continuous signals primarily because in our model the speed of learning, λ_0 and λ_1, depends on the type of the student. In a continuous Brownian Motion signal setting, this would be equivalent to having different volatilities for the signal process. However, Merton (1980) and Nelson and Foster (1994) point out that an observer of a continuous path generated by a diffusion process can estimate a constant or a smoothly time-varying volatility term with arbitrary precision over an arbitrarily short period of calendar time provided she has access to arbitrarily high frequency data. Of course, introducing a discrete signal is not the only way to get around this problem but it is analytically more efficient than other possibilities, such as introducing stochastic volatility or writing a discrete time model. These alternatives require belief updating about the mean and variance of the corresponding stochastic process which complicates the analysis without adding anything to the intuition.

Table 4 Probability of receiving different grades based on student's type

	Fail	Pass	Excellent
$\mu = 0$	$\lambda_0 dt$	$1 - \lambda_0 dt$	0
$\mu = 1$	0	$1 - \lambda_1 dt$	$\lambda_1 dt$

Each value in the table is the probability of receiving a given grade on the exam per unit of time dt, conditional of the student's true ability level.

A student initially enrolled in college chooses her consumption stream $\{c(t)\}_{t\geq 0}$ and whether to remain a student or to drop out and join the workforce, in order to maximize her time-separable expected discounted lifetime utility derived from consumption,

$$\mathbb{E}\left\{ \int_0^\infty e^{-\rho t} \frac{c(t)^{1-\gamma}}{1-\gamma} \,\middle|\, p(0), x(0) \right\},$$

where γ is the coefficient of Constant Relative Risk Aversion (CRRA).

We let $J(x, p, \tau)$ denote the value for a student with current wealth level x, prior p, and $T - \tau$ time already spent in school. Also, $V(x, \mu, \tau)$ denotes the value for a worker of true type μ with current wealth level x who spent $T-\tau$ time in school.

3.1 The Problem of a Worker

An individual with $T-\tau$ years of schooling who joins the workforce maximizes her lifetime discounted utility $\int_0^\infty e^{-\rho t} c^{1-\gamma} / (1-\gamma)\, dt$, subject to the law of motion for wealth, $dx/dt = rx + \tilde{w}(\mu,\tau) - c$. It will prove useful to characterize $W(\mu,\tau)$, the present discounted value of lifetime earnings. This object is simply

$$W(\mu,\tau) = \int_0^\infty e^{-rt} \tilde{w}(\mu,\tau) dt.$$

The Hamilton-Jacobi-Bellman representation of the worker's problem is

$$\rho V(x,\mu,\tau) = \max_c \frac{c^{1-\gamma}}{1-\gamma} + V_x(x,\mu,\tau)(rx + \tilde{w}(\mu,\tau) - c),$$

which states that the flow value of being a worker must be equal to the instantaneous utility derived from consumption in addition to the change in value that happens through the change in wealth.

The first-order condition of the worker's problem reads $c^{-\gamma} = V_x(x,\mu,\tau)$. Plugging it back and operating provides an equation that a worker's maximized value function needs to satisfy,

$$\rho V_x(x,\mu,\tau) = \frac{\gamma}{1-\gamma} \left[V_x(x,\mu,\tau)\right]^{1-\frac{1}{\gamma}} + r(x + W(\mu,\tau)) V_x(x,\mu,\tau), \tag{1}$$

with the solution given by

$$V(x,\mu,\tau) = A \left(r \left[x + W(\mu,\tau)\right]\right)^{1-\gamma}, \tag{2}$$

where

$$A \equiv \left[(1-\gamma) r\right]^{\gamma-1} \left[(\rho - (1-\gamma) r) \frac{1-\gamma}{\gamma}\right]^{-\gamma}. \tag{3}$$

The term $x + W(\mu,\tau)$ in equation (2) represents the worker's wealth also accounting for the discounted lifetime labor earnings.

Although we have assumed that the wages earned after college graduation are constant, this solution also holds if the wages depend upon on-the-job experience. Once we redefine $\tilde{w}(\mu, \tau, s)$ as the worker's wage depending on how long she has been working, s, we can replace $W(\mu, \tau)$ with $W(\mu, \tau) = \int_0^\infty e^{-rs}\tilde{w}(\mu, \tau, s)ds$ in the solution. Therefore, our results are robust to different specifications of the lifetime wage profile for students with different durations of college education.

3.2 The problem of a student of type μ and the natural borrowing limit

If the student knows that she has high ability, $\mu = 1$, her value function satisfies

$$\rho J(x, 1, \tau) = \max_c \frac{c^{1-\gamma}}{1-\gamma} + (rx - a - c)J_x(x, 1, \tau) + J_\tau(x, 1, \tau)\frac{d\tau}{dt}, \qquad (4)$$

subject to the terminal condition $J(x, 1, 0) = V(x, 1, 0)$. This terminal condition states that, upon graduation, the student's value function is equivalent to the value of being a worker in the high-skill sector. We are also using an implicit condition guaranteeing that high-skilled students would never find it profitable to drop out. This condition is later derived in Lemma 1.

This Hamilton-Jacobi-Bellman equation states that the desired return on being a student with high ability, $\mu = 1$, with wealth level x, and with τ periods left to graduation equals the instant utility derived from consumption plus the change in value due to the change in wealth and time to graduation, both due to the change accrued in time. Note that $d\tau/dt = -1$.

The first-order condition of this problem states that $c^{-\gamma} = J_x(x, 1, \tau)$. Solving for c and plugging the result back into equation (4) provides the student's maximized value function. We guess and verify that the solution to this object is $J(x, 1, \tau) = A\left[rx + B(\tau)\right]^{1-\gamma}$, where A is defined as in equation (3) and $B(\tau)$ needs to be solved for. Intuitively, $B(\tau)$ accounts for the change in the value function due to the time to graduation and the terminal payoff. Plugging the guess into the maximized value function provides

$$rB(\tau) + B'(\tau) + ra = 0$$

with boundary condition, $B(0) = rW(1,0)$, that follows from the terminal condition of the problem presented in equation (4). This is an Ordinary Differential Equation in one variable with a terminal condition. The solution is $B(\tau) = (rW(1,0) + a)\, e^{-r\tau} - a$, and therefore the value function of a student of type $\mu = 1$ is

$$J(x,1,\tau) = A \left(rx - a + e^{-r\tau} \left(rW(1,0) + a\right)\right)^{1-\gamma} \tag{5}$$

or,

$$J(x,1,\tau) = A \left[r \left(x + e^{-r\tau} W(1,0) - \frac{1 - e^{-r\tau}}{r} a\right)\right]^{1-\gamma},$$

where the term in parentheses gives the student's net lifetime wealth after accounting for the discounted value of future wages and the remaining college costs.

The next lemma characterizes the condition guaranteeing that high-ability students who know their type will not find it profitable to drop out of college.

Lemma 1 *A student of type μ with current wealth level x and $T - \tau$ time spent in college will choose to remain as a student until $\tau = 0$ if $re^{-r\tau} W(1,0) - a\left(1 - e^{-r\tau}\right) \geq rW(1,\tau)$.*

Lemma 1 follows from noting that for a student to remain in college it has to be the case that, for every value of τ, $J(x,1,\tau) \geq V(x,1,\tau)$. This condition simply indicates that the graduation premium for a high-ability student is high enough so that a student who knows she has high ability will remain in college until graduation. We assume throughout the paper that this condition holds.[13]

Another important assumption of the model is that low-ability students who know their type will always find it profitable to drop out of college. If this were not the case, for some values of time to gradution, τ, there would be no dropouts by

[13]This condition at $\tau = T$ also guarantees that there are at least some students with optimistic prior beliefs willing to enroll in college.

construction. The next lemma characterizes the condition guaranteeing that low-ability students who know their type will decide to drop out and join the workforce immediately, i.e. $J(x, 0, \tau) = V(x, 0, \tau)$.

Lemma 2 *A student who knows that she is of type $\mu = 0$ will drop out immediately if $a + rW(0, \tau) + W_\tau(0, \tau) > 0$.*

Proof. See Appendix A. ■

Intuitively, the marginal cost of attending college for another period of time is adt. Moreover, the marginal increase in the present value of earnings after an additional period of college education is

$$e^{-rdt}W(0, \tau - dt) - W(0, \tau) = -[rW(0, \tau) + W_\tau(0, \tau)]\, dt + O(dt)^2,$$

where we used a Taylor series expansion. Subtracting the increase in marginal earnings from marginal cost, dividing by dt and taking the limit as $dt \to 0$ gives the condition stated above.

Although our model does not entail any explicit borrowing constraint this last assumption leads to an implicit borrowing constraint for a student who does not know her true ability level. Every student with $p < 1$ faces ex-ante a positive probability of receiving a shock revealing that she has low ability. Such a shock would force her to drop out of college and join the low-skilled workforce. Since the marginal value of wealth for a college dropout goes to infinity as x goes to $-W(0, \tau)$ the student will never borrow more than her discounted value of lifetime earnings $W(0, \tau)$. As long as the wage profile is common knowledge, as it is the case in our model, this ”natural” borrowing constraint should also be the actual one because the student can always repay the borrowed money using his earnings when $x > -W(0, \tau)$.

3.3 The Problem of a Student of Unknown Type

The problem of a student who does not know her ability level is more difficult to solve because of three reasons. First, the wage upon graduation depends on the agent's true ability. Second, the arrival of new information via exams results in updating of the student's belief. Third, some students drop out.

Before constructing the Hamilton-Jacobi-Bellman equation for this case first it is useful to consider how the information obtained through exams can be used to update beliefs. Consider a student with the belief $p(t)$, where $p(t)$ is the probability of being type 1 conditional on the information available at time t. Table 4 can be used to construct the posterior conditional on the grade received during $(t, t+dt)$. If the student receives a failing grade, it is clearly revealed that she has low ability and thus $p(t+dt)=0$. If the student receives a grade of excellent, $p(t+dt)=1$. Conditional on not receiving a failing or excellent grade through period $(t, t+dt)$, receving a passing grade in the current exam implies that Bayes' rule can be used to update beliefs,

$$p(t+dt) = \frac{p(t)\left[1-\lambda_1 dt\right]}{p(t)\left[1-\lambda_1 dt\right] + \left[1-p(t)\right]\left[1-\lambda_0 dt\right]}.$$

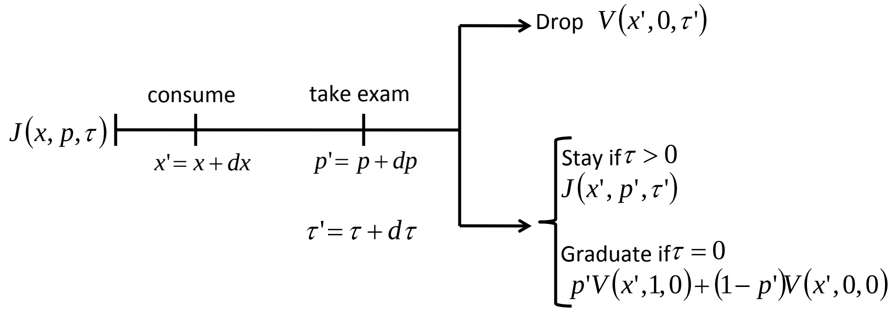
Substracting $p(t)$, dividing by dt, and taking the limit as $dt \to 0$ provides the Bayes' rule in its continuous time formulation,

$$\frac{dp}{dt} = -\left(\lambda_1 - \lambda_0\right) p(1-p). \tag{6}$$

Figure 5 describes the timeline of the student's problem in a given period. The student enters the period with wealth level x, prior p, and remaining time until graduation τ. Her value function is therefore $J(x,p,\tau)$. At the beginning of the period, the student chooses her consumption level and thus produces the wealth level x' for next period. Before the end of the period she takes an exam and with the grade at hand updates her beliefs to p'. By the end of the period she accumulates more time in school and therefore the distance to graduation is reduced to τ'. At the beginning of next period the student compares the value of remaining in school, that is either $J\left(x',p',\tau'\right)$ if $\tau'>0$ or $p'V\left(x',1,0\right)+\left(1-p'\right)V\left(x',0,0\right)$ if $\tau'=0$ (i.e. graduation), with the value of joining the workforce $V\left(x',0,\tau\right)$ to decide between staying in college or becoming a dropout.

The Hamilton-Jacobi-Bellman equation for a student with current wealth level

Figure 5 Timeline



A student starts the current period with wealth level x, prior p, and remaining time in college τ. At the beginning of the period a student chooses her consumption level and thus provides the new value for wealth x'. Before the end of the period the student takes an exam used to produce the posterior p' and reduces the time left to graduation to τ'. At the beginning of next period the student chooses between dropping out or remaining as a student (or graduation is $\tau' = 0$).

x, prior p, and τ periods left in school is

$$\begin{aligned}\rho J(x,p,\tau) \;&=\; \max_c \tfrac{c^{1-\gamma}}{1-\gamma} + (rx - a - c) J_x(x,p,\tau) - J_\tau(x,p,\tau) \\ &\quad - (\lambda_1 - \lambda_0)\, p(1-p) J_p(x,p,\tau) + \lambda_1 p \left[J(x,1,\tau) - J(x,p,\tau)\right] \\ &\quad + \lambda_0 (1-p) \left[V(x,0,\tau) - J(x,p,\tau)\right]. \end{aligned} \tag{7}$$

This equation states that the desired return on being a student with current wealth level x, prior p, and time to graduation τ equals the instant utility derived from consumption plus the change in value of being a student through (i) the change in wealth, (ii) the change in τ, and (iii) belief updating. When a passing grade arrives the belief adjustment is continuous through the Bayesian updating of p; when an excellent grade arrives, expected with unconditional probability $\lambda_1 p$, the change in value occurs through switching from having $p(t) = p$ to $p(t + dt) = 1$; and when a failing grade arrives, expected with unconditional probability $\lambda_0(1-p)$, the change in value is through switching from having $p(t) = p$ to $p(t + dt) = 0$. Also note that the problem faced by a high-ability student who knows her type presented in equation (4) is a particular case of the problem presented here, that follows by setting $p = 1$ in equation (7) and noting that $d\tau/dt = -1$.

A student faces the problem presented in equation (7) subject to a set of boundary conditions,

$$\begin{aligned} J(x,p,0) &= pV(x,1,0) + (1-p)V(x,0,0) \\ J(x,p^*(x,\tau),\tau) &= V(x,0,\tau) \\ J_p(x,p^*(x,\tau),\tau) &= 0 \\ J_x(x,p^*(x,\tau),\tau) &= V_x(x,0,\tau) \\ J_\tau(x,p^*(x,\tau),\tau) &= V_\tau(x,0,\tau). \end{aligned} \tag{8}$$

The first equation gives the Terminal Condition (TC) and states that the value of being a student with no time remaining until graduation has to equal the expected value of being a worker. Note that with probability p a student expects to be of type $\mu = 1$ and therefore would earn lifetime discounted labor income $W(1,0)$, while with probability $1-p$ she expects to be of type $\mu = 0$ and therefore earn lifetime discounted labor income $W(0,0)$. To understand the second to fifth equations $p^*(x,\tau)$ needs to be defined. Let $p^*(x,\tau)$ be the belief threshold such that students with $p \leq p^*(x,\tau)$ drop out and join the workforce. The second equation states that a student with $p = p^*(x,\tau)$, wealth level x, and τ periods away from graduation has to be indifferent between staying in school and dropping out and enjoying lifetime discounted labor income $W(0,\tau)$. This equation is also known as the Value Matching Condition (VMC). The third, fourth, and fifth equations are known as Smooth Pasting Conditions (SPC) required for the optimality of $p^*(x,\tau)$.[14]

The first-order condition of the problem presented in equation (7) is $c^{-\gamma} = J_x(x,p,\tau)$. Plugging it back into equation (7) together with the terminal, value matching, and smooth pasting conditions provides the equation that the threshold $p^*(x,\tau)$ needs to satisfy,

$$\begin{aligned} \rho V(x,0,\tau) &= \tfrac{\gamma}{1-\gamma} V_x(x,0,\tau)^{1-\frac{1}{\gamma}} + (rx-a)V_x(x,0,\tau) \\ &\quad + \lambda_1 p^*(x,\tau)\left[J(x,1,\tau) - V(x,0,\tau)\right] - V_\tau(x,0,\tau), \end{aligned} \tag{9}$$

[14]For a treatment of Value Matching and Smooth Pasting Conditions see Dumas (1991) and Dixit (1993).

which we can rewrite as

$$\lambda_1 p^* (x, \tau) [J (x, 1, \tau) - V (x, 0, \tau)] = [a + rW (0, \tau) + W_\tau (0, \tau)] V_x (x, 0, \tau) .$$

This equation provides provides intuition about the belief threshold $p^* (x, \tau)$. The left-side of this equation is the expected utility gain from delaying the dropout decision by dt, whereas the right side represents the marginal net utility loss due to delaying the dropout decision. The student chooses the optimal dropout time, by equalizing the marginal gain and loss from delaying the dropout decision.

Solving for $p^*(x, \tau)$ allows for a close-form representation of the student's dropout threshold,

$$p^*(x, \tau) = \frac{a + rW(0, \tau) + W_\tau(0, \tau)}{\lambda_1} \frac{V_x(x, 0, \tau)}{J(x, 1, \tau) - V(x, 0, \tau)} > 0, \qquad (10)$$

provided that Lemma 2 holds. This threshold is decreasing with the wealth level. That is,

$$\frac{\partial p^*(x, \tau)}{\partial x} < 0, \qquad (11)$$

where the details of the calculations can be found in Appendix B.

Our main result is that, conditional on their beliefs, college students from wealthier families drop out later and are less likely to drop out than are poor students. The result that the threshold p^* is decreasing with the wealth level is not enough to argue this result because the consumption profiles during college tenure can overcome the initial difference in wealth.[15] The next proposition and corollary deal with this.

Let τ^* denote the time to graduation at the moment the individual joins the workforce. For example, $\tau^* = T$ if the individual joins the workforce directly after high school graduation and $\tau^* = 0$ if the individual joins the workforce with a college degree. The following proposition states that, conditional on abilities and initial prior beliefs, the distribution of dropout times for richer students first-order stochastically dominate the distribution for poor students in the model, which

[15] Miao and Wang (2007) entrepreneurial survival model is a special case of our model where they also show that the boundary p^* is decreasing in the wealth level. However, they immediately conclude that richer entrepreneurs survive longer without providing an explicit proof.

explains the pattern in Figure 1 to Figure 3.

Proposition 1 *Let* $x^i(0)$ *and* $x^j(0)$ *denote the initial wealth levels at time 0 of students* i *and* j. *If* $x^i(0) > x^j(0)$ *then, for any* $\bar{\tau}$, $Pr\left\{\tau^* \leq \bar{\tau} \middle| x^i(0), p(0), \mu\right\} \geq Pr\left\{\tau^* \leq \bar{\tau} \middle| x^j(0), p(0), \mu\right\}$. *In other words, given a skill level* μ *and initial belief* $p(0)$, *richer students tend to drop out later and have longer expected tenures in college.*

Proof. See Appendix C. ■

The next corollary extends the result to show that, conditioning in the initial prior $p(0)$ and ability level μ, students from richer families are less likely to drop out and, therefore more likely to graduate.

Corollary 1 *Let* $x^i(0)$ *and* $x^j(0)$ *denote the wealth levels at time 0 of students* i *and* j. *If* $x^i(0) > x^j(0)$ *then* $\Pr\left[\widetilde{\tau} = 0 \middle| x^i(0), p(0), \mu\right] \geq \Pr\left[\widetilde{\tau} = 0 \middle| x^j(0), p(0), \mu\right]$. *That is, once conditioned on the initial prior* $p(0)$ *and skill level* μ, *richer students are more likely to graduate from college.*

Proof. Set $\bar{\tau} = 0$ in Proposition 1. ■

Both Proposition 1 and Corollary 1 are driven by the fact that ex-ante uncertainty regarding the outcome of college education cannot be diversified away, which makes the investment in college education more risky than financial investments in the model.[16] Given students' abilities and beliefs, a student with higher wealth chooses to increase the number of dollars invested in the risky asset, i.e. stay longer in college, if absolute risk aversion is decreasing in wealth as it is the case with the CRRA preferences.[17]

[16] Of course, financial assets in real life are risky due to macroeconomic fluctuations not modeled in this paper. However, these macroeconomic fluctuations also affect the payoff of college education through their effect on wages, and investing in college education is still riskier than investing in financial assets due to undiversifiable idiosyncratic risk.

[17] In Appendix D we extend the model to allow for hyperbolic risk aversion (HARA) preferences, which include CRRA as a special case. We show that the belief threshold is decreasing in wealth if and only if the absolute risk aversion is decreasing. Furthermore, Proposition 1 and Corollary 1 also apply here.

4 Conclusion

In this paper, we provide evidence regarding the skewed distribution of college dropouts with respect to the student's family wealth. Poor students are more likely to drop out and they tend to do so earlier than rich students. We explore whether a model that treats college education as a risky investment and incorporates Bayesian learning about own's unobserved ability can explain the skewness in dropout behavior.

Our main results rely on the fact that the outcome of obtaining a college education is subject to uncertainty against which students cannot insure themselves. When we combine this fact with CRRA preferences so that the absolute risk aversion decreases with wealth we arrive at the conclusion that poor students are less willing to accept the risk associated with pursuing a college education. This mechanism generates the skewness observed in dropout behavior.

We provide a closed-form characterization of a student's optimal choice as a function of (i) the expected future income due to graduation (through the prior, p), and (ii) the direct and indirect costs of remaining in college (through the time remaining to graduation, τ). We exploit the model's simplicity to show that it is able to fit the data qualitatively: (i) poorer students are more likely to drop out than are rich students, and (ii) if the poor students drop out, they do so earlier than students from wealthier families.

To motivate the theory we run a series of reduced-form regressions, conditioning by measures of the student's unobserved ability and prior beliefs. The regressions' results are in line with the model's predictions. We estimate that poor college students are at least 27 percent more likely to drop out than rich students and if they around a year earlier.

Our goal is not to claim that borrowing constraints are not part of the story behind the high and skewed college dropout rates. Instead, we provide a complementary story that is able to explain the skewed distribution of the time to drop out. Furthermore, our story is consistent with Stinebrickner and Stinebrickner (2008) which finds that borrowing constraints are not the main determinant of dropout decision, and Stinebrickner and Stinebrickner (2009) which shows that bad grades are a good predictor of dropout behavior.

Since our model generates these results without including explicit borrowing constraints, it also suggests that policies that are geared towards reducing borrowing constraints, such as student loan programs, are not likely to eliminate the differences in dropout rates between rich and poor students. Moreover, a direct subsidy to poor students for their college education would not only increase their graduation rates but also their tenure in college, by both reducing the cost of spending additional time in college and increasing the expected gain from delaying the dropout decision. The optimal subsidy would depend on the wealth level of the student's family and the distribution of ability for a given wealth level, which is the topic of future research.

Finally, it is plausible that poor students are more likely to participate in the labor force during college and have less time to devote to study, making them take longer to finish college. Although we have not included the time allocation decision in our model, an extension of the model with the decision to work versus study during college can potentially generate this result, following the basic intuition in this paper: Labor force participation while attending college provides a safe income today, and poor students are more likely to work during college because they are more risk averse, and hence poor students are more willing to invest in a safe asset. The empirical and theoretical analysis of the student's work-study decision during college is an interesting question that is left for future research.

References

Aiyagari, S Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *The Quarterly Journal of Economics* 109(3): 659–84.

Chen, Stacey H. 2008. "Estimating the Variance of Wages in the Presence of Selection and Unobserved Heterogeneity." *The Review of Economics and Statistics* 90(2): 275–289.

Dixit, Avinash K. 1993. *The Art of Smooth Pasting*. Routledge.

Dumas, Bernard. 1991. "Super contact and related optimality conditions." *Journal of Economic Dynamics and Control* 15(4): 675–685.

Han, Aaron, and Jerry A. Hausman. 1990. "Flexible Parametric Estimation of Duration and Competing Risk Models." *Journal of Applied Econometrics* 5(1): 1–28.

Heckman, James J., and Sergio Urzua. 2008. "The Option Value of Educational Choices and the Rate of Return to Educational Choices." Working paper. Cowles Foundation Structural Conference, Yale University.

Lovenheim, Michael F., and C. Lockwood Reynolds. 2010. "The Effect of Housing Wealth on College Choice: Evidence from the Housing Boom."

Manski, Charles, and David Wise. 1983. *College Choice in America*. Harvard University Press.

Merton, Robert C. 1980. "On estimating the expected return on the market : An exploratory investigation." *Journal of Financial Economics* 8(4): 323–361.

Miao, Jianjun, and Neng Wang. 2007. "Experimentation under Uninsurable Idiosyncratic Risk: An Application to Entrepreneurial Survival." Working paper.

Mincer, Jacob A. 1974. *Schooling, Experience, and Earnings*. No. minc74-1 in NBER Books. National Bureau of Economic Research, Inc.

Nelson, Daniel B., and Dean P. Foster. 1994. "Asymptotic Filtering Theory for Univariate ARCH Models." *Econometrica* 62(1): 1–41.

Stange, Kevin M. 2011. "An Empirical Examination of the Option Value of College Enrollment." *American Economic Journal: Applied Economics* forthcoming.

Stinebrickner, Ralph, and Todd Stinebrickner. 2008. "The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study." *American Economic Review* 98(5): 2163–84.

Stinebrickner, Todd R., and Ralph Stinebrickner. 2009. "Learning about Academic Ability and the College Drop-out Decision." NBER Working Papers 14810. National Bureau of Economic Research, Inc.

Trachter, Nicholas. 2010. "Option Value and Transitions in a Model of Postsecondary Education." Working paper.

Appendix

A Proof of Lemma 2

Let τ^* denote the threshold for τ so that students drop out for $\tau < \tau^*$. A student of type $\mu = 0$ faces the problem given by

$$\rho J(x, 0, \tau) = \max_c \frac{c^{1-\gamma}}{1-\gamma} + (rx - c - a)J_x(x, 0, \tau) - J_\tau(x, 0, \tau),$$

subject to the terminal condition $J(x, 0, 0) = V(x, 0, 0)$, the boundary condition $J(x, 0, \tau^*) = V(x, 0, \tau^*)$, and smooth pasting conditions $J_x(x, 0, \tau^*) = V_x(x, 0, \tau^*)$ and $J_\tau(x, 0, \tau^*) = V_\tau(x, 0, \tau^*)$.

Plugging in the first-order condition provides that

$$\rho V_x(x, 0, \tau^*) = \frac{\gamma}{1-\gamma} \left[V_x(x, 0, \tau^*)\right]^{1-\frac{1}{\gamma}} + (rx - a)V_x(x, 0, \tau) - V_\tau(x, 0, \tau^*).$$

Using equation (1) this equation can be reduced to $a + rW(0, \tau^*) + W_\tau(0, \tau^*) = 0$. Hence, if $a + rW(0, \tau^*) + W_\tau(0, \tau^*) > 0$ for all $\tau \leq T$ the boundary condition for an interior dropout boundary is not satisfied. Moreover, the desired return from continuing education in terms of utility, the left side of Bellman equation, is greater than the continuation value, the right side of Bellman equation, at the default boundary and hence it is optimal to drop out immediately.

B Proof of $\frac{\partial p^*}{\partial x} < 0$

Differentiating the threshold p^* (see equation (10)) with respect to x provides

$$\begin{aligned}
\frac{\partial p^*}{\partial x} &= \frac{a+rW(0,\tau)+W_\tau(0,\tau)}{\lambda_1}\left[\begin{array}{c} \frac{V_x(x,0,\tau)}{J(x,1,\tau)-V(x,0,\tau)}\frac{V_{xx}(x,0,\tau)}{V_x(x,0,\tau)} \\ -\frac{V_x(x,0,\tau)}{J(x,1,\tau)-V(x,0,\tau)}\frac{J_x(x,1,\tau)-V_x(x,0,\tau)}{J(x,1,\tau)-V(x,0,\tau)} \end{array}\right] \\
&= \frac{a+rW(0,\tau)+W_\tau(0,\tau)}{\lambda_1}\frac{V_x(x,0,\tau)}{J(x,1,\tau)-V(x,0,\tau)}\left[\frac{V_{xx}(x,0,\tau)}{V_x(x,0,\tau)} - \frac{J_x(x,1,\tau)-V_x(x,0,\tau)}{J(x,1,\tau)-V(x,0,\tau)}\right] \\
&= p^*(x,\tau)\left[\frac{V_{xx}(x,0,\tau)}{V_x(x,0,\tau)} - \frac{J_x(x,1,\tau)-V_x(x,0,\tau)}{J(x,1,\tau)-V(x,0,\tau)}\right] \\
&= -p^*(x,\tau)\left[\begin{array}{c} \frac{\gamma r}{r[x+W(0,\tau)]} \\ +(1-\gamma)r\frac{\left(r\left[x+e^{-r\tau}W(1,0)-\frac{a}{r}\left(1-e^{-r\tau}\right)\right]\right)^{-\gamma}-(r[x+W(0,\tau)])^{-\gamma}}{\left(r\left[x+e^{-r\tau}W(1,0)-\frac{a}{r}(1-e^{-r\tau})\right]\right)^{1-\gamma}-(r[x+W(0,\tau)])^{1-\gamma}} \end{array}\right] \\
&= -p^*(x,\tau)\left[\frac{\gamma r}{r[x+W(0,\tau)]} + \frac{1}{x+W(0,\tau)}\frac{y^{-\gamma}-1}{\frac{y^{1-\gamma}-1}{1-\gamma}}\right] \\
&= -\frac{p^*(x,\tau)}{x+W(0,\tau)}\left[\gamma + \frac{y^{-\gamma}-1}{\frac{y^{1-\gamma}-1}{1-\gamma}}\right],
\end{aligned}$$

where $y \equiv \frac{x+e^{-r\tau}W(1,0)-\frac{a}{r}\left(1-e^{-r\tau}\right)}{x+W(0,\tau)}$, with $y \geq 1$ provided the condition on Lemma 1 holds. Next, it will be proved by contradiction that $\gamma + (1-\gamma)\frac{y^{-\gamma}-1}{y^{1-\gamma}-1} > 0$ when $y \geq 1$.

Consider first the case where $\gamma < 1$. Suppose that $\gamma + (1-\gamma)\frac{y^{-\gamma}-1}{y^{1-\gamma}-1} < 0$. Because $\gamma \in (0,1)$ and hence $y^{1-\gamma}-1 > 0$ we can multiply both sides of this inequality by $y^{1-\gamma}-1$ to get $\gamma(y^{1-\gamma}-1)+(1-\gamma)(y^{-\gamma}-1) < 0$. The left-hand-side of this in equality is strictly increasing in y and therefore attains its minimum at $y=1$ with value equal to 0. Therefore, $\gamma(y^{1-\gamma}-1)+(1-\gamma)(y^{-\gamma}-1) < 0$ and hence $\gamma + (1-\gamma)\frac{y^{-\gamma}-1}{y^{1-\gamma}-1} < 0$ is not possible.

Now consider the case where $\gamma > 1$. Suppose that $\gamma + (1-\gamma)\frac{y^{-\gamma}-1}{y^{1-\gamma}-1} < 0$. Because $\gamma > 1$ and hence $y^{1-\gamma}-1 < 0$ we can multiply both sides of this inequality by $y^{1-\gamma}-1$ to get $\gamma(y^{1-\gamma}-1)+(1-\gamma)(y^{1-\gamma}-1)>0$. The left-hand-side of this equation is strictly decreasing in y and therefore attains its maximum at $y=1$ with value equal to 0. Therefore, $\gamma(y^{1-\gamma}-1)+(1-\gamma)(y^{1-\gamma}-1)>0$ and hence $\gamma+(1-\gamma)\frac{y^{-\gamma}-1}{y^{1-\gamma}-1}<0$ is not possible.

As $\gamma+(1-\gamma)\frac{y^{-\gamma}-1}{y^{1-\gamma}-1}>0$ for every $\gamma>0$, $\frac{\partial p^*}{\partial x}<0$.

C Proof of Proposition 1

Two students with the same skill level μ are equally likely to receive a failing or an excellent grade at any point in time. Therefore, although the grade earned affects the behavior of an individual student it does directly affect the distribution of dropout times for students with different wealth levels. Therefore, we look at the dropout behavior of students that do not receive any signals that reveal their true types.

Suppose we have two students i and j with the same initial belief, i.e. $p^i(0)=p^j(0)$, and with initial wealth levels $x^i(0)>x^j(0)$ so that the first student is initially richer. There are two possible outcomes conditional on not receiving a signal. First, if $p(0)$ is high enough both students wait until they graduate which does not violate our proposition as $\tau^{*i}=\tau^{*j}=0$. Second, if $p(0)$ is not high enough at least one of the students drops out. Let $t_0<T$ be the first point in time when one of the students drop out. Then we have $p^i(t)=p^j(t)$ for all $t\leq t_0$ because $p^i(0)=p^j(0)$ and the belief evolution is the same for both students conditional on not receiving a signal. Moreover, we know that the richer student does not drop out earlier than the poorer student, that is $\tau^{*i}\leq\tau^{*j}$, if $x^i(t)\geq x^j(t)$ for all $t\leq t_0$ because $p^*\left(x^i,\tau\right)\leq p^*\left(x^j,\tau\right)$ as long as $x^i\geq x^j$. Therefore, we can prove our proposition by showing that $x^i(t)\geq x^j(t)$ for all $t\leq t_0$. Suppose $x^i(t)<x^j(t)$ for some $t<t_0$. Then, since $x^i(t)$ and $x^j(t)$ have continuous paths there exists a $\bar{t}\leq t_0$ where $x^i\left(\bar{t}\right)=x^j\left(\bar{t}\right)$ by the intermediate value theorem. Moreover, since $p^i(t)=p^j(t)$ for all $t\leq t_0$ we have $p^i\left(\bar{t}\right)=p^j\left(\bar{t}\right)$. As a result, both students' consumption decisions are synchronized from time $\bar{t}$ on because they are forward looking. Hence, $x^i(t)=x^j(t)$ for $\bar{t}\leq t\leq t_0$ which is a contradiction.

D Hyperbolic Risk Aversion

In this appendix we extend the model to allow for hyperbolic risk aversion (HARA) preferences and show that the belief threshold is decreasing in wealth if and only if we have decreasing risk aversion. The absolute risk aversion for this class of preferences is given by

$$RA = -\frac{u''(c)}{u'(c)} = \frac{1}{ac+b}$$

where a and b are constants. For $a = 0$ we have exponential preferences, for $a > 0$ we have decreasing absolute risk aversion, and for $a < 0$ we have increasing absolute risk aversion. Moreover, we obtain the CRRA preferences in the model for $b = 0$. The solution of this differential equation is given by

$$u(c) = \kappa_1 \frac{(ac+b)^{1-1/a}}{a-1} + \kappa_2,$$

where κ_1 and κ_2 are constants of integration. Let $\gamma \equiv 1/a$ and $\bar{c} \equiv -b/a$. Then,

$$u(c) = \kappa_1 a^{-1/a} \frac{(c+b/a)^{1-1/a}}{(a-1)/a} + \kappa_2 = \bar{\kappa}_1 \frac{(c-\bar{c})^{1-\gamma}}{1-\gamma} + \kappa_2.$$

Since $a > 0$, hence $\gamma > 0$, determines if we have decreasing risk-aversion we stick to the linear transformation of this utility function, $u(c) = (c-\bar{c})^{1-\gamma} / (1-\gamma)$.

Let us define $\hat{c} \equiv c - \bar{c}$ as the surplus consumption and the $\hat{x} \equiv x - \bar{c}/r$ as the surplus wealth. Then, $d\hat{x}/dt = dx/dt$ and all $(rx - c)$ terms in the law of motion of x become $r\hat{x} - \hat{c}$. Moreover, the utility function becomes $u(\hat{c}) = (\hat{c})^{1-\gamma} / (1-\gamma)$. Therefore, the new belief boundary will be $\hat{p}^*(x,\tau) = p^*(\hat{x},\tau) = p^*(x - \bar{c}/r, \tau)$. Hence, in order to show that $\partial\hat{p}^*/\partial x < 0$ iff $\gamma > 0$, it is enough to show that $\partial p^*/\partial x < 0$ iff $\gamma > 0$. If $\gamma > 0$, $\partial p^*/\partial x < 0$ follows immediately from Appendix B. So, we only have to show that $\gamma > 0$ if $\partial p^*/\partial x < 0$.

From the derivation in Appendix B and Lemma 1, it follows that $\partial p^*/\partial x < 0$ implies that $\gamma + (1-\gamma)\frac{y^{-\gamma}-1}{y^{1-\gamma}-1} > 0$ for all $y \geq 1$. Suppose $\partial p^*/\partial x < 0$ holds but

$\gamma \leq 0$. Then, $\gamma + (1-\gamma)\frac{y^{-\gamma}-1}{y^{1-\gamma}-1} \leq 0$ because

$$signum\left[\gamma + (1-\gamma)\frac{y^{-\gamma}-1}{y^{1-\gamma}-1}\right] = signum\left[\gamma\left(y^{1-\gamma}-1\right) + (1-\gamma)\left(y^{-\gamma}-1\right)\right]$$

and

$$\max_{y \geq 1} \gamma\left(y^{1-\gamma}-1\right) + (1-\gamma)\left(y^{-\gamma}-1\right) = 0.$$

This contradicts $\partial p^*/\partial x < 0$.

E Other Tables

Table 5 NLSY97: marginal effect of socioeconomic status on dropout probability

	coef.	std. err.
male	0.1224	0.0185
asvab	-0.0024	0.0003
born-US	0.071	0.023
hh-size	-0.0009	0.0067
minority	0.0675	0.0219
socio-low	0.0265	0.0266
socio-high	-0.0901	0.0187
# of obs.	1948	
Sociolow=Sociohigh	$prob > \chi^2$	0.02

To compute the marginal effects we run a logit regression on the probability of dropping out. Male: =1 if male . asvab: Armed Services Vocational Aptitude Battery. born-US: =1 if born in the United States and =0 otherwise. hh-sze: size of household. minority: =1 if black or hispanic and =0 otherwise. Socio-low: =1 if student reported to be of low socioeconomic status. Socio-high: =1 if student reported to be of high socioeconomic status.

Table 6 NLSY79: marginal effect of socioeconomic status on dropout probability

	coef.	std. err.
male	0.0377	0.0412
afqt	-0.0047	0.0009
home-abroad	0.1059	0.0979
city	0.014	0.052
siblings	0.0029	0.0108
country-mother	0.0616	0.0752
minority	0.0656	0.0523
socio-low	0.1576	0.063
socio-high	-0.0588	0.0468
# of obs.	635	
Sociolow=Sociohigh	$prob > \chi^2$	0.03

To compute the marginal effects we run a logit regression on the probability of dropping out. Male: =1 if male . afqt: Armed Forced Qualification Test. home-abroad: =1 if born in the United States and =0 otherwise. city: =1 if living in a city and =0 otherwise. siblings: numbers of siblings. country-mother: =1 if mother was born in the United States and =0 otherwise. minority: =1 if black or hispanic and =0 otherwise. Socio-low: =1 if student reported to be of low socioeconomic status. Socio-high: =1 if student reported to be of high socioeconomic status.

Table 7 NLS-72: marginal effect of socioeconomic status on dropout probability

	coef.	std. err.
male	0.017	0.0192
rank	0.0004	0.0000
minority	-0.0375	0.0275
socio-low	0.0256	0.0298
socio-high	-0.138	0.0211
# of obs.	2705	
Sociolow=Sociohigh	$prob > \chi^2$	0.00

To compute the marginal effects we run a logit regression on the probability of dropping out. Male: =1 if male. Rank: ratio of rank in high-school senior class to total class size. Minority: =1 if race is not white. Socio-low: =1 if student reported to be of low socioeconomic status. Socio-high: =1 if student reported to be of high socioeconomic status.

Table 8 NLSY97: effect of socioeconomic status on time to dropout

	coef.	std. err.
male	-0.0379	0.1866
asvab	0.0138	0.0042
born-US	0.4049	0.338
hh-size	-0.0059	0.074
minority	0.5586	0.2187
socio-low	-0.3793	0.2569
socio-high	0.3365	0.2352
cutoff 1	-0.8217	0.6047
cutoff 2	0.5858	0.5983
cutoff 3	1.3686	0.6025
cutoff 4	1.9868	0.6070
cutoff 5	2.7301	0.6135
cutoff 6	3.8652	0.6311
Pseudo R-2	0.0154	
# of obs.	361	
Sociolow=Sociohigh	$prob > \chi^2$	0.00

Results of an ordered logit regression of the time to drop out. Male: =1 if male . asvab: Armed Services Vocational Aptitude Battery. born-US: =1 if born in the United States and =0 otherwise. hh-sze: size of household. minority: =1 if black or hispanic and =0 otherwise. Socio-low: =1 if student reported to be of low socioeconomic status. Socio-high: =1 if student reported to be of high socioeconomic status.

Table 9 NLSY79: effect of socioeconomic status on time to dropout

	coef.	std. err.
male	0.4057	0.2935
afqt	0.008	0.0063
home-abroad	-0.0366	0.7184
city	-0.0619	0.4248
no-mother	0.9776	1.1664
siblings	-0.0518	0.0652
country-mother	-0.1982	0.5102
minority	-0.3258	0.3412
socio-low	-0.9225	0.3871
socio-high	0.0343	0.3631
cutoff 1	-2.684	0.9707
cutoff 2	-0.7625	0.933
cutoff 3	0.1705	0.9329
cutoff 4	0.7995	0.9365
cutoff 5	1.6166	0.946
cutoff 6	2.4725	0.9694
Pseudo R-2	0.0326	
# of obs.	158	
Sociolow=Sociohigh	$prob > \chi^2$	0.00

Results of an ordered logit regression of the time to drop out. Male: =1 if male . afqt: Armed Forced Qualification Test. home-abroad: =1 if born in the United States and =0 otherwise. city: =1 if living in a city and =0 otherwise. siblings: numbers of siblinigs. country-mother: =1 if mother was born in the United States and =0 otherwise. minority: =1 if black or hispanic and =0 otherwise. Socio-low: =1 if student reported to be of low socioeconomic status. Socio-high: =1 if student reported to be of high socioeconomic status.

Table 10 NLS72: effect of socioeconomic status on time to dropout

	coef.	std. err.
male	0.2148	0.0911
rank	-0.0008	0.0003
minority	0.3444	0.1225
socio-low	-0.425	0.1339
socio-high	0.5285	0.1023
cutoff 1	-0.3672	0.0926
cutoff 2	0.9031	0.095
cutoff 3	1.3934	0.0988
cutoff 4	2.2977	0.1125
cutoff 5	3.3964	0.1492
cutoff 6	4.099	0.1917
Pseudo R-2	0.0116	
$\#$ of obs.	1610	
Sociolow=Sociohigh	$prob > \chi^2$	0.00

Results of an ordered logit regression of the time to drop out. Male: =1 if male. Rank: ratio of rank in high school senior class to total class size. Minority: =1 if race is not white. Socio-low: =1 if student reported to be of low socioeconomic status. Socio-high: =1 if student reported to be of high socioeconomic status.