Economic Quarterly—Volume 100, Number 4—Fourth Quarter 2014—Pages 279–295

The Dropout Option in a Simple Model of College Education

Ali Ozdagli and Nicholas Trachter

It has been estimated that college graduates earn substantially more than those who do not pursue higher education (Kane and Rouse 1995; Heckman, Lochner, and Todd 2008). But the high rates of return associated with college graduation seem at odds with the low enrollment and graduation rates observed in the data (Athreya and Eberly 2015). Most of the literature addresses this apparent inconsistency in two alternative ways. The first way introduces credit constraints that prevent some students from enrolling and some others from graduating (Keane and Wolpin 2001, Cameron and Taber 2004, Stinebrickner and Stinebrickner 2008, and Lochner and Monge-Naranjo 2011). The second way models college education as a process of learning about self-ability (Stinebrickner and Stinebrickner 2012, Strange 2012, Hendricks and Leukhina 2014, and Trachter 2015). The evidence in favor of the first way is mixed, and thus the second method gained attention in recent years. In this article we aim to provide a simple laboratory to provide some sense of the workings of the learning model.

We build a simple continuous-time model where high school graduates are uncertain about their innate ability to accumulate human capital in college.[1] They are endowed with a belief about their initial ability level. Those with pessimistic initial beliefs find it optimal not to pursue higher education and instead join the workforce, while

Ozdagli is a senior economist with the Boston Fed. Trachter is an economist with the Richmond Fed. The views expressed in this article are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Boston, the Federal Reserve Bank of Richmond, or the Federal Reserve System. E-mail: nicholas.trachter@rich.frb.org.

[1] The modeling approach relates to the ones in Miao and Wang (2007) and Ozdagli and Trachter (2011).

those with optimistic beliefs enroll in college. During college education students are confronted with random events, which we label as exams, that convey information about the student's true ability level. This information causes students to update their beliefs, making students either more or less optimistic about the expected wage increase upon obtaining a college degree. Those that become very pessimistic about it find it optimal to drop out and join the workforce without a college degree.

We make several assumptions to keep the model simple and tractable. Although some of the assumptions look stark, they allow not only for an extremely tractable model, but also allow us to solve the model in closed form for the college enrollment rate, dropout rate, and wage premium upon graduation. These three objects are at the core of the puzzling nature of college education described above. With this in mind, it is natural to calibrate the model to match these three figures. The model, for its given extreme simplicity, does a good job in matching these objects.

We use the model to gauge the importance of the dropout option. We do so by calculating the fraction of the returns to college enrollment that comes from the availability of this option relative to an alternative scenario where students are not allowed to drop out. We find that a large fraction of measured returns are associated with the fact that students are allowed to drop out.

The rest of the article is organized as follows. Section 1 presents the framework. Section 2 characterizes the problem of a worker, while Section 3 characterizes the problem of a student. Section 4 pertains to the calibration exercise, while Section 5 presents a simple measurement of the value attached to the dropout option. Finally, Section 6 concludes.

1. FRAMEWORK

We build a simple, continuous-time model of postsecondary education. Upon high school graduation, an individual has to decide between joining the workforce or pursuing a degree in a (four-year) college.[2] Agents are endowed with an initial wealth level a_0, and they differ in their ability to accumulate human capital in college. For simplicity, we assume that ability can take two levels: low and high. Let $\mu \in \{0, 1\}$ denote the ability level, with $\mu = 0$ denoting low ability. The ability level is

[2] Although relevant, for simplicity, we abstract from the decision to enroll or transfer to/from two-year colleges.

not observable, and the high school graduate is endowed with a belief $p_0 = \Pr(\mu = 1)$.

At any point in time, an agent can be working or attending college. We assume that working is an absorbing state, and that a college student graduates with intensity (the analogous of probability in a discrete time setup) ϕ, so that the expected time until obtaining a degree is $1/\phi$. Upon college graduation, the true ability level of the agent is revealed. We denote with τ college tuition, and by $h(\mu, GS)$ the wage, in any given period, of an agent joining the workforce with true ability level μ, and graduation status $GS \in \{0, 1\}$, where 0 implies that the agent does not hold a college degree, and 1 does. In particular, we let $h(0,0) = h(1,0) \equiv \underline{h}$, and $h(1,1) \equiv \bar{h} > h(0,1) = \underline{h}$. That is, there is a graduation premium only if the ability level of the agent is high. Notice that it follows that the process for wealth accumulation can be written as $da_t/dt = ra_t + h(\mu, GS) - c_t$ if the agent is working, and by $da_t/dt = ra_t - \tau - c_t$ if the agent is enrolled in college, where c_t denotes the consumption level at time t.

While attending college, students are faced with exams every period. The exams have two alternative outcomes: a passing grade or a good grade. We assume that only high-ability students are able to obtain a good grade, something that occurs with intensity λ.[3] Because the grade obtained in a particular exam is correlated with the ability level of the student, grades convey information that the students use to update their beliefs: Upon receiving a good grade, the belief of a student "jumps" from p to 1, because only high-ability students can obtain a good grade, while, upon receiving a passing grade, students update their beliefs by Bayes' rule. Suppose that the period length is given by Δ. Bayes' rule implies that the belief at period $t + \Delta$ is given by

$$p_{t+\Delta} = \frac{(1 - \lambda\Delta)p_t}{p_t(1 - \lambda\Delta) + (1 - p_t)} \ .$$

The denominator, $p_t(1-\lambda\Delta)+(1-p_t)$, accounts for the probability that we observe belief $p_{t+\Delta}$ at time $t+\Delta$, while the numerator accounts for the probability that we observe the high ability level $\mu = 1$ (which is conditional on the student being of high ability) times the belief that this event could occur, p_t. Subtracting p_t on both sides and dividing

[3] It is easy to add more grades and to correlate them with the ability level of students. However, this would add more parameters to the model, increasing the complexity of the calibration exercise.

by Δ provides

$$\frac{p_{t+\Delta}-p_t}{\Delta} = -p_t(1-p_t)\frac{\lambda}{p_t(1-\lambda\Delta)+(1-p_t)} \ .$$

Notice that $dp_t/dt = \lim_{\Delta\downarrow 0}\frac{p_{t+\Delta}-p_t}{\Delta}$. Then, taking the limit as Δ approaches zero provides that

$$dp_t/dt = -\lambda p_t(1-p_t) \ .$$

The learning equation shows that, conditional on the student not receiving a good grade, the beliefs of a student fall through time. This follows because as time passes and no fully revealing signal is received, the student updates her beliefs toward believing that her ability level is low. Further, notice that the speed of learning depends on (i) the intensity parameter λ regulating the probability of receiving a fully revealing signal, and (ii) the current level of beliefs p_t. A higher λ implies that signals arrive at a faster rate and thus not receiving the signal conveys more information too (the student updates her beliefs faster toward being of low ability). The process depends on p_t, as this regulates how uncertain is the student. In particular, the highest uncertainty level is $p_t = 1/2$, which maximizes the value $p_t(1-p_t)$. The higher the degree of uncertainty, the more relevant the arrival of information becomes.

A high school graduate, endowed with p_0 and a_0, chooses her consumption stream c_t, whether to enroll and/or remain in college at every point in time, in order to maximize her time-separable expected discounted lifetime utility, with period utility function $u(c_t) = e^{\gamma c_t}/-\gamma$, where γ is the coefficient of constant absolute risk aversion. We further assume that the rate of discount equals the interest rate, and we denote both by r.

Let $V(p,a)$ denote the value for a college student with current belief p and wealth level a, and let $W(h(\mu,GS),a)$ denote the value for a worker with wage $h(\mu,GS)$ and wealth a. Given the structure of the problem, we start by solving the problem of workers; we then tackle the problem of students, and we finalize by obtaining the optimal policy of high school graduates.

2. THE PROBLEM OF A WORKER

We start by describing the problem of a worker. The value function $W(h(\mu,GS),a)$ solves

$$rW(h(\mu,GS),a) = \max_c \frac{e^{-\gamma c}}{-\gamma} + W_a(h(\mu,GS),a)(ra + h(\mu,GS) - c) \ .$$

The equation states that the flow value of being a worker with wage $h(\mu, GS)$ and wealth a, is equal to the instantaneous utility derived from consumption, in addition to the change in value accrued to the change in the wealth level of the worker. Notice that this is a standard savings problem, with solution given by $W(h(\mu, GS), a) = -\frac{e^{-\gamma(ra+h(\mu,GS))}}{\gamma r}$.

3. THE PROBLEM OF A STUDENT

The value function for a student with current belief p and wealth level a solves

$$\begin{aligned} rV(p,a) &= \max_c \frac{e^{-\gamma c}}{-\gamma} + p\lambda[V(1,a) - V(p,a)] \\ &\quad +\phi[pW(\bar{h},a) + (1-p)W(\underline{h},a) - V(p,a)] \\ &\quad -V_p(p,a)\lambda p(1-p) + V_a(p,a)(ra - \tau - c) \ . \end{aligned}$$

Let us provide some intuition to this Bellman equation. The left-hand side of the equation measures the flow value of being a student with current belief p and wealth a. This value has to be equal to the sum of five different terms. The first term is the instantaneous utility derived from consumption. The second term accounts for the change in value if the student is revealed to be of high ability, which, conditional on the student being of high ability (which occurs with probability p), happens with intensity λ. The third term accounts for the change in value if the student graduates (remember that upon graduation ability is revealed). Finally, the fourth and fifth terms account for the change in value accrued for the change in beliefs and wealth, if the student remains enrolled in school and no fully revealing signal had arrived.

Solving for the value function of students of unknown type requires us first to obtain an expression for the value function when the student is of high ability, $V(1, a)$. In Appendix A we solve for this value function. We obtain that $V(1, a) = -b\frac{e^{-\gamma(ra+\bar{h})}}{\gamma r}$, where $b > 1$ is the unique solution to $\phi/b = \phi - \gamma r(\bar{h} + \tau) + r \ln b$. We will restrict our attention to cases where $V(1, a) > W(\underline{h}, a)$, so that high-ability students always find it worthwhile to enroll in college and remain until graduation.

Using the expressions for the value function for a worker and for a student with $p = 1$, we can now solve for the value function of a student of unknown type. Details on the derivation can be found in Appendix B. We obtain that $V(p, a) = -\frac{e^{-\gamma(ra+f(p))}}{\gamma r}$, where $f(p)$ solves

$$(\phi + p\lambda - \gamma f'^{-\gamma f(p)} = p\lambda b e^{-\gamma\bar{h}} + \phi\left[pe^{-\gamma\bar{h}} + (1-p)e^{-\gamma\underline{h}}\right] \ .$$

In Appendix C we show that $f(p)$ is increasing. Although the proof is somewhat involved, the result is intuitive. The idea is that high-ability students are the ones who obtain a wage boost after graduation, and also are the ones likely to receive positive signals. As a result, a higher belief p implies a higher expected gain from college enrollment.

The decision to enroll in college involves comparing the value accrued if an agent enrolls, $V(p, a)$, with the value if she drops out, $W(\underline{h}, a)$. Notice that these two functions are multiplicative in $e^{-\gamma ra}$, and thus the optimal enrollment policy should not depend on the wealth level a. This is a natural implication of the exponential utility function used in the article that implies no income effects. As a result, the enrollment decision depends solely on the belief of the student. Because the value of college enrollment is increasing in the student's belief p—because $f(p)$ is increasing in p, we conjecture that there exists a threshold p^* such that those students with beliefs $p < p^*$ do not attend college, while those with $p > p^*$ pursue higher education. Furthermore, because there is no direct cost to be paid after enrollment or dropping out from college, the threshold p^* also regulates who drops out. That is, the decision about whether to remain enrolled at college with belief p is identical to the decision of whether to start college with a belief of p. Because the value function $V(p, a)$ is continuous in p, it has to be the case that a student with prior p^* and wealth level a is indifferent between working and enrolling in college. That is, $V(p^*, a) = W(\underline{h}, a)$, which is known as the value matching condition. Also, at the indifference point, it has to be the case that there is no extra value of staying in school for the marginal student. That is, $V_p(p^*, a) = 0$, a condition known as the smooth pasting condition. Using the expressions we found for $V(p, a)$ and $W(\underline{h}, a)$, we can reduce these two conditions to $f(p^*) = \underline{h}$, and $f'^*) = 0$. We can combine these two conditions with the expression that $f(p)$ has to satisfy to solve for p^*,

$$p^* = \frac{\gamma r(\tau + \underline{h})}{\phi\left(1 - e^{-\gamma(\bar{h} - \underline{h})}\right) + \lambda\left(1 - be^{-\gamma(\bar{h} - \underline{h})}\right)} \ .$$

Before concluding this section, it is useful to define the value of college attendance. Let $\sigma(p)$ denote the maximum amount of wealth a high school graduate with belief p is willing to forgo to have the option to attend college. Notice that $\sigma(p)$ solves $V(p, a - \sigma(p)) = W(\underline{h}, a)$. Immediate calculations provide that $\sigma(p) = (f(p) - \underline{h})/r$. Notice that the value of attending college for the student at the margin is zero. That is, $\sigma(p^*) = 0$.

4. CALIBRATION: A PARAMETRIC EXAMPLE

In this section we aim to provide a simple calibration of the model in order to be able to gauge the option value provided to students by having the opportunity to drop out from school. We will propose a simple way for constructing initial beliefs, i.e., p_0, and we will parameterize the model by a combination of direct imputation and calibration exercise.

Imputed Parameters

We set the interest rate r to be 0.02, which implies a yearly discount factor of 0.98. We standardize $\underline{h} \equiv 1$, we set the risk aversion parameter $\gamma = 8$, and, following Trachter (2015), we set $\tau = 0.32$.[4]

Calibrated Parameters

The remaining parameters, λ (the learning parameter), $\bar{h}$ (the wage premium parameter), and s (the share of high-ability students), are estimated by a method of moments. With this in mind, we build three moments that depend on these three parameters: the college enrollment rate (C^e), the college dropout rate (C^d), and the average wage premium of college graduates (W^p).

To calibrate the model, we need to specify the true ability level of agents and their initial beliefs. We do so by setting up density functions from where the ability levels are drawn, and we then compute, using Bayes' rule, the initial belief of each individual. There are alternative ways of producing initial beliefs. Some studies, such as Hendricks and Leukhina (2014) and Trachter (2015), estimate initial beliefs by the means of parametric models that link them to observable and unobservable measures of initial ability. A simpler approach that does not rely on microdata would be to specify a structure of initial beliefs and figure out the mass of the individual at each belief level so as to match moments in the data. The method we follow in this article follows this idea, but with the further assumption that we impose structure on how we relate individual types, initial beliefs, and the measure of agents at each belief level.

We normalize the total mass of high school graduates to be one. We let s denote the share of high-ability students. Each agent receives a signal, q_0, drawn from a distribution that is type dependent, $\Gamma_\mu(q_0)$.

[4] The value for τ follows from obtaining the average expenditures in tuition (including room and board), and then normalizing by the wage of those not pursuing college education and dropouts. More details can be found in Trachter (2015).

For any given q_0, then, p_0 is just the result of Bayes' rule, and the gamma distributions, transformed by Bayes' rule, provide the distribution of initial beliefs p_0. For simplicity, we assume that $\Gamma_1(q_0) = 2q_0$, and $\Gamma_0(q_0) = 2 - 2q_0$, for $q_0 \in [0, 1]$, to capture the idea that a high q_0 (and thus a high p_0) is more likely to be drawn from the distribution for high-ability students. Bayes' rule provides that

$$p_0 \equiv \Pr(\mu = 1 | q_0) = \frac{\Pr(q_0 | \mu = 1) \Pr(\mu = 1)}{\Pr(q_0)} = \frac{q_0 s}{s q_0 + (1 - s)(1 - q_0)} . \tag{1}$$

We now compute the three required moments. We start by providing an expression for the college enrollment rate, which is given by

$$\begin{aligned} C^e &= \int_{q(p^*)}^{1} (h_0(q_0)(1 - s) + h_1(q_0)s) dq_0 \\ &= 1 - 2\left[(1 - s)q(p^*) + (2s - 1)\frac{q(p^*)^2}{2}\right] , \end{aligned}$$

where $q(p^*)$ follows from inverting the expression in (1).

The college dropout rate is given by

$$\begin{aligned} C^d &= s\int_{q(p^*)}^{1} Q(p(q_0), 1)\frac{h_1(q_0)}{1 - H_1(q^*)} dq_0 \\ &\quad +(1 - s)\int_{q(p^*)}^{1} Q(p(q_0), 0)\frac{h_0(q_0)}{1 - H_0(q^*)} dq_0 , \end{aligned}$$

where $Q(p, \mu)$ denotes the dropout probability of a college student of type μ and belief p (see Appendix D for a derivation of the expressions for $Q(p, \mu)$). Some algebra provides that

$$\begin{aligned} C^d &= s\frac{2}{1 - q(p^*)^2}\left(\frac{p^*}{1 - p^*}\frac{1 - s}{s}\right)^{\frac{\phi+\lambda}{\lambda}} Beta\left(q(p^*), 2 + \frac{\phi}{\lambda}, 1 - \frac{\phi}{\lambda}\right) \\ &\quad +(1 - s)\frac{2}{1 - 2q(p^*) + q(p^*)^2}\left(\frac{p^*}{1 - p^*}\frac{1 - s}{s}\right)^{\frac{\phi}{\lambda}} \\ &\quad Beta\left(q(p^*), 2 + \frac{\phi}{\lambda}, 1 - \frac{\phi}{\lambda}\right) , \end{aligned}$$

where details of the derivation can be obtained in Appendix E, and where $Beta(\cdot, \cdot, \cdot)$ denotes an incomplete Beta function, and where it is required that $\lambda > \phi$.

Finally, the wage premium upon graduation is given by

$$W^p = \bar{h}\frac{sN_1}{sN_1 + (1 - s)N_0} + \underline{h}\frac{(1 - s)N_0}{sN_1 + (1 - s)N_0} ,$$

Table 1 Moments in Model and Data

	Model	Data
Enrollment rate	0.34	0.25
Dropout rate	0.38	0.4
Wage premium	0.34	0.21

where the amount sN_1 ($(1-s)N_0$) accounts for the mass of high- (low-) ability students who graduate, where $\int_{q(p^*)}^{1}(1-Q(p(q_0),1))\,h_1(q_0)dq_0$ and $N_0=\int_{q(p^*)}^{1}(1-Q(p(q_0),0))\,h_0(q_0)dq_0$.

Following a similar method as the one used for the dropout rate, we can find expressions for N_1 and N_0,

$$\begin{aligned} N_1 &= 1-q(p^*)^2-2\left(\frac{p^*}{1-p^*}\frac{1-s}{s}\right)^{\frac{\phi+\lambda}{\lambda}} Beta\left(q(p^*),2+\frac{\phi}{\lambda},1-\frac{\phi}{\lambda}\right), \\ N_0 &= 1+q(p^*)^2-2q(p^*)-2\left(\frac{p^*}{1-p^*}\frac{1-s}{s}\right)^{\frac{\phi}{\lambda}} \\ &\quad Beta\left(q(p^*),2+\frac{\phi}{\lambda},1-\frac{\phi}{\lambda}\right). \end{aligned}$$

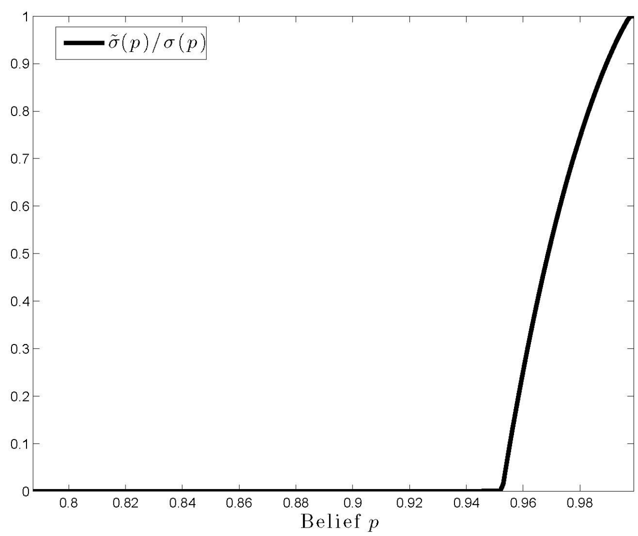
We aim to calibrate $\bar{h}$, λ, and s by looking for the values of these three parameters that make the model implied enrollment rate, dropout rate, and average wage premium to be as close as possible to its empirical counterpart. We again borrow the empirical figures from Trachter (2015), which computed these figures from the *National Longitudinal Study of the High School Class of 1972* (NLS-72). Table 1 presents the data and model-implied moments. The model, for its given simplicity, does a good job in matching the data.[5] It overly predicts college enrollment and wage premium of graduates, while the college dropout rate is very close to its empirical counterpart. This exercise provided us with estimates for the remaining parameters: $\bar{h}=1.36$, $\lambda=0.241$, and $s=0.6$.

5. THE VALUE OF THE DROPOUT OPTION

In this section, we provide a simple measurement on the valued added by the dropout option. We do so by comparing the gains from college enrollment that a student gets if the option is available, $\sigma(p)$, with those

[5] The fit of the model can be improved in several ways, as done in Trachter (2015). For example, by using a less restrictive belief construction or learning process.

Figure 1 Decomposition of Returns



returns if the option is removed, which we define by $\tilde{\sigma}(p)$.[6] Notice that $\tilde{\sigma}(p)/\sigma(p)$ will measure the fraction of the gains that follow from the fact that students are allowed to enroll, while $1 - \tilde{\sigma}(p)/\sigma(p)$ measures the fraction of the gains that follow from the fact that students are also allowed to drop out.

Figure 1 presents the decomposition of returns to enrollment for all the beliefs that imply college enrollment (i.e., $p > p^* = 0.7875$). For a large fraction of beliefs (up to $p \approx 0.95$), the dropout option accounts for all of the returns to college enrollment. This follows because students with beliefs in this range do not find it very likely to graduate and be of high ability, so they value highly the dropout opportunity. As the belief keeps rising toward one, the enrollment option gains value as students are more likely to be of high ability and so value less the dropout option. Overall, the figure shows that the dropout option accounts for a large part of the returns to enrollment. In fact, using the densities for the initial types, we can compute the average value added of the dropout option: The dropout option accounts for 85 percent of the value of college enrollment.

[6] We do not include this second model in the article, but it can be easily computed following a similar procedure to the one used for the model with the dropout option.

6. CONCLUDING REMARKS

This article proposes a simple and highly tractable model of postsecondary education. In the model, students are allowed to drop out at any point in time as they update their beliefs about the gains accrued from college education. For its apparent simplicity, the parameterized version of the model is able to generate patterns consistent with the data. As a result, this article showcases the potential of learning models to explain the patterns observed in postsecondary education.

The simplified nature of the model implied that we abstracted from several noteworthy elements that any exercise attempting to gauge the returns to education might need to address. First, the model does not provide any wage premium for graduates of low ability. Lifting this assumption is very simple and would allow us to gauge more precisely the importance of the dropout option as, without the wage premium for low-ability students, the model overestimates the importance of the dropout option. Although, as shown in Trachter (2015), the qualitative properties of this article would survive. Second, the model does not allow for any wealth effects, something undesirable if we want to explore the connection between wealth differences and education outcomes. In Ozdagli and Trachter (2011), we evaluate a setup similar to the one developed here but where wealth differences affect the risk preferences of students. We show that, consistent with the data, wealthier students are less likely to drop out and, conditional on this event, they drop out later than poorer students.

The goal of this article is to introduce the reader to a particular class of models being used to analyze postsecondary patterns of education. These models are highly tractable, are easy to handle, and are able to fit several salient features of the data (for example, see Hendricks and Leukhina [2014] and Trachter [2015]). The ongoing success of these models calls for further research in the area.

APPENDIX A: A STUDENT OF HIGH ABILITY

Evaluating the problem of a student at $p = 1$ provides

$$rV(1,a) = \max_{c} \frac{e^{-\gamma c}}{-\gamma} + \phi[W(\bar{h},a) - V(1,a)] + V_a(1,a)(ra - \tau - c) \ .$$

Notice that standard techniques imply that the solution to this problem is unique. Plugging in the first-order condition provides

$$rV(1,a) = \frac{V_a(1,a)}{-\gamma} + \phi[W(\bar{h},a) - V(1,a)] + V_a(p,a)(ra - \tau + \gamma^{-1}\ln V_a(1,a)) \, .$$

We look for the solution to this problem by a guess and verify method, which readily provides the expression provided in the main text.

APPENDIX B: A STUDENT OF UNKNOWN TYPE

The problem of the student can be rewritten as

$$\begin{aligned}(r+\phi+p\lambda)V(p,a) \quad = \quad & \max_c \frac{e^{-\gamma c}}{-\gamma} + p\lambda V(1,a) \\ & +\phi[pW(\bar{h},a) + (1-p)W(\underline{h},a)] \\ & -V_p(p,a)\lambda p(1-p) + V_a(p,a)(ra - \tau - c) \, .\end{aligned}$$

As with the problem of a worker, it is straightforward to argue that this problem has a unique solution. The first-order condition provides $e^{-\gamma c} = V_a(p,a)$. Solving for c, we get that $c = -\gamma^{-1}\ln V_a(p,a)$. Using these two expressions, we obtain that

$$\begin{aligned}(r+\phi+p\lambda)V(p,a) \quad = \quad & \frac{V_a(p,a)}{-\gamma} + p\lambda V(1,a) + \phi[pW(\bar{h},a) \\ & +(1-p)W(\underline{h},a)] - V_p(p,a)\lambda p(1-p) \\ & +V_a(p,a)(ra - \tau + \gamma^{-1}\ln V_a(p,a)) \, .\end{aligned}$$

Notice that this last expression implies that the value function $V(p,a)$ has to satisfy a partial differential equation. We solve for $V(p,a)$ by a guess and verify method. We guess that $V(p,a) = -\frac{e^{-\gamma(ra+f(p))}}{\gamma r}$, where $f(p)$ is a function to be determined. Under this guess, the previous expression reduces to

$$(\phi + p\lambda - \gamma f'^{-\gamma f(p)} = p\lambda b e^{-\gamma\bar{h}} + \phi\left[pe^{-\gamma\bar{h}} + (1-p)e^{-\gamma\underline{h}}\right] \, ,$$

where we also used the solution to the worker problem and the problem of a student with belief $p = 1$.

APPENDIX C: DETAILS ON $f(p)$

Here we show that $f(p)$ is increasing in p. To do so, notice that we can rewrite the expression defining $f(p)$ as $-\gamma f'(p)\lambda p(1-p) = \theta(p, f(p))$, where

$$\begin{aligned}\theta(p, f(p)) \;\; &\equiv \;\; p(\lambda b+\phi)e^{-\gamma(\bar{h}-f(p))} + \phi(1-p)e^{-\gamma(\underline{h}-f(p))} \\ &\quad -(\phi + p\lambda - \gamma r(\tau + f(p))) \; .\end{aligned}$$

Notice that

$$\frac{\partial\theta(p, f(p))}{\partial f(p)} = \gamma p(\lambda b+\phi)e^{-\gamma(\bar{h}-f(p))} + \gamma\phi(1-p)e^{-\gamma(\underline{h}-f(p))} + \gamma r > 0 \; ,$$

and

$$\begin{aligned}\frac{\partial\theta(p, f(p))}{\partial p} \;\; &= \;\; (\lambda b+\phi)be^{-\gamma(\bar{h}-f(p))} - \phi e^{-\gamma(\underline{h}-f(p))} - \lambda \\ &= \;\; \phi e^{\gamma f(p)}\left(e^{-\gamma\bar{h}} - e^{-\gamma\underline{h}}\right) + \lambda\left(be^{-\gamma(\bar{h}-f(p))} - 1\right) \; .\end{aligned}$$

To sign this last expression we need to notice first that the highest attainable value of college enrollment is $V(a, 1)$, so that $V(a, p) \leq V(a, 1)$ for all p. This condition reduces to $be^{-\gamma(\bar{h}-f(p))} - 1$. Then, we obtain that $\frac{\partial\theta(p,f(p))}{\partial p} < 0$. We use the sign of these two derivatives to show that $f(p)$ increases with p.

We prove that $f'(p) > 0$ by a contradiction argument. Suppose there exists a belief p_1, $p_1 > p^*$, such that $f'(p_1) \leq 0$. Then, it has to be the case that $\theta(p_1, f(p_1)) \geq 0$. Pick a belief p_2 in the neighborhood of p_1 that satisfies $p_2 < p_1$. Because $f'(p_1) \leq 0$, we have that $f(p_2) \geq f(p_1)$. Then, because $\theta(\cdot)$ increases in p and decreases in $f(p)$, we have that $\theta(p_2, f(p_2)) > \theta(p_1, f(p_1)) \geq 0$. This provides that $f'(p_2) < 0$. Iterating on this process provides that $f'^{*}) < 0$, which contradicts the fact that, by construction, $f'^{*}) = 0$. Then, $f'(p) > 0$ for all p.

APPENDIX D: DROPOUT PROBABILITY

Let $Q(p, \mu)$ denote the dropout probability of a student with current belief p and true ability level μ. Notice that we are not including the wealth level a as a state, as we already concluded that the dropout threshold is independent of a. Also, notice that, for a given belief p, a high-ability student is less likely to drop out than a low-ability student given that $\lambda > 0$, $V(1,a) > W(\underline{h},a)$.

For a high-ability student, we have that

$$Q(p,1) = \phi dt \cdot 0 + (1-\phi dt)\left[\lambda dt Q(1,1) + (1-\lambda dt)Q(p-\lambda p(1-p)dt,1)\right],$$

where $dp/dt = -\lambda p(1-p)$. Given that, $V(1,a) > W(\underline{h},a)$ implies that $Q(1,1) = 0$. Also, noticing that $Q(p-\lambda p(1-p)dt,1) \approx Q(p,1) - Q_p(p,1)\lambda p(1-p)dt$, we can rewrite the previous expression as

$$\begin{aligned}
Q(p,1) &= (1-\phi dt)(1-\lambda dt)\left[Q(p,1) - Q_p(p,1)\lambda p(1-p)dt\right] \ , \\
Q(p,1) &= (1-(\phi+\lambda)dt + \phi\lambda dt^2)Q(p,1) \\
&\quad -(1-\phi dt)(1-\lambda dt)Q_p(p,1)\lambda p(1-p)dt \ , \\
Q(p,1) &= -\frac{(1-\phi dt)(1-\lambda dt)Q_p(p,1)\lambda p(1-p)}{((\phi+\lambda)-\phi\lambda dt)} \ .
\end{aligned}$$

Taking the limit as dt approaches zero provides that $Q(p,1)$ satisfies a first-order linear ordinary differential equation, $\frac{Q_p(p,1)}{Q(p,1)} = -\frac{\phi+\lambda}{\lambda p(1-p)}$ with boundary condition $Q(p^*,1) = 1$. The general solution to the differential equation is $Q(p,1) = C_1\left(\frac{1-p}{p}\right)^{\frac{\phi+\lambda}{\lambda}}$, where C_1 is such that $Q(p^*,1) = 1$. It is immediate to obtain that $C_1 = \left(\frac{p^*}{1-p^*}\right)^{\frac{\phi+\lambda}{\lambda}}$, so that $Q(p,1) = \left(\frac{p^*}{p}\frac{1-p}{1-p^*}\right)^{\frac{\phi+\lambda}{\lambda}}$.

We now derive $Q(p,0)$. This value solves

$$Q(p,0) = \phi dt \cdot 0 + (1-\phi dt)Q(p-\lambda p(1-p)dt,0) \ .$$

As we did above, from this expression we obtain that $Q(p,0)$ satisfies $\frac{Q_p(p,0)}{Q(p,0)} = -\frac{\phi}{\lambda p(1-p)}$, with boundary condition $Q(p^*,0) = 1$. It follows that $Q(p,0) = \left(\frac{p^*}{p}\frac{1-p}{1-p^*}\right)^{\frac{\phi}{\lambda}}$.

APPENDIX E: COMPUTING DROPOUT RATES

We begin by computing $\int_{q(p^*)}^{1} Q(p(q_0),1)\frac{h_1(q_0)}{1-H_1(q^*)}dq_0$. Using the expression for the density $h_1(q_0)$, together with the expression for $Q(p,1)$

derived in Appendix D, simple algebra provides that

$$\begin{aligned}&\int_{q(p^*)}^{1} Q(p(q_0),1)\frac{h_1(q_0)}{1-H_1(q(p^*))}dq_0\\ =\;& \frac{2}{1-q(p^*)^2}\left(\frac{p^*}{1-p^*}\frac{1-s}{s}\right)^{\frac{\phi+\lambda}{\lambda}}\int_{q(p^*)}^{1} q_0^{1-\frac{\phi+\lambda}{\lambda}}(1-q_0)^{\frac{\phi+\lambda}{\lambda}}dq_0\ ,\end{aligned}$$

where, from (1), we can obtain that $\frac{1-p}{p}=\frac{1-q_0}{q_0}\frac{1-s}{s}$. Define $y\equiv 1-q_0$. This change of variables allows us to rewrite the previous expression as

$$\begin{aligned}&\int_{q(p^*)}^{1} Q(p(q_0),1)\frac{h_1(q_0)}{1-H_1(q(p^*))}dq_0\\ =\;& \frac{2}{1-q(p^*)^2}\left(\frac{p^*}{1-p^*}\frac{1-s}{s}\right)^{\frac{\phi+\lambda}{\lambda}}\int_{0}^{1-y(p^*)} y^{\frac{\phi+\lambda}{\lambda}}(1-y)^{1-\frac{\phi+\lambda}{\lambda}}dy\ .\end{aligned}$$

Notice that the expression $\int_0^{1-y(p^*)} y^{\frac{\phi+\lambda}{\lambda}}(1-y)^{1-\frac{\phi+\lambda}{\lambda}}dy$ accounts for an incomplete Beta function $Beta(q(p^*),1+\frac{\phi+\lambda}{\lambda},2-\frac{\phi+\lambda}{\lambda})$. Then,

$$\begin{aligned}&\int_{q(p^*)}^{1} \frac{Q(p(q_0),1)h_1(q_0)}{1-H_1(q(p^*))}dq_0\\ =\;& \frac{2}{1-q(p^*)^2}\left(\frac{p^*}{1-p^*}\frac{1-s}{s}\right)^{\frac{\phi+\lambda}{\lambda}} Beta\left(q(p^*),1+\frac{\phi+\lambda}{\lambda},2-\frac{\phi+\lambda}{\lambda}\right)\ .\end{aligned}$$

We now turn to compute the term $\int_{q(p^*)}^{1} Q(p(q_0),0)\frac{h_0(q_0)}{1-H_0(q^*)}dq_0$. Again, we use the expression for $h_0(q_0)$ and the expression for $Q(p,0)$ derived in Appendix D to obtain

$$\begin{aligned}&\int_{q(p^*)}^{1} Q(p(q_0),0)\frac{h_0(q_0)}{1-H_0(q^*)}dq_0\\ =\;& \frac{2}{1-2q(p^*)+q(p^*)^2}\left(\frac{p^*}{1-p^*}\frac{1-s}{s}\right)^{\frac{\phi}{\lambda}}\int_{q(p^*)}^{1}\left(\frac{1-q_0}{q_0}\right)^{\frac{\phi}{\lambda}}(1-q_0)dq_0\ .\end{aligned}$$

The same change of variables we used above provides

$$\begin{aligned}&\int_{q(p^*)}^{1} Q(p(q_0),0)\frac{h_0(q_0)}{1-H_0(q^*)}dq_0\\ =\;& \frac{2}{1-2q(p^*)+q(p^*)^2}\left(\frac{p^*}{1-p^*}\frac{1-s}{s}\right)^{\frac{\phi}{\lambda}}\int_{0}^{1-y(p^*)} y^{1+\frac{\phi}{\lambda}}(1-y)^{-\frac{\phi}{\lambda}}dy\ ,\end{aligned}$$

from where it follows that

$$\begin{aligned}&\int_{q(p^*)}^{1} Q(p(q_0),0)\frac{h_0(q_0)}{1-H_0(q^*)}dq_0\\ =\;& \frac{2}{1-2q(p^*)+q(p^*)^2}\left(\frac{p^*}{1-p^*}\frac{1-s}{s}\right)^{\frac{\phi}{\lambda}} Beta\left(q(p^*),2+\frac{\phi}{\lambda},1-\frac{\phi}{\lambda}\right)\ .\end{aligned}$$

REFERENCES

Athreya, Kartik, and Janice Eberly. 2015. "The College Premium, College Noncompletion, and Human Capital Investment." Federal Reserve Bank of Richmond Working Paper 13-02R (February).

Cameron, Stephen V., and Christopher Taber. 2004. "Borrowing Constraints and the Returns to Schooling." *Journal of Political Economy* 112 (February): 132–82.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2008. "Earnings Functions and Rates of Return." *Journal of Human Capital* 2 (1): 1–31.

Hendricks, Lutz, and Oksana Leukhina. 2014. "The Return to College: Selection and Dropout Risk." Center for Financial Studies Working Paper 487 (November).

Kane, Thomas J., and Cecilia Elena Rouse. 1995. "Labor-market Returns to Two- and Four-year College." *American Economic Review* 85 (June): 600–14.

Keane, Michael P., and Kenneth I. Wolpin. 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review* 42 (November): 1,051–103.

Lochner, Lance J., and Alexander Monge-Naranjo. 2011. "The Nature of Credit Constraints and Human Capital." *American Economic Review* 101 (October): 2,487–529.

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

Ozdagli, Ali K., and Nicholas Trachter. 2011. "On the Distribution of College Dropouts: Household Wealth and Uninsurable Idiosyncratic Risk." Federal Reserve Bank of Boston Working Paper 11-8 (July).

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

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 (December): 2,163–84.

Stinebrickner, Todd, and Ralph Stinebrickner. 2012. "Learning about Academic Ability and the College Dropout Decision." *Journal of Labor Economics* 30 (October): 707–48.

Trachter, Nicholas. 2015. "Stepping Stone and Option Value in a Model of Postsecondary Education." *Quantitative Economics* 6 (March): 223–56.