College Enrollment, Dropouts and Option Value of Education

Ozdagli, Ali[*] Trachter, Nicholas[‡]

October 11, 2009

Abstract

Returns to postsecondary education have been found to be unreasonably high relative to the low enrollment and graduation rates observed in data. The consensus at the time is that there exists underinvestment in postsecondary education. We model college enrollment and dropout decisions in a real options model similar to Miao and Wang (2007) where attitude towards schooling stems from the initial belief of students about their skill level and the option value of droping out stems from the Bayesian learning process of one's skill level. The option value of droping out is a novel feature of this paper and allows for an easy and straightforward characterization. Our paper has four main results. First, our model shows that the mainstream models in the literature that assume risk neutrality overestimate the value of college, college enrollment and graduation rates because they omit the uncertainty regarding the outcome of college education. Second, we show that the option value of learning is much more important when agents are risk averse rather than risk neutral. Further, the effect of the option value is much more important at the margin so it becomes important for understanding drop-out rates. Finally, we show that ex-ante returns predicted by our model are smaller suggesting that uncertainty about future income stream can potentially solve the high returns to education puzzle.

Returns to postsecondary education have been found to be unreasonably high relative to the low enrollment and graduation rates observed in data. The consensus at the time is that there exists underinvestment in postsecondary education.

[Cunha-Heckman (2007)] construct a counterfactual analysis in order to obtain estimates for college returns. They obtain estimates between 20% to 37%

[*]University of Chicago.

[†]University of Chicago. Contact: trachter@uchicago.edu

[‡]We wish to thank Fernando Alvarez, Lars Hansen, James Heckman, Fabian Lange and Jarda Borovicka for very helpful comments and suggestions. We also want to thank seminar participants of the SED Meetings 2008, UTDT workshop, the Capital Theory Working Group, the Macroeconomics Dynamics Working Group and Education Working Group at The University of Chicago. All errors are our own.

ignoring tuition costs. [Judd (2000)] performs a portfolio analysis to compare whether there is underinvestment in education by comparing returns to college with assets of same risk and illiquidity. Judd concludes that there is underinvestment in education. Both studies abstracted from the dropout option in order to simplify the analysis.

A key aspect of [Cunha-Heckman (2007)] and other influential papers (such as [Keane-Wolpin (1997)] or [Carneiro-Hansen-Heckman (2003)]) that try to match empirial facts related with college enrollment of high school graduates are what are known as Psychic costs. Psychic costs are modeled as an unobservable idiosyncratic component of utility function or income process and are estimated to have high predictive power in terms of college enrollment. Still, as they are known by the high-school graduate prior to deciding whether or not to attend college they fail to explain dropout rates.

Another interesting result in [Cunha-Heckman (2007)] is that *ex-ante* heterogeneity has a very important explanatory power of *ex-post* earnings, while *ex-ante* uncertainty is almost unimportant. The problem is that if uncertainty is not important and the estimated *ex-post* returns are a good approximation of *ex-ante* returns then, even if the dropout option is included in the model, it would be impossible to match the low graduation rates present in data.

In this paper, we focus on college enrollment and dropout decisions in a traceable options model similar to [Miao-Wang (2007)] where attitude towards schooling stems from the initial belief of students about their skill level. The students with very pessimistic initial beliefs do not enroll in college. If a student chooses to enroll in college her initial belief is updated during the college education according to a Bayesian learning process that may eventually lead to dropout decision.

According to our model college education has two main benefits. The first one is the sheepskin or credential effect because many of the high-skill jobs require a college degree. The amount of time spent in college does not add anything to the students' earning potential if they do not finish the college with a degree. The second benefit comes from learning and the option to drop out of college. College students can choose to drop out of college if their beliefs become so pessimistic during the learning process that decide to cut their losses. This second effect can be easily analyzed in our model and can be considered the novel feature of our paper.

We also assume that the students do not face any borrowing constraints because of the availability of college loans. [Cameron-Heckman (2001)], [Keane-Wolpin (2001)], and [Cameron-Taber (2004)] provide empirical support for this assumption.

Our paper has four main results. First, our model shows that papers that assume risk neutrality, such as [Keane-Wolpin (1997)], [Cunha-Heckman (2007)] and [Heckman-Navarro (2006)], overestimate the value of college, college enrollment and graduation rates because they omit the uncertainty regarding the outcome of college education. Second, we show that the option value of learning is much more important when agents are risk averse rather than risk neutral[1].

[1] We can intrepret the risk-aversion also as ambiguity aversion because students assign

Further, the effect of the option value is much more important at the margin so it becomes important for understanding drop-out rates. These two results imply that any serious model of college education should deviate from risk neutrality assumption and any model with risk averse students should take the option value of dropping out into account. Finally, we show that *ex-ante* returns predicted by our model are smaller suggesting that uncertainty about future income stream can potentially solve the high returns to education puzzle.

1 The Model

We start by describing the time zero problem of high school graduates. The high school graduates differ in their types, in particular, they can be skilled (S) or unskilled (U). A skilled student earns w_S upon graduation from college whereas an unskilled student or a college dropout earns $w_U < w_S$. Although a high school graduate, indexed by j, does not know her type she has an initial belief about her probability of being an unskilled worker which we denote as p_0^j.

Each high school graduate is endowed with initial wealth x_0^j. She chooses her consumption stream, $\{c_t : t \geq 0\}$, and whether to enroll in and dropout from college given her initial belief and wealth in order to maximize her time-separable expected discounted utility from consumption:

$$E_0 \left[\int_0^\infty e^{-\rho t} u\left(c_t^j\right) dt \middle| \mathcal{F}_0^j \right]$$

where $\mathcal{F}_0^j = \left\{p_0^j, x_0^j\right\}$ and $\rho > 0$ is the discount rate. Also, $u_c > 0$ and $u_{cc} < 0$. Risk aversion will be important here due to market incompleteness that arises as a failure to hedge the risk from future income flow. This is an important difference with, for example [Keane-Wolpin (1997)], and [Cunha-Heckman (2007)].

The agents can borrow and lend at a risk-free interest rate $0 < r \leq \rho$. A worker that earns wage w_i accumulates wealth according to

$$\frac{dx_t^j}{dt} = rx_t^j + w_i - c_t^j,\ i \in \{U, S\}$$

A high-school graduate faces the decision of attending college, with costs of a per unit of time, or joining the work-force directly as an unskilled worker. It follows that the wealth dynamics for a college student are given by,

$$\frac{dx_t^j}{dt} = rx_t^j - c_t^j - a$$

subjective beliefs to different outcomes of schooling. Hansen and Sargent (2001) show that the implications of risk-aversion and ambiguity aversion are similar.

1.1 Time 0 problem

For notation simplicity we will drop the agent's index j in the following. Let $V(x; w_i)$ be the value function of an agent working in sector i and wealth x. Also, let $W(x, p)$ be the value for an agent of unknown type but with prior p currently enrolled in college. We assume that the graduation time for the students that do not drop out are exponentially distributed, in order to simplify the solution of the model. In particular, the college students receive a shock that makes them graduate at rate ϕ. Further, we assume that if a student receives a college degree (i.e. gets hit by the ϕ shock) at date t she goes to the market with the belief that she will work in sector U with probability $p(t)$ and the terminal payoff will be the corresponding expected continuation value given $p(t)$.[2]

Figure (1) presents the timing of the agent's problem. Given her initial belief she decides whether to enroll in college or not. If she decides to enroll she becomes a college student, where she updates her belief. At any moment in time she can decide to drop out and join the U-sector. Also, with probability ϕ she graduates and joins her true type' sector.

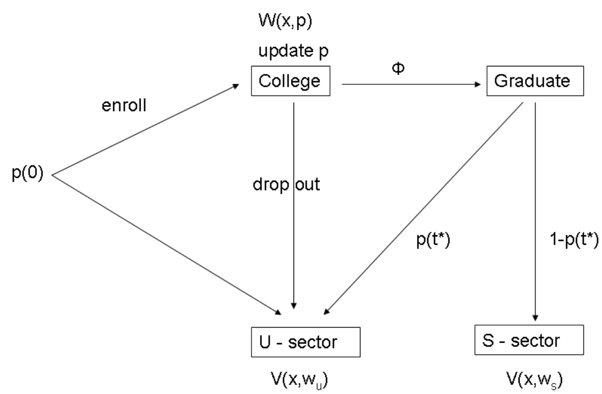

Figure 1: Timing of the model. High-school graduate with prior p_0 decides wether or not to attend college. While in college, she updates her beliefs. Depending on them, she can decide to dropout. Graduation occurs with random probability ϕ and the student's true type is revealed. Then, upon graduation, she joins her true sector.

1.2 Belief' updating

During her college education a student receives good news at rate λ that tells her that she is type S, that is $p^j(t) = 0$ once good news are received. Therefore, a student that does not receive a shock at date t updates her belief using Bayes'

[2] Upon graduation the student takes an exam that fully reveals her true type.

rule. This gives us the evolution of $p(t)$, the belief about the probability of being type U:[3]

$$\begin{aligned}\frac{dp^j(t)}{dt} &= \lambda p^j(t)\left(1-p^j(t)\right) \\ \text{with } p^j(0) &= p_0^j\end{aligned}$$

Note that $\frac{dp^j(t)}{dt} \geq 0$ because not receiving any good news makes the agent more pessimistic.

2 Solving the model with CARA utility

We will use $u(c) = -\frac{e^{-\gamma c}}{\gamma}$, where $\gamma > 0$ denotes the coefficient of absolute risk aversion. Due to this utility function, wealth has no effect on schooling decisions which simplifies considerably all the calculations. If wealth effect would be important a student with $p = 0$ (i.e. knows that she is type S) might decide to drop out depending on her current wealth level. Because this is not the case in our framework the student with $p = 0$ stays in college until graduation. Figure (2) introduces this fact and the learning process discussed before into the model.

2.1 Worker of type i

Let $V(x; w_i)$ be the value function of an agent working in sector i and wealth x. The problem faced by this agent can be written as,

$$\rho V(x; w_i) = \max_c \left(-\frac{e^{-\gamma c}}{\gamma}\right) + V_x(x; w_i)\frac{dx}{dt} \tag{1}$$

where

$$\frac{dx}{dt} = rx + w_i - c \tag{2}$$

Proposition 1 (Value at Work) *The solution to (1) is given by*

$$\begin{aligned}V(x; w_i) &= -\frac{e^{-\gamma\left(rx+w_i+\frac{\rho-r}{r\gamma}\right)}}{\gamma r} \\ c(x; w_i) &= rx + w_i + \frac{\rho - r}{r\gamma}\end{aligned}$$

[3] Using Bayes' rule,

$$p(t+dt) = \frac{p(t)}{p(t) + (1-p(t))(1-\lambda dt)}$$

substracting $p(t)$ from both sides and dividing by dt,

$$\frac{p(t+dt) - p(t)}{dt} = \frac{p(t)(1-p(t))\lambda}{1-(1-p(t))\lambda dt}$$

Now, $\lim_{dt\to 0} \frac{p(t+dt)-p(t)}{dt} = p'(t) = \lambda p(t)(1-p(t))$

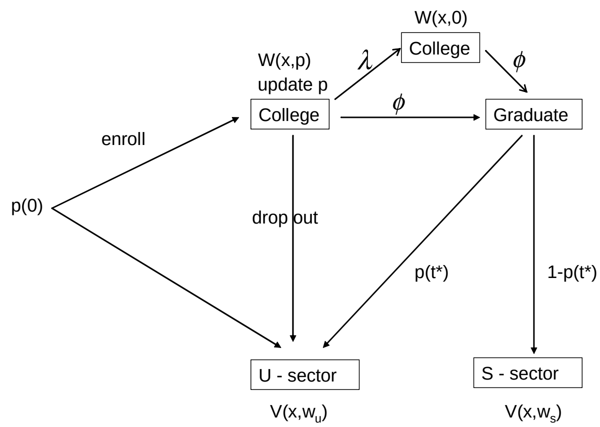

Figure 2: Timing of the model and updating. High-school graduate with prior p_0 decides wether or not to attend college. While in college, she updates her beliefs using Bayes' rule when possible. Bayes' rule is not possible if the agent gets hit by the shock that fully reveals her type (with prob. λ) Depending on the beliefs, she can decide to dropout. If $p = 0$, the student will never decide to dropout because nobody would go to college at the very first stage if a student with $p = 0$ will drop out. Graduation occurs with random probability ϕ and the student's true type is revealed. Then, upon graduation, she joins her true sector.

provided $\rho > r > 0$.

Proof. *See appendix.* ■

Note that, given $w_S > w_U$, $V(x; w_S) > V(x; w_U)$ and so a worker of known type S enjoys a higher level of welfare than a worker of type U. A natural extension of this problem would be letting the w_S and w_U to be randomly drawn from a given distribution in order to analyze the effects of the wage dispersion among college graduates and high school graduates. For a simple example[4], suppose that both of the wages are normally distributed with mean $\bar{w}_i$ and standard deviation σ_i with $\sigma_S > \sigma_U$. Then the results of our model are preserved once we let $w_i \equiv \bar{w}_i - \frac{1}{2}\gamma\sigma_i^2$.

2.2 The problem of a high-school graduate with no dropout option

First we solve the problem faced by a high-school graduate who doesn not have the option of drop out. The fact that students are not allowed to drop out

[4]The normal distribution of wages implies that the wages can be negative with some positive probability. Not withstanding this shortcoming we will proceed with this example due to its simplicity. Moreover, as $\bar{w}_i/\sigma_i$ increases the probability of a negative wage draw becomes arbitrarilly small.

implies that the learning process provides no value to them. In this setup, high school graduates will only account for their initial prior and then decide whether to enroll in college or not.

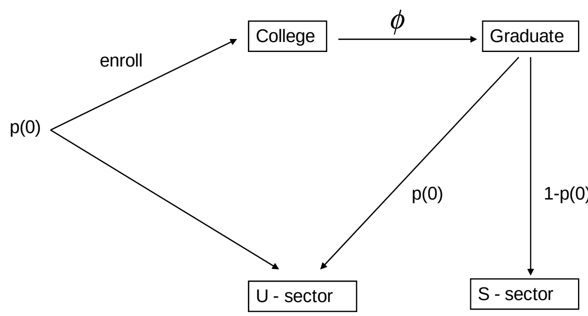

Figure 3: Timing and solution strategy in a model where high-school graduates are not allowed to dropout once enrolled in college.

Let $W^{no\ drop}(x;p)$ denote the current value for a high-school graduate with wealth x and prior p.[5] The problem faced by this agent is,

$$\begin{aligned}\rho W^{no\ drop}(x;p) &= \max_c \frac{e^{-\gamma c}}{-\gamma} + \phi\left[pV(x;w_U) + (1-p)V(x;w_S) - W^{no\ drop}(x;p)\right] \\ &\quad + W_x^{no\ drop}(x;p)(rx - c - a)\end{aligned} \tag{3}$$

The next proposition summarizes the solution to this problem.

Proposition 2 (Value no dropout) *The solution to (3) is given by*

$$\begin{aligned}W^{no\ drop}(x,p) &= \frac{e^{-\gamma(rx+g(p))}}{-\gamma r} \\ c &= rx + g(p)\end{aligned}$$

where $g(p)$ is the solution to

$$(\rho + \phi - r - \gamma r(g(p) + a))e^{-\gamma g(p)} = \phi\left[pe^{-\gamma\left(w_U + \frac{\rho - r}{r\gamma}\right)} + (1-p)e^{-\gamma\left(w_S + \frac{\rho - r}{r\gamma}\right)}\right] \tag{4}$$

Proof. *see appendix.* ■

Corollary 1 $g'(p) < 0$.

Proof. *The left side of the equation (4) is decreasing in $g(p)$ whereas the right side of the same equation is increasing in p. Therefore, $g'(p) < 0$.* ■

[5] In this problem p is not a state variable.

Proposition 3 *If* $(\phi - \gamma r (w_U + a)) e^{\gamma(w_S - w_U)} - \phi > 0$*, there exists a unique threshold* $p_* \in (0,1)$ *such that high-school graduates will attend college if* $p_0 < p_*$ *and will not if* $p \geq p_*$*. Further,*

$$p_* = \frac{(\phi - \gamma r (w_U + a)) e^{\gamma(w_S - w_U)} - \phi}{\phi e^{\gamma(w_S - w_U)} - \phi} \tag{5}$$

Proof. *Assume for a moment that* $g(0)$ *and* $g(1)$ *are such that* $W^{no\ drop}(x;0) > V(x;w_U) > W^{no\ drop}(x;1)$*. Continuity of* $W^{no\ drop}(x;p)$ *and the fact that* $g(p)$ *is strictly decreasing implies that there exists a unique* $p_* \in (0,1)$ *such that* $W^{no\ drop}(x;p_*) = V(x;w_U)$*. This condition is also known as the Value Matching Condition (VM).*

VM implies that it has to be the case that $g(p_*) = w_U + \frac{\rho - r}{r\gamma}$*. Once we applied this condition into (4) we obtained (5). Also, note that* $p_* \in (0,1)$ *is the proposed condition holds. Later, if* $p_* \in (0,1)$ *it has to be the case that* $W^{no\ drop}(x;0) > V(x;w_U) > W^{no\ drop}(x;1)$*.* ■

Lemma 1 $g(p) > 0$*.*

Proof. Although the student does not have the option to drop out she can choose not to enroll in college right after high school. Therefore $W^{no\ drop}(x,p) \geq V(x,w_U)$ and hence $g(p) \geq w_U + \frac{\rho - r}{r\gamma} > 0$. ■

We want to price the option of going to college. Given market incompleteness standard techniques are not available. The approach that we are going to use follow from the Equivalent Variation analysis framework. Let Σ be the maximum amount of wealth that a high-school graduate is willing to pay to have the option of attending college.

2.2.1 Value added of College

The value added of college can be obtained from the following equation:

$$V\left(x; w^U\right) = W^{P,no\ drop}\left(x - \Sigma^{no\ drop}; p\right)$$

where $\Sigma^{no\ drop}$ is the maximum amount (on top of typical college costs) that a high-school student with initial prior p is willing to pay to have the option of attending college. In other words, $\Sigma^{no\ drop}$ is the value added by college.

Solving for $\Sigma^{no\ drop}$,

$$\Sigma^{no\ drop} = \frac{g(p) - w_U + \frac{r - \rho}{r\gamma}}{r}$$

Note that

$$\frac{\partial \Sigma^{no\ drop}}{\partial p} = \frac{g'(p)}{r} < 0$$

So the value added by college is higher the lower is the initial prior.

2.3 The problem of a college student of unknown type

Let $W(x,p)$ denote the value for a college student with prior p. After doing the pertinent approximation from discrete time, we have

$$\begin{aligned}\rho W(x,p) &= \max_c \frac{e^{-\gamma c}}{-\gamma} + \lambda(1-p)\left[W(x,0)-W(x,p)\right] \\ &\quad +\phi\left[pV(x;w_U)+(1-p)V(x;w_S)-W(x,p)\right] \\ &\quad +W_x(x,p)\frac{dx}{dt}+W_p(x,p)\frac{dp}{dt} \\ &\quad \text{where}\begin{cases}\frac{dx}{dt}=rx-c-a\\ \frac{dp}{dt}=\lambda p(1-p)\end{cases}\end{aligned} \tag{6}$$

This equation states that the value flow of a student (the LHS) has to the sum of the following terms: (1) the instant utility flow from consumption, (2) expected jump in value once of receiving the signal that fully reveals your type while in college, (3) expected jump in value from graduation, and (4) the change in value through the change in the state variables.

To solve the previous problem first we need to obtain $W(x,0)$. If we evaluate (6) at $p=0$ we get

$$\begin{aligned}\rho W(x,0) &= \max_c \frac{e^{-\gamma c}}{-\gamma} + \phi\left[V(x;w_S)-W(x,0)\right] \\ &\quad +W_x(x,0)\frac{dx}{dt} \\ &\quad \text{where}\begin{cases}\frac{dx}{dt}=rx-c-a\end{cases}\end{aligned} \tag{7}$$

This equation is telling us that the flow value of being a student with wealth x accounts for the instant utility, the change of value of switching to be a worker (with the correspondent probability of this happening), and for the marginal value of the change in wealth.

Proposition 4 (Value of college with $p=0$) *The solution to (7) is given by*

$$\begin{aligned}W(x,0) &= -B\frac{e^{-\gamma\left(rx+w_S+\frac{\rho-r}{r\gamma}\right)}}{\gamma r} \\ c^S &= rx-a+\frac{\rho-r}{r\gamma}-\frac{\phi}{\gamma r}\left(\frac{1}{B}-1\right)\end{aligned} \tag{8}$$

where B solves

$$h\left(\tilde{B}\right)=0 \tag{9}$$

where $h\left(\tilde{B}\right)\equiv\left(\phi+r\ln\tilde{B}-\gamma r\left(w_S+a\right)\right)\tilde{B}-\phi.$

Proof. *See appendix.* ■

Remark 1 *B exists, it is unique and it is the case that*

$$B \in \left(\max\left(1, e^{\left(\gamma(w_S+a)-\frac{\phi}{r}\right)}\right), +\infty \right)$$

In the appendix we show the details about the properties of B.

Lemma 2 *$W(x,0)$ is stricly increasing and strictly concave.*
Proof.

$$\begin{aligned} \frac{\partial W(x,0)}{\partial x} &= Be^{-\gamma\left(rx+w_S+\frac{\rho-r}{r\gamma}\right)} > 0 \\ \frac{\partial^2 W(x,0)}{\partial x^2} &= -\gamma r Be^{-\gamma\left(rx+w_S+\frac{\rho-r}{r\gamma}\right)} < 0 \end{aligned}$$

provided $B > 0$. ■

Now that we found the solution for $W(x,0)$ we present the solution to the general problem.

Proposition 5 *The solution to (6) is given by*

$$\begin{aligned} W(x,p) &= -\frac{1}{\gamma r} e^{-\gamma(rx+f(p))} \\ c &= rx + f(p) \end{aligned}$$

where $f(p)$ is the solution to the differential equation

$$\begin{aligned} &(\rho - r + \phi + \lambda(1-p) + \gamma f'(p)(\lambda p(1-p)) - \gamma r(f(p)+a))\, e^{-\gamma f(p)} \\ = \;& \\ &\lambda(1-p) Be^{-\gamma\left(w_S+\frac{\rho-r}{\gamma r}\right)} + \phi\left[pe^{-\gamma\left(w_U+\frac{\rho-r}{r\gamma}\right)} + (1-p)e^{-\gamma\left(w_S+\frac{\rho-r}{r\gamma}\right)}\right] \end{aligned} \tag{10}$$

Proof. *See appendix.* ■

Conjecture 1 *There exists p^* such that a student with belief $p \in [p^*, 1]$ drop-out of school. Further, high-school graduates with initial belief $p_0 \in [p^*, 1]$ will not attend college.*

As the problem is casted in continuous time and there is no initial cost of attending college, p^* also dictates college enrollment, i.e. students with initial prior $p_0^j \geq p^*$ will not enroll in college.

Under this conjecture we can rewrite the value function of a student as

$$W(x,p) = \begin{cases} W(x,0) \text{ for } p = 0 \\ W(x,p) \text{ for } p \in (0, p^*) \\ V(x; w_U) \text{ for } p \in [p^*, 1] \end{cases}$$

Note that there is an underlying conjecture in constructing $W(x,p)$. The conjecture is that $W(x,0) \geq W(x,p) \geq V(x; w_U)$ for the corresponding priors.

This conjecture is very intuitive. If a high-school graduates joins college has to be the case that he is better of than being a worker with wage w_U. Also, it has to be the case that a student of known type S has to enjoy higher value provided his expected flow of lifetime earning is higher.

The fact that such a p^* exists implies the existence of an optimal barrier. Further, it must be the case that the marginal student with belief p^* must be indifferent between staying at school and becoming a drop-out,

$$W(x, p^*) = V(x; w_U) \tag{11}$$

This condition is the Value Matching condition. Note that (11) implies that

$$f(p^*) = w_U + \frac{\rho - r}{r\gamma} \tag{12}$$

Also, it must happen that there is no extra value of staying in school for this marginal student,

$$W_p(x, p^*) = 0 \tag{13}$$

This condition is known as Smooth Pasting[6] (SP from now on) and implies that it must be the case that

$$f'(p^*) = 0 \tag{14}$$

The following lemma characterize upper and lower limits for $W(x, p)$.

Lemma 3 *$V(x; w_U) < W(x, p) < V(x; w_S)$ for $p < p^*$.*
Proof. *$V(x; w_U) < W(x, p)$ follows from the fact that an agent can always drop out of school and get value $V(x; w_U)$, and so if he remains a student it must have higher value. Then, if he is studying (i.e. $p < p^*$), $V(x; w_U) < W(x, p)$. If he is studying is because the agent wants to become an S type worker and so $W(x, p) < V(x; w_S)$[7].* ■

Lemma 4 *$w_U + \frac{\rho - r}{r\gamma} < f(p) < w_S + \frac{\rho - r}{r\gamma}$ and $f(p)$ is positive and decreasing for $p \in [0, p^*)$.*
Proof. *From the previous lemma we have that*

$$V(x; w_U) < W(x, p) < V(x; w_S)$$

substituting into this condition the functional forms for $V(x; w_U)$, $V(x; w_S)$ and $W(x, p)$ we obtain

$$w_U + \frac{\rho - r}{r\gamma} < f(p) < w_S + \frac{\rho - r}{r\gamma}, \ 0 \le p < p^*$$

Further, given $\rho \ge r$, $f(p) > 0$.

[6] This can be derived by setting up the problem in discrete time with initial prior $p = p^*$ and then approximate into continuous time.

[7] Provided $a > 0$ it can't be the case that $V(x; w_S) = W(x, p)$ even for an agent that knows he is type S.

To prove that $f'(p) < 0$ *let's conjecture first that* $W(x,0) > W(x,p)$ *because the student who is aware that he is skilled has to be strictly better off (provided that* $\phi > 0$*) than a student who is not sure about it. From this we find that*

$$1 > Be^{-\gamma\left(w_S + \frac{\rho - r}{\gamma r} - f(p)\right)} \tag{15}$$

Moreover, note that $V(x; w_U) < W(x,p) < V(x; w_S)$ *for* $p \in (0, p^*)$ *implies*

$$w_U + \frac{\rho - r}{\gamma r} < f(p) < w_S + \frac{\rho - r}{\gamma r}$$

We can rewrite the differential equation that $f(p)$ *has to satisfy as*

$$\gamma f'(p)(\lambda p(1-p)) = \theta(p, f(p)) \tag{16}$$

where

$$\begin{aligned} \theta(p, f(p)) \quad \equiv \quad & \lambda(1-p)\left[Be^{-\gamma\left(w_S + \frac{\rho - r}{\gamma r} - f(p)\right)} - 1\right] \\ & + \phi\left[pe^{-\gamma\left(w_U + \frac{\rho - r}{r\gamma} - f(p)\right)} + (1-p)e^{-\gamma\left(w_S + \frac{\rho - r}{r\gamma} - f(p)\right)} - 1\right] \\ & - (\rho - r) + \gamma r(f(p) + a) \end{aligned}$$

Note that

$$\begin{aligned} \frac{\partial \theta}{\partial f(p)} \quad = \quad & \lambda(1-p)B\gamma e^{-\gamma\left(w_S + \frac{\rho - r}{\gamma r} - f(p)\right)} \\ & + \phi\gamma pe^{-\gamma\left(w_U + \frac{\rho - r}{r\gamma} - f(p)\right)} + (1-p)e^{-\gamma\left(w_S + \frac{\rho - r}{r\gamma} - f(p)\right)} + \gamma r \end{aligned}$$

Which implies that $\frac{\partial \theta}{\partial f(p)} > 0$*. Further,*

$$\begin{aligned} \frac{\partial \theta}{\partial p} \quad = \quad & \lambda\left[1 - Be^{-\gamma\left(w_S + \frac{\rho - r}{\gamma r} - f(p)\right)}\right] \\ & + \phi\left[e^{-\gamma\left(w_U + \frac{\rho - r}{r\gamma} - f(p)\right)} - e^{-\gamma\left(w_S + \frac{\rho - r}{r\gamma} - f(p)\right)}\right] \end{aligned}$$

Then, $\frac{\partial \theta}{\partial p} > 0$ *because* $w_S > w_U$ *and provided that (15) needs to hold.*

Now suppose that there exist a value of $p = p_1 < p^*$ *such that* $f'(p_1) \geq 0$*. Therefore,* $\theta(p_1, f(p_1)) > 0$*. Then there exist a* $p_2 > p_1$ *in a sufficiently close neighborhood of* p_1 *so that* $f(p_2) \geq f(p_1)$*. Then, we should have* $\theta(p_2, f(p_2)) > \theta(p_1, f(p_1)) > 0$ *because* $\frac{\partial \theta}{\partial f(p)} > 0$ *and* $\frac{\partial \theta}{\partial p} > 0$*. Therefore, from equation (16)* $f'(p_2) > 0$*. If we would repeat this procedure for* p_2 *and further we can easily figure that if* $f(p)$ *is increasing for a value of* $p = p_1 < p^*$ *then it should be increasing for all* $p \in (p_1, p^*)$*. Given* $f(p)$ *is a continuous function this implies that* $f(p^*) > f(p_1) > w_U + \frac{\rho - r}{\gamma r}$ *contradicting the boundary condition that* $f(p^*) = w_U + \frac{\rho - r}{\gamma r}$*. As a result,* $f'(p) < 0$ *for all* $p < p^*$*.* ■

Provided VM and SP we can solve for the optimal threshold p^*. Applying (12) and (14) to (10) and then solving for p^*,

$$p^* = \frac{(\phi + \lambda - \gamma r (w_U + a)) e^{\gamma(w_S - w_U)} - \lambda B - \phi}{(\phi + \lambda) e^{\gamma(w_S - w_U)} - \lambda B - \phi}$$

Further, $p^* \in (0, 1)$ provided

$$(\phi + \lambda - \gamma r (w_U + a)) e^{\gamma(w_S - w_U)} - \lambda B - \phi > 0 \tag{17}$$

or,

$$\frac{(\phi + \lambda - \gamma r (w_U + a)) e^{\gamma(w_S - w_U)} - \phi}{\lambda} > B \tag{18}$$

This two conditions will prove to be useful in the comparative statics calculations below.

2.3.1 Comparative Statics

An advantage of our modeling decisions is that the solution to the problem is almost a close-form solution except the value B which depends endogenously on model parameters. The advantage of this is that allows for easy-to-get comparative statics. Table (1) presents the results and the details are in the appendix.

An increase in w_S increases the expected terminal payoff of the project and thus increases the initial value of the option. Further, this implies an increase in the likelihood of execution of the option. An increase in the schooling cost a decreases the initial value of the project and thus decreases p^*. An increase in w_U has two different effects. First, as w_U increases the terminal payoff increases (increasing p^*) because an increase in w_U provides insurance. Second, an increase in w_U increases the opportunity cost of attending college (a college student can always work in the U-sector with wage w_U), generating downward pressure over p^*. In our model the later effect is stronger and thus an increase in w_U decreases p^*. An increase in the graduation parameter ϕ implies that the expected time until graduation decreases, as ϕ is exponentially distributed with mean $1/\phi$. Then, in expectations, the terminal payoff is closer the higher is ϕ. This implies that an increase in ϕ should increase the total value of the project and thus p^*.[8]

An increase in the learning parameter λ implies that, conditional on not receiving any good news, a student gets pessimistic faster. This would suggest that p^* should be lower because not receiving good news is more informative in a negative way. However, higher λ also means that skilled students receive the jump to $p(t) = 0$ and to the higher continuation value $W(x, 0)$ faster. In our model, this latter effect dominates the former one indicating that p^* should increase[9].

[8] It is important to remember that upon graduation all uncertainty is resolved.

[9] This result is very intuitive for one of the potential extensions of the model where we interpret λ to be a proxy for cognitive abilities that differs among the students that is different,

Parameter	Enrollment	Dropout	p^*
$\uparrow w_S$	$\uparrow$	$\downarrow$	$\uparrow$
$\uparrow w_U$	$\downarrow$	$\uparrow$	$\downarrow$
$\uparrow a$	$\downarrow$	$\uparrow$	$\downarrow$
$\uparrow \phi$	$\uparrow$	$\downarrow$	$\uparrow$
$\uparrow \lambda$	$\uparrow$	$\downarrow$	$\uparrow$
$\uparrow r$	$\downarrow$	$\uparrow$	$\downarrow$
$\uparrow \rho$	$=$	$=$	$=$
$\uparrow \gamma$	?	?	?

Table 1: Comparative Statics. Relation between primitives of the model and college enrollment and dropout

To understand why an increase in the interest rate r decreases p^* we can think of schooling as a project that causes utility loss today and utility gain tomorrow. If interest rate increases both the discounted value of utility loss and utility gain will decrease. However, the decrease in the discounted value of utility gain will be higher since it occurs later in the future. (To see this more clearly think of a project that requires investment $-A$ at the beginning for T periods and then will provide returns B after that and have positive discounted present value. The discounted value of this project is $-A/r + (B + A)(1 + r)^{-T}/r > 0$. Taking the derivative shows that this value is decreasing in r.) As a result an increase in r decreases the option value of schooling. This leads to a decrease in p^*.

As shown in Table (1) there is no effect over p^* when we change ρ. This is the case because ρ affect only the intertemporal allocation.

Finally, the effect of γ over p^* is ambiguous. The reason is that increased risk aversion decreases the utility you get from the outcome of schooling once you get the graduation shock, given that the student's type is not revealed by a λ-shock. On the other hand, it also increases the option value of schooling and hence propensity to stay in school. These two effects are acting against eachother.

2.3.2 Value added of college with drop out option

We want to price the option of going to college. As before,

$$V\left(x; w^U\right) = W\left(x - \Sigma, p\right)$$

Using the functional forms we obtain for $V\left(x; w_U\right)$ and $W\left(x, p\right)$, and solving for Σ,

but not necessarily independent, from being of type S or U. In particular, we can assume that students with better cognitive abilities learn faster about their type. Therefore, $\partial p^*/\partial\lambda > 0$ implies that students with better cognitive abilities are more likely to enroll in and graduate from college.

$$\Sigma(p) = \frac{f(p) - w_U - \frac{\rho - r}{r\gamma}}{r}$$

Note that $\Sigma(p^*) = 0$ because college has no value for an agent with prior p^*. Also,

$$\frac{\partial \Sigma}{\partial p} < 0$$

2.4 Comparison of p^* and p_*

First we rewrite p^* as

$$p^* = \frac{(\phi - \gamma r(w_U + a)) e^{\gamma(w_S - w_U)} - \phi + \lambda\left(e^{\gamma(w_S - w_U)} - B\right)}{\phi e^{\gamma(w_S - w_U)} - \phi + \lambda\left(e^{\gamma(w_S - w_U)} - B\right)} \tag{19}$$

In order to compare p^* and p_* first we need to sign $e^{\gamma(w_S - w_U)} - B$. Recall that B solves $h(B) = 0$. Also, because h is increasing for $B > 1$, $e^{\gamma(w_s - w_u)} > B$ iff $h\left(e^{\gamma(w_s - w_u)}\right) > 0$ or, by using the definition of $h(\cdot)$,

$$\phi\left(e^{\gamma(w_s - w_u)} - 1\right) - r\gamma(w_U + w_a) e^{\gamma(w_s - w_u)} > 0$$

given that $e^{\gamma(w_s - w_u)} > 1$,

$$\phi\left(e^{\gamma(w_s - w_u)} - 1\right) - r\gamma(w_U + w_a) > 0$$

Also, from (17),

$$[\phi - \gamma r(w_u + a)] e^{\gamma(w_s - w_u)} - \phi > \lambda\left(B - e^{\gamma(w_s - w_u)}\right) \tag{20}$$

Now suppose $e^{\gamma(w_s - w_u)} \leq B$.Then, we should have

$$\begin{aligned} \phi\left(e^{\gamma(w_s - w_u)} - 1\right) - \gamma r(w_u + a) &\leq 0 \text{ and} \\ \lambda\left(B - e^{\gamma(w_s - w_u)}\right) &\geq 0 \end{aligned}$$

which contradicts (20) above. Then, $B < e^{\gamma(w_S - w_U)}$.

Now we proceed to compare the thresholds. Define $J \equiv \lambda\left(e^{\gamma(w_S - w_U)} - B\right)$ and rewrite p^* as follows:

$$\begin{aligned} p^* &= \frac{(\phi - \gamma r(w_U + a)) e^{\gamma(w_S - w_U)} - \phi + J}{\phi e^{\gamma(w_S - w_U)} - \phi + J} \\ &= \frac{(\phi - \gamma r(w_U + a)) e^{\gamma(w_S - w_U)} - \phi}{\phi e^{\gamma(w_S - w_U)} - \phi + J} + \frac{J}{\phi e^{\gamma(w_S - w_U)} - \phi + J} \\ &= \left(\frac{\phi e^{\gamma(w_S - w_U)} - \phi}{\phi e^{\gamma(w_S - w_U)} - \phi + J}\right) p_* + \frac{J}{\phi e^{\gamma(w_S - w_U)} - \phi + J} \\ &= \frac{\left(\phi e^{\gamma(w_S - w_U)} - \phi\right) p_* + J}{\phi e^{\gamma(w_S - w_U)} - \phi + J} \end{aligned}$$

Given $\left(\phi e^{\gamma(w_S-w_U)}-\phi\right)p_* < \phi e^{\gamma(w_S-w_U)}-\phi$ and $J>0$,

$$p^* > p_*$$

As a result we conclude that the enrollment and graduation rates should be lower in the absense of the option to drop out of college.

3 Solving the model using the Net Present Value (NPV) approach

The risk-neutral NPV approach is widely used in studies of education choice. This approach simply compares income/cost flows, without taking into account income or attitude towards risk. First we will consider the case where a college student is not allowed to dropout. Then, we will extend the model to allow for this possibility.

3.1 NPV without option of drop out

Define $V^{NPV}(w_i)$ and $W^{NPV}(p)$ as the net present value of being employed with constant wage w_i and net present value of being a student with initial prior $p=p_0$.

Straightforward calculations provide $V^{NPV}(w_i)=\int_0^\infty e^{-rt}w_i dt=\frac{w_i}{r}$. The problem faced by a high-school graduate is

$$rW^{NPV}(p)=-a+\phi\left(pV^{NPV}(w_U)+(1-p)V^{NPV}(w_S)-W^{NPV}(p)\right)$$

After substituting for $V^{NPV}(w_U)$ and $V^{NPV}(w_S)$ we can solve for $W^{NPV}(p)$:

$$W^{NPV}(p)=\frac{-a+\phi\left(p\frac{w_U}{r}+(1-p)\frac{w_S}{r}\right)}{(r+\phi)} \tag{21}$$

Proposition 6 *Provided $\phi(w_S-w_U)>r(a+w_U)$, there exists a unique threshold $\underline{p}=\frac{\phi(w_S-w_U)-r(w_U+a)}{\phi(w_S-w_U)}\in(0,1)$ such that only high-school graduates with initial prior $p_0<\underline{p}$ will attend college.*

Proof. *First note that $W^{NPV}(1)<V^{NPV}(w_U)$ and that $W^{NPV}(0)>V^{NPV}(w_U)$ if $\phi(w_S-w_U)>r(a+w_U)$. Continuity of $W^{NPV}(p)$ implies that there exists at least one $\underline{p}$ such that $W^{NPV}(\underline{p})=V^{NPV}(w_U)$. This is the VM condition. Further, $\frac{\partial W^{NPV}(p)}{\partial p}=-\frac{\phi(w_S-w_U)}{r(r+\phi)}<0$ provides uniqueness. Using the VM condition in (21) we can solve for $\underline{p}$. This also guarantees $\underline{p}\in(0,1)$.* ■

3.2 NPV with option of drop out

Define $W^{NPV,P}(p)$ as the net present value of being a student of unknown type. The NPV of an education can be constructed as follows:

$$\begin{aligned}(r+\phi+\lambda(1-p))\,W^{NPV,P}(p) &= -a+\lambda(1-p)\,W^{NPV} \\ &\quad +\phi\left[pV^{NPV}(w_U)+(1-p)\,V^{NPV}(w_S)\right] \\ &\quad +W_p^{NPV,P}(p)\,\lambda p(1-p)\end{aligned} \tag{22}$$

Conjecture 2 *There exists a threshold $\bar{p}$ such that a student with prior $p \geq \bar{p}$ drops-out of college.*

VM and SP conditions:

$$\begin{aligned} W^{NPV,P}(\bar{p}) &= \frac{w_U}{r} \\ W_p^{NPV,P}(\bar{p}) &= 0 \end{aligned}$$

Proposition 7 (NPV value function) *The solution to (22) is given by*

$$W^{NPV,P}(p)=\frac{\frac{\phi}{r}w_S-a}{\phi+r}+\left[\frac{(1-\bar{p})\,\lambda}{(1-\bar{p})\,\lambda+r+\phi}\left(\frac{p}{\bar{p}}\frac{1-\bar{p}}{1-p}\right)^{\frac{r+\phi}{\lambda}}-1\right]\frac{p\frac{\phi}{r}(w_S-w_U)}{\phi+r}$$

where

$$\bar{p}=\frac{\left(1+\frac{\lambda}{r+\phi}\right)\left[\phi(w_S-w_U)-r(w_U+a)\right]}{\left(1+\frac{\lambda}{r+\phi}\right)\left[\phi(w_S-w_U)-r(w_U+a)\right]+r(w_U+a)}$$

Proof. *see appendix.* ■

In order to have $\bar{p} \in (0,1)$ it must be the case that

$$r(w_U+a)<\phi(w_S-w_U) \tag{23}$$

which is the same condition we derived for the case with no dropout option.

3.3 Comparison of $\bar{p}$ and $\underline{p}$

First rewritte $\bar{p}$ as follows,

$$\bar{p}=\frac{\phi(w_S-w_U)-r(w_U+a)}{\phi(w_S-w_U)-\left(\frac{\lambda}{r+\phi+\lambda}\right)r(w_U+a)}$$

Then it is clear that $\bar{p}>\underline{p}$ because $\left(\frac{\lambda}{r+\phi+\lambda}\right)r(w_U+a)>0$. This implies that the likelihood of exercising the option of attending college is higher if students are allowed to dropout. This makes sense because the dropout option provides extra value for students.

4 Construction of initial belief p_0

We already solved the problem faced by a high-school graduate with initial prior p_0. We still need to link these initial beliefs with the agent's true type.

We normalize the total size of the high-school graduates as a continuum over the unit interval. Also, let $s_U \in (0, 1)$ represent the proportion of type U high-school graduates. Also, let q_0 denote the objective prior for each high-school graduate. The difference between q_0 and p_0 is significant. q_0^j is the belief of being type U only taking into account observable characteristics of the agent, as are the GPA average in high-school, parents' income, and so on. This information is also generally available to the econometrician. On the other hand, p_0^j is the subjective belief and is just the updating over q_0^j using the unobserved characteristics of the agent or his individual expectations.

We assume that the initial prior of each agent, denoted by q_0, has the following distribution:

$$h_U(q_0) \equiv h(q_0|U) = 2q_0,\ q_0 \in [0, 1]$$

and

$$h_S(q_0) \equiv h(q_0|S) = 2 - 2q_0,\ q_0 \in [0, 1]$$

The choice of this particular $h(\cdot)$ was purely due to the simplicity of these distributions. Also, it is the case that $H_S(q_0) \underset{FOSD}{\succ} H_U(q_0)$ so that agents with higher q_0 are more likely to have higher p_0.

Even though $h_S(\cdot)$ and $h_U(\cdot)$ differ, both share the same support and thus every agent uses Bayes' rule to update her beliefs.

As a side note we can think about this model as a game with two players and 4 nodes (as shown in Figure (4)). The players are nature and the high-school graduate. The nodes are: (1) Nature playes and choses the type of the high-school graduate with the mixed strategy $\{(U, s_U), (S, 1 - s_U)\}$[10], (2) Nature plays and, conditional on the result of the first stage picks a q_0 for the high-school graduate using the mixed strategy $h_U(q_0)$ if the high-school graduate is type U and uses $h_S(q_0)$ if type S, (3) the high-school graduate decides whether to go to college or not, and (4) conditional on attending college, the student decides between continuing education or dropping-out.

Once in stage 3 (where the high-school graduate starts participating), and with q_0 in hand, the high-school graduate uses Bayes' rule to infer the probability of being type U:

$$\Pr(U|q_0) = \frac{\Pr(q_0|U)\Pr(U)}{\Pr(q_0)}$$

Using that $\Pr(U) = s_U$, $\Pr(q_0|U) = h_U(q_0)$, and that

$$\begin{aligned}\Pr(q_0) &= h_U(q_0)s_U + h_S(q_0)(1 - s_U) \\ &= 2q_0 s_U + (2 - 2q_0)(1 - s_U)\end{aligned}$$

[10] Still, Nature observes the result of this bet.

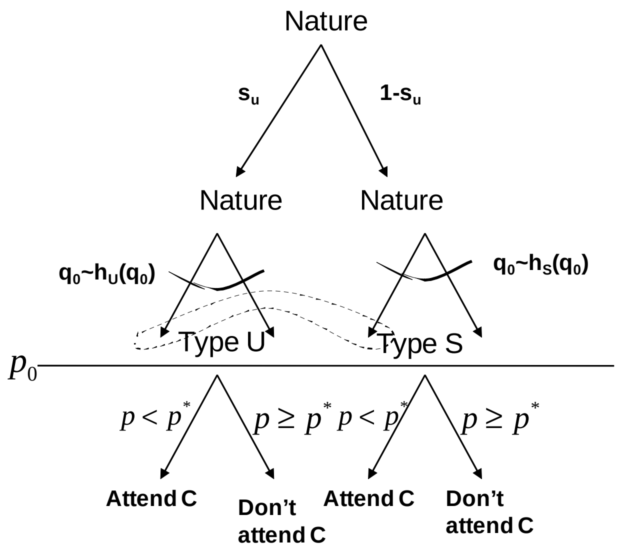

Figure 4: Structure of the game. Nature plays first and assings a type and an initial belief q_0. The agent updates her beliefs using Bayes' rule to compute his subjective initial belief p_0 and decides wether to attend college or not by comparing it with the treshold p^*.

we get that

$$p(q_0) \equiv \Pr(U|q_0) = \frac{q_0 s_U}{q_0 s_U + (1-q_0)(1-s_U)} \tag{24}$$

5 Computing College Enrollment and Dropout rates

We will attempt to calibrate this model to data. The two moments that we are interested on are College Enrollment and Dropout rates.

As a side step we can invert (24) to find $q(p_0)$,

$$q(p_0) = \frac{(1-s_U)p_0}{(1-s_U)p_0 + s_U(1-p_0)} \tag{25}$$

Let C^E denote college enrollment rate and D^O denote dropout rate. Note that

$$C^E = \int_0^{q(p^*)} (h_S(q_0)(1-s_U) + h_U(q_0)s_U)\, dq_0$$

where $q(p^*)$ is such that agents with $q \geq q(p^*)$ will not attend college.

Solving the definite integral,

$$C^E = 2(1-s_U)\, q(p^*) + (2s_U - 1)\, q(p^*)^2$$

Evaluating (25) at $p_0 = p^*$,

$$C^E = 2(1-s_U)\, q^* + (2s_U - 1)(q^*)^2 \tag{26}$$

In the same fashion,

$$D^O = \sum_{i\in\{U,S\}} s_i \int_0^{q^*} Q_i(p(q_0)) \frac{h_i(q_0)}{H_i(q^*)} dq_0$$

where $Q_i(p)$ is the probability of a college student of type i hitting p^* and thus dropping out. Further, $H_i(q^*)$, the CDF of $h_i(q)$, accounts for the fact that Q_i is conditional on attending college. In our setup,

$$Q_S(p_0) = \left(\frac{1-p^*}{1-p_0}\frac{p_0}{p^*}\right)^{\frac{\phi+\lambda}{\lambda}}$$

and

$$Q_U(p_0) = \left(\frac{1-p^*}{1-p_0}\frac{p_0}{p^*}\right)^{\frac{\phi}{\lambda}}$$

(the details can be found in the appendix).
Using that

$$\frac{p(q_0)}{1-p(q_0)} = \frac{q_0 s_U}{(1-q_0)(1-s_U)} = \frac{s_U}{1-s_U}\frac{q_0}{1-q_0} \tag{27}$$

we get that

$$\begin{aligned} D^O &= \left[\frac{1-p^*}{p^*}\right]^{\frac{\phi}{\lambda}} \frac{s_U}{(q^*)^2} \int_0^{q^*} \left[\frac{p(q_0)}{1-p(q_0)}\right]^{\frac{\phi}{\lambda}} 2q_0 dq_0 \\ &\quad + \left[\frac{1-p^*}{p^*}\right]^{1+\frac{\phi}{\lambda}} \frac{1-s_U}{2q^*-(q^*)^2} \int_0^{q^*} \left[\frac{p(q_0)}{1-p(q_0)}\right]^{1+\frac{\phi}{\lambda}} 2(1-q_0)\, dq_0 \end{aligned} \tag{28}$$

Note that, using (27), we can write $\int_0^{q^*}\left[\frac{p(q_0)}{1-p(q_0)}\right]^{\alpha} q_0 dq_0$ as

$$\int_0^{q^*} \left[\frac{p(q_0)}{1-p(q_0)}\right]^{\alpha} q_0 dq_0 = \left(\frac{s_U}{1-s_U}\right)^{\alpha} \int_0^{q^*} q_0^{\alpha+1}(1-q_0)^{-\alpha} dq_0$$

Where

$$\int_0^{q^*} q_0^{\alpha+1}(1-q_0)^{-\alpha} dq_0 = Beta\left[q^*; 2+\alpha; \alpha\right]$$

where $Beta\left[q^*; 2+\alpha; \alpha\right]$ is an incomplete Beta density that is already available in many computational packages.

6 Calibration

We discuss how we calibrated the model with CARA utility function and dropout option.[11]

We set $\rho = 0.03$ so that $\beta = e^{-\rho} = 0.97$. Further, we set $r = \rho$ in order to abstract from intetemporal issues regarding differences between ρ and r.

Using data from the National Center for Education Studies (see Appendix), or NCES, we can obtain estimates for the other parameters. Before doing so, we have to think about what to look for in the data. Agents with Associate degrees and Vocational Certificates are in esence, even though working in the white-collar sector, unskilled agents. This implies that we have to adjust the data for this fact (this is done in the Appendix).

The estimates for w_U and w_S and a (net cost of attending college) are the following:

$w_U = 28369$
$w_S = 41100$
$a = 9934$

To calibrate γ we use estimates by Navarro (2005). Navarro estimates the coefficient of relative risk aversion, σ, using data from NLSY. He estimates σ to be 2.15. The relative risk aversion with CARA utility function used here is

$$-\frac{u''(c)\,c}{u'(c)} = \gamma c$$

Then it follows that it should be the case

$$\gamma c = \sigma$$

or,

$$\gamma = \frac{\sigma}{c}$$

Abstracting from wealth[12], and provided $\rho = r$, consumption by a student should be a convex combination between w_U and w_S. As a first order approximation, we assume that

$$\gamma \approx \frac{\sigma}{\frac{1}{2}(w_S + w_U)} = 6.19$$

To calibrate ϕ we look at the average expected time in college until completion. In data, an average student requires 4.58 years to leave college with a degree. Provided that the arrival rate of earning a degree ϕ is exponentially distributed, the average expected time until completion is $1/\phi$ so

$$\phi = 0.21834$$

[11] the method used to calibrate the NPV approach model is similar.

[12] You can always think about wealth being close to 0 in our model because of the inability of CARA models of generating wealth effects.

For obtaining estimates for λ and s_U we have to rely in numerical methods. For that, first we need to construct a set of moments in order to identify these parameters. The moments that we are going to use are college enrollment and drop out rates. From the data appendix we obtain that $C^E = 0.435$ and $D^O = 0.3253$. Note that (26) and (28) can be represented as

$$\begin{aligned} C^E &= C^E(\lambda, s_U) \\ D^O &= D^O(\lambda, s_U) \end{aligned}$$

Then the two moments are

$$\begin{cases} 0.435 = C^E(\lambda, s_U) \\ 0.3253 = D^O(\lambda, s_U) \end{cases}$$

Using these tow moments we find

$$\lambda = 0.38$$

and

$$s_U = 0.67$$

7 Results

With the calibration complete we proceed to compare the models and discuss the results. Figure (5) presents the value added by college for the CARA and NPV approach with the dropout option. As we discussed before, the value is decreasing in the initial prior until the enrollment threshold is reached. Further, it can be observed in the figure that the threshold in the CARA case is lower than in the NPV approach. Third, and most important, the value added by college in the NPV approach lies above the CARA approach. Several studies (for example [Cunha-Heckman (2007)] or [Judd (2000)]) have pointed out that returns to college education are very high relative to other investment opportunities and thus generates the "Returns to Education puzzle" that we already discussed. Our results are important in the sense that show that if the true model is the CARA approach, then the expected returns should be lower and thus mitigates the size of the puzzle.[13]

In order to compute the value added by the dropout option we computed the value added by college for an agent where the dropout option was not available. As seen in Figure (6) the threshold is higher when the dropout option is present. Further, the value added in the case with the dropout option is always higher.[14]

Finally, we produce the total value added by the dropout option for both CARA and NPV approach. The results can be seen in Figure (7). There are three important things to point out here.

[13] C-H 2007 discuss models without uncertainty in the sense we discussed here that would further increase the returns and so increase the size of the puzzle.

[14] the same applies for the NPV approach.

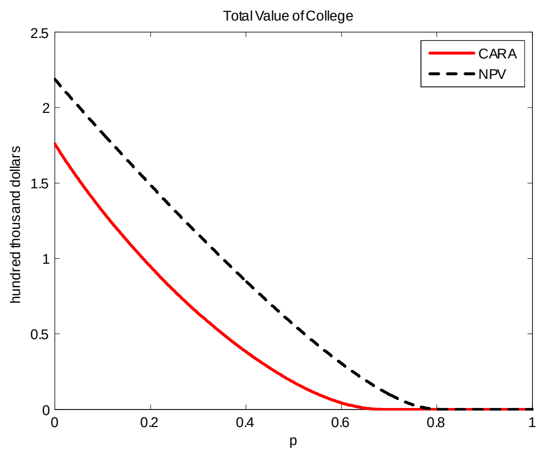

Figure 5: Value added by college for CARA and NPV with dropout option. Note: for each case a different moment calibration was performed.

First, the option value has an inverted v shape, reaching a maximum somewhere between 0 and the respective threshold. This is due to the fact that the option has to have higher value the higher the uncertainty. Agents with priors close to 0 are pretty certain about their true type. Also, agents with priors close to the threshold, even though they might be somewhat uncertain about their true type, are certain enough that they are of type U and so will probably not attend college (as it happens for agents with priors over the threshold). However, agents with intermediate values of p face higher uncertainty and therefore value the drop out option more.

Second, the option value exhibits some skewness to the left because students with higher priors are more certain of being type U and so have more to loose from college education. As a result, the option value decreases sharply.

Third we have that it is not the case that the option value is always higher/lower in the CARA approach relative to the NPV approach because the threshold is higher for the NPV case. This suggests that there is a non-linear relation between risk-aversion and the option value.

8 Extensions and Future Work

Our model provides a traceable framework which can be extended to capture the properties of the data better by relaxing some assumptions.

First of all, because the model assumes constant wages in both skilled and unskilled sectors and because the threshold value of beliefs do not depend on wealth, dropouts would not consider reapplying for college at a later date. One way to allow for return to college is having stochastic wage processes in skilled

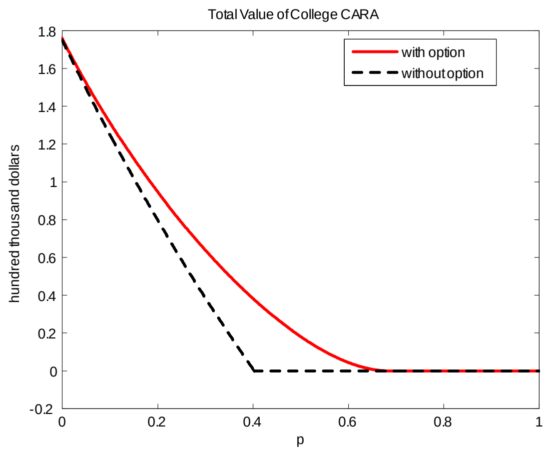

Figure 6: Value added by college in the CARA approach with and without dropout option.

and unskilled sectors so that students choose to go back to college if the gap between skilled and unskilled wages widens. This type of phenomenon has been observed in the 80s.

Second, the constant absolute risk aversion assumption leads to the simple property of the model that the enrollment and dropout decisions do not depend on the amount of wealth. However, we observe that the richer students are more likely to go to college and graduate. We can use constant relative risk aversion utility function in order to capture the difference between rich and poor students because the threshold level of beliefs will depend on wealth. But this model is much less straightforward that the CARA model. First, we do not have an explicit solution. Second, the effect of wealth on the enrollment and dropout decision cannot be qualitatively determined at the beginning. In the CRRA case richer students will act as if they are less risk averse in absolute terms whereas the effect of absolute risk aversion on the belief threshold can increase or decrease with absolute risk aversion. Finally, because p^* does not depend on wealth in CARA case the students would not choose to re-enter the college later even if they had the option to do so. However, in the CRRA case the student can choose to increase its wealth while not in college. This may in turn decrease his threshold value for p^* if $\partial p^*/\partial x > 0$ and make him return to the college.

Third, we can make the rate of arrival of good news and graduation shocks idiosyncratic and endogenous. For this purpose, we should separate working skills and cognitive abilities. Students with better cognitive skills learn their type faster and hence receive the good news at a higher rate, conditional that they are skilled workers. We should also consider the possibility that the working

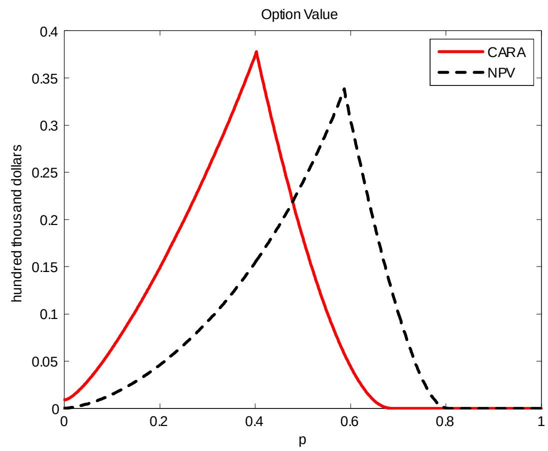

Figure 7: Value added by the option value for both CARA and NPV approach.

skills and cognitive abilities are correlated. Moreover, we can think of the rate of graduation shocks as increasing with time or let it be different before and after the receipt of good news. This way, students that receive good news stay less in college instead of lingering around campus for a long time. We can endogenize the arrival rate of these shocks by letting the students adjust them at some monetary or utility expense.

Fourth, we can provide some empirical content for the initial prior p_0. Imagine a model where

$$p_0^i = \Theta\left(X_i\right) + \varepsilon_i$$

where X_i is a vector of observables characterizing high-school graduate i and ε_i is the unobservable part (both to the econometrician and the high-school graduate) such that $X_i \perp \varepsilon_i$, $E\left(\varepsilon_i\right) = 0$ so ε_i is unforcastable. X_i can include the GPA average, income of parents, education of parents, and so on. We can run a regression on observed wages (that we thing is a good predictor of ability) on this characteristics in order to compute $\hat{p}_0^i$. Then, we can try to figure out the distribution of ε_i that better fits the data.

Finally, we can exploit the dynamic nature of our model to fit the trajectories of different variables over time as a robustness check.

9 Conclusion

In this paper, we focus on college enrollment and dropout decisions in a traceable options model where attitude towards schooling stems from the initial belief of students about their skill level, providing a less controversial explanation for the once obscure psychic costs. In this model the two main benefits of college education are credential effects and learning.

Our paper has four main results. First, our model shows that papers that assume risk neutrality, such as [Keane-Wolpin (1997)], [Cunha-Heckman (2007)] and [Heckman-Navarro (2006)], overestimate the value of college, college enrollment and graduation rates because they omit the uncertainty regarding the outcome of college education. Second, we show that the option value of learning is much more important when agents are risk averse rather than risk neutral. Further, the effect of the option value is much more important at the margin so it becomes important for understanding drop-out rates. These two results imply that any serious model of college education should deviate from risk neutrality assumption and any model with risk averse students should take the option value of dropping out into account. Finally, we show that *ex-ante* returns predicted by our model are smaller suggesting that uncertainty about future income stream can potentially solve the high returns to education puzzle.

A Proof of Proposition (Value at Work)

Substitute the wealth evolution (2) into (1) to get

$$\rho V(x;w_i) = \max_c \left(-\frac{e^{-\gamma c}}{\gamma}\right) + V_x(x;w_i)(rx + w_i - c) \tag{29}$$

The First Order Condition (FOC) states that

$$e^{-\gamma c} = V_x(x;w_i)$$

Further, it implies:

$$\begin{aligned} -\frac{e^{-\gamma c}}{\gamma} &= -\frac{V_x(x;w_i)}{\gamma} \\ c &= -\frac{\ln V_x(x;w_i)}{\gamma} \end{aligned}$$

Plugging these two expressions into (29),

$$\rho V(x;w_i) = -\frac{V_x(x;w_i)}{\gamma} + V_x(x;w_i)\left(rx + w_i + \frac{\ln V_x(x;w_i)}{\gamma}\right)$$

Guess $V(x;w_i) = -\frac{e^{-\gamma(rx+w_i+A)}}{\gamma r}$ and evaluate the last equation,

$$-\rho\frac{e^{-\gamma(rx+w_i+A)}}{\gamma r} = -\frac{e^{-\gamma(rx+w_i+A)}}{\gamma} + e^{-\gamma(rx+w_i+A)}\left(rx + w_i - (rx + w_i + A)\right)$$

Solving for A,

$$A = \frac{\rho - r}{\gamma r}$$

Thus,

$$V(x; w_i) = -\frac{e^{-\gamma\left(rx + w_i - \frac{r-\rho}{\gamma r}\right)}}{\gamma r}$$

Using the FOC,

$$c(x; w_i) = rx + w_i - \frac{r-\rho}{\gamma r}$$

That $\rho > r$ is needed follows from the fact that otherwise the transversality condition will not hold (the agent can save every period his whole income and consume ∞ at $t = \infty$).

The transversality condition is

$$\lim_{t\to\infty} e^{-\rho t} |V(x; w_i)| = \lim_{t\to\infty} e^{-\rho t} |\frac{e^{-\gamma\left(rx + w_i - \frac{r-\rho}{\gamma r}\right)}}{\gamma r}| = 0$$

B Proof of proposition (Value no dropout)

(3) can be rewritten as

$$\begin{aligned}(\rho + \phi) W^{no\ drop}(x; p) &= \max_c \frac{e^{-\gamma c}}{-\gamma} - \phi \left[\begin{array}{c} p \frac{e^{-\gamma\left(rx + w_U - \frac{r-\rho}{r\gamma}\right)}}{\gamma r} \\ + (1-p) \frac{e^{-\gamma\left(rx + w_S - \frac{r-\rho}{r\gamma}\right)}}{\gamma r} \end{array} \right] \\ &\quad + W_x^{no\ drop}(x; p)(rx - c - a)\end{aligned}$$

The First Order Condition (FOC) states that

$$e^{-\gamma c} = W_x^{no\ drop}(x; p)$$

Further, it implies:

$$\begin{aligned} -\frac{e^{-\gamma c}}{\gamma} &= -\frac{W_x^{no\ drop}(x; p)}{\gamma} \\ c &= -\frac{1}{\gamma} \ln W_x^{no\ drop}(x; p) \end{aligned}$$

Substituting into the previous equation,

$$\begin{aligned}(\rho + \phi) W^{no\ drop}(x; p) &= -\frac{W_x^{no\ drop}(x; p)}{\gamma} - \phi \left[\begin{array}{c} p \frac{e^{-\gamma\left(rx + w_U - \frac{r-\rho}{r\gamma}\right)}}{\gamma r} \\ + (1-p) \frac{e^{-\gamma\left(rx + w_S - \frac{r-\rho}{r\gamma}\right)}}{\gamma r} \end{array} \right] \\ &\quad + W_x^{no\ drop}(x; p) \left(rx + \frac{1}{\gamma} \ln W_x^{no\ drop}(x; p) - a \right)\end{aligned}$$

Conjecture that $W^{no\ drop}(x;p) = \frac{e^{-\gamma(rx+g(p))}}{-\gamma r}$. Then,

$$(\rho + \phi - r - \gamma r (g(p) + a)) e^{-\gamma g(p)} = \phi \left[p e^{-\gamma \left(w_U - \frac{r-\rho}{r\gamma} \right)} + (1-p) e^{-\gamma \left(w_S - \frac{r-\rho}{r\gamma} \right)} \right]$$

The transversality condition is $\lim_{t\to\infty} e^{-\rho t} |W^{no\ drop}(x;p)| = 0$. This condition holds because for $p < p_*$,

$$V(x; w_U) < W^{no\ drop}(x;p) < V(x; w_S)$$

which by continuity implies $0 = \lim_{t\to\infty} e^{-\rho t} V(x; w_U) < \lim_{t\to\infty} e^{-\rho t} W^{no\ drop}(x;p) < \lim_{t\to\infty} e^{-\rho t} V(x; w_S) = 0$.

C Proof of proposition (Value of college with $p = 0$)

(7) can be written as

$$\begin{aligned}
(\rho + \phi) W(x,0) &= \max_c \left(-\frac{e^{-\gamma c}}{\gamma} \right) + \phi V^S(x; w_S) + W_x(x,0)(rx - c - a) \\
(\rho + \phi) W(x,0) &= \max_c \left(-\frac{e^{-\gamma c}}{\gamma} \right) - \frac{\phi}{r\gamma} e^{-\gamma \left(rx + w_S - \frac{r-\rho}{r\gamma} \right)} + W_x(x,0)(rx - c - a)
\end{aligned}$$

The FOC wrt c implies,

$$\begin{aligned}
-\frac{e^{-\gamma c}}{\gamma} &= -\frac{W_x(x,0)}{\gamma} \\
c &= -\frac{1}{\gamma} \ln W_x(x,0)
\end{aligned}$$

Substituting back,

$$(\rho + \phi) W(x,0) = -\frac{W_x(x,0)}{\gamma} - \frac{\phi}{r\gamma} e^{-\gamma \left(rx + w_S - \frac{r-\rho}{r\gamma} \right)} + W_x(x,0) \left(rx + \frac{1}{\gamma} \ln W_x^S(x) - a \right) \tag{30}$$

Guess that the value function is of the form

$$W(x,0) = -B \frac{e^{-\gamma(rx + w_S + A)}}{\gamma r}$$

Now, apply the guess into (30),

$$\begin{aligned}
-\frac{(\rho+\phi)}{\gamma r} B e^{-\gamma(rx+w_S+A)} &= -B \frac{e^{-\gamma(rx+w_S+A)}}{\gamma} - \frac{\phi}{r\gamma} e^{-\gamma \left(rx + w_S - \frac{r-\rho}{r\gamma} \right)} \\
&\quad + B e^{-\gamma(rx+w_S+A)} \left(rx + \frac{1}{\gamma} \ln \left(B e^{-\gamma(rx+w_S+A)} \right) - a \right)
\end{aligned}$$

guess that $A=-\frac{r-\rho}{r\gamma}$. Then the equation can be reduced to,

$$\left(\phi+r\ln B-\gamma r\left(w_S+a\right)\right)B=\phi$$

it can also be written as,

$$\ln B=\frac{\phi}{r}\left(\frac{1}{B}-1\right)+\gamma\left(w_S+a\right)$$

Finally,

$$\begin{aligned} c &= -\frac{1}{\gamma}\ln W_x\left(x,0\right)=-\frac{1}{\gamma}\ln B+\left(rx+w_S-\frac{r-\rho}{r\gamma}\right) \\ &= rx-a-\frac{r-\rho}{r\gamma}-\frac{\phi}{\gamma r}\left(\frac{1}{B}-1\right) \end{aligned}$$

Provided that B exists we can check that the transversality condition holds:

$$\lim_{t\to\infty}e^{-\rho t}|W\left(x,0\right)|=\lim_{t\to\infty}e^{-\rho t}|-B\frac{e^{-\gamma\left(rx+w_S+A\right)}}{\gamma r}|=0$$

C.1 Existence and Uniqueness of B

The next two lemma summarizes the properties of B.

Lemma 5 (Existence of B) *There exists $B>1$ such that (9) is satisfied.*

Proof. Define $h\left(\tilde{B}\right)\equiv\left(\phi+r\ln\tilde{B}-\gamma r\left(w_S+a\right)\right)\tilde{B}-\phi$. Note that $h\left(\tilde{B}\right)$ is continuous. Further,

$$h\left(1\right)=-\gamma r\left(w_S+a\right)<0$$

And,

$$\lim_{\tilde{B}\to+\infty}h\left(\tilde{B}\right)=+\infty$$

Then, by the mean-value theorem there exists a $B\in\left(1,+\infty\right)$ such that $h\left(B\right)=0$. ■

Lemma 6 (Uniqueness of B) *B is unique.*
Proof. *Note that*

$$\lim_{\tilde{B}\to0+}h\left(\tilde{B}\right)=-\phi$$

also,

$$h'\left(\tilde{B}\right)=\phi-\gamma r\left(w_S+a\right)+r\left(1+\ln\tilde{B}\right)$$

and,

$$\lim_{\tilde{B}\to0+}h'\left(\tilde{B}\right)<0$$

We also have that

$$h''\left(\tilde{B}\right) = \frac{r}{\tilde{B}} > 0$$

This implies that $h\left(\tilde{B}\right)$ is strictly convex. It is decreasing for $\tilde{B} < \bar{B}$ where $h'\left(\bar{B}\right) = 0$ and increasing for $\tilde{B} > \bar{B}$. Moreover, $h\left(\bar{B}\right) < 0$. Therefore, there exists a unique B such that $h(B) = 0$. ■

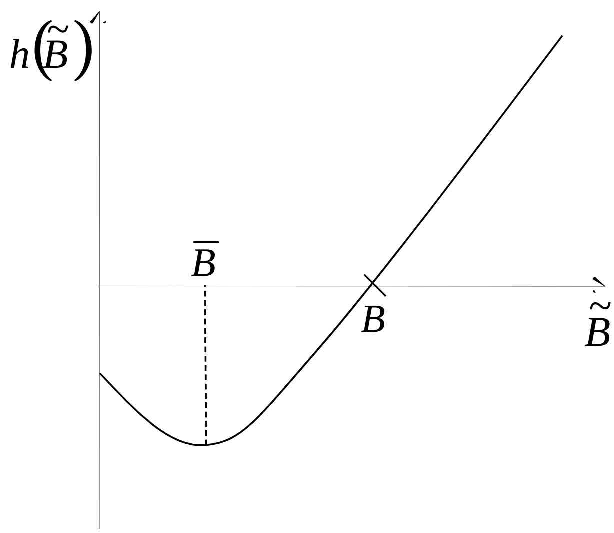

Figure 8: $h\left(\tilde{B}\right)$ and location of B.

Corollary 2 *The fact that $h(B) = 0$ implies that*

$$\left(\phi + r\ln B - \gamma r\left(w_S + a\right)\right) B > 0$$

Then,

$$\begin{aligned} r\ln B &> \gamma r\left(w_S + a\right) - \phi \\ B &> e^{\gamma\left(w_S + a\right) - \phi/r} \end{aligned}$$

and thus,

$$B \in \left(\max\left(1, e^{\left(\gamma\left(w_S+a\right) - \frac{\phi}{r}\right)}\right), +\infty\right)$$

This corollary will show to be useful for different calculations.

C.2 How does B changes with the different parameters of the model?

This exercise will prove to be useful later. Note that

$$\begin{aligned}
\frac{dB}{dw_S} &= \frac{\gamma r B}{r+\phi+r\ln B-\gamma r\left(w_S+a\right)} \\
\frac{dB}{da} &= \frac{\gamma r B}{r+\phi+r\ln B-\gamma r\left(w_S+a\right)} \\
\frac{dB}{dr} &= \frac{\gamma\left(w_S+a\right)B-\ln B}{r+\phi+r\ln B-\gamma r\left(w_S+a\right)} \\
\frac{dB}{d\phi} &= \frac{1-B}{r+\phi+r\ln B-\gamma r\left(w_S+a\right)} \\
\frac{dB}{d\gamma} &= \frac{r\left(w_S+a\right)B}{r+\phi+r\ln B-\gamma r\left(w_S+a\right)}
\end{aligned}$$

Further,

$$r+\phi+r\ln B-\gamma r\left(w_S+a\right)=\frac{Br+\left(\phi+r\ln B-\gamma r\left(w_S+a\right)\right)B}{B}=\frac{Br+\phi}{B}>0$$

and so,

$$\begin{aligned}
\frac{dB}{dw_S} &= \frac{\gamma r B^2}{Br+\phi}>0 \\
\frac{dB}{da} &= \frac{\gamma r B^2}{Br+\phi}>0 \\
\frac{dB}{dr} &= \frac{\gamma\left(w_S+a\right)B-\ln B}{\frac{Br+\phi}{B}}=-\left(1-B\right)\frac{\gamma\left(w_S+a\right)B+\frac{\phi}{r}}{Br+\phi}>0 \\
\frac{dB}{d\phi} &= \frac{\left(1-B\right)B}{Br+\phi}<0 \\
\frac{dB}{d\gamma} &= \frac{r\left(w_S+a\right)B^2}{Br+\phi}>0 \\
\frac{dB}{d\rho} &= 0
\end{aligned}$$

D Proof of proposition (Value of an unknown student)

(6) can be written as

$$\begin{aligned}
\left(\rho+\phi+\lambda\left(1-p\right)\right)W\left(x,p\right) &= \max_c\left(-\frac{e^{-\gamma c}}{\gamma}\right)+\lambda\left(1-p\right)W\left(x,0\right) \\
&\quad +\phi\left[pV\left(x;w_U\right)+\left(1-p\right)V\left(x;w_S\right)\right] \\
&\quad +W_x\left(rx-c-a\right)+W_p\left(\lambda p\left(1-p\right)\right)
\end{aligned}$$

The FOC wrt to c provides

$$\begin{aligned} -\frac{e^{-\gamma c}}{\gamma} &= -\frac{W_x}{\gamma} \\ c &= -\frac{1}{\gamma}\ln W_x \end{aligned}$$

Plugging into the original problem,

$$\begin{aligned} (\rho+\phi+\lambda(1-p))\,W(x,p) &= -\frac{W_x}{\gamma}-\lambda(1-p)\,B\frac{e^{-\gamma\left(rx+w_S-\frac{r-\rho}{\gamma r}\right)}}{\gamma r} \\ &\quad -\phi\left[p\frac{e^{-\gamma\left(rx+w_U-\frac{r-\rho}{r\gamma}\right)}}{\gamma r}+(1-p)\,\frac{e^{-\gamma\left(rx+w_S-\frac{r-\rho}{r\gamma}\right)}}{\gamma r}\right] \\ &\quad +W_x\,(rx-c-a)+W_p\,(\lambda p\,(1-p)) \end{aligned}$$

Guess that $W(x,p)=-\frac{e^{-\gamma(rx+f(p))}}{\gamma r}$ and apply into this last equation. After some algebra,

$$\begin{aligned} (\rho+\phi+\lambda(1-p))\,\frac{e^{-\gamma f(p)}}{\gamma r} &= \frac{e^{-\gamma f(p)}}{\gamma}+\lambda(1-p)\,B\frac{e^{-\gamma\left(w_S-\frac{r-\rho}{\gamma r}\right)}}{\gamma r} \\ &\quad +\phi\left[p\frac{e^{-\gamma\left(w_U-\frac{r-\rho}{r\gamma}\right)}}{\gamma r}+(1-p)\,\frac{e^{-\gamma\left(w_S-\frac{r-\rho}{r\gamma}\right)}}{\gamma r}\right] \\ &\quad +e^{-\gamma f(p)}\,(f(p)+a)-\frac{e^{-\gamma f(p)}}{r}f'(p)\,(\lambda p\,(1-p)) \end{aligned}$$

or

$$\begin{aligned} &(\rho-r+\phi+\lambda(1-p)+\gamma f'(p)\,(\lambda p\,(1-p))-\gamma r\,(f(p)+a))\,e^{-\gamma f(p)} \\ = \\ &\lambda(1-p)\,Be^{-\gamma\left(w_S-\frac{r-\rho}{\gamma r}\right)}+\phi\left[pe^{-\gamma\left(w_U-\frac{r-\rho}{r\gamma}\right)}+(1-p)\,e^{-\gamma\left(w_S-\frac{r-\rho}{r\gamma}\right)}\right] \end{aligned}$$

D.1 Comparative statics for p^*

D.1.1 Tuition

$$\begin{aligned} \frac{\partial p^*}{\partial a} &= -\frac{\gamma r e^{\gamma(w_S-w_U)}+\lambda\frac{dB}{da}}{(\phi+\lambda)\,e^{\gamma(w_S-w_U)}-\lambda B-\phi}-\lambda\frac{dB}{da}\frac{(\phi+\lambda-\gamma r\,(w_U+a))\,e^{\gamma(w_S-w_U)}-\lambda B-\phi}{\left[(\phi+\lambda)\,e^{\gamma(w_S-w_U)}-\lambda B-\phi\right]^2} \\ &= -\frac{\gamma r e^{\gamma(w_S-w_U)}+\lambda\frac{dB}{da}}{(\phi+\lambda)\,e^{\gamma(w_S-w_U)}-\lambda B-\phi}-\lambda\frac{dB}{da}\frac{(\phi+\lambda-\gamma r\,(w_U+a))\,e^{\gamma(w_S-w_U)}-\lambda B-\phi}{\left[(\phi+\lambda)\,e^{\gamma(w_S-w_U)}-\lambda B-\phi\right]^2} \end{aligned}$$

we know that $(\phi+\lambda-\gamma r\,(w_U+a))\,e^{\gamma(w_S-w_U)}-\lambda B-\phi>0$ and so $(\phi+\lambda)\,e^{\gamma(w_S-w_U)}-\lambda B-\phi>0$. Also we know that $\frac{dB}{da}>0$. Then,

$$\frac{\partial p^*}{\partial a}<0$$

D.1.2 Learning parameter

$$
\begin{aligned}
\frac{\partial p^*}{\partial \lambda} &= \left(e^{\gamma(w_S-w_U)}-B\right)\left(\begin{array}{c}\frac{1}{(\phi+\lambda)e^{\gamma(w_S-w_U)}-\lambda B-\phi} \\ -\frac{(\phi+\lambda-\gamma r(w_U+a))e^{\gamma(w_S-w_U)}-\lambda B-\phi}{\left[(\phi+\lambda)e^{\gamma(w_S-w_U)}-\lambda B-\phi\right]^2}\end{array}\right) \\
&= \frac{e^{\gamma(w_S-w_U)}-B}{\left[(\phi+\lambda)\,e^{\gamma(w_S-w_U)}-\lambda B-\phi\right]^2}\gamma r\,(w_U+a)\,e^{\gamma(w_S-w_U)} \\
&= \frac{e^{\gamma(w_S-w_U)}-B}{\left[(\phi+\lambda)\,e^{\gamma(w_S-w_U)}-\lambda B-\phi\right]^2}\gamma r\,(w_U+a)\,e^{\gamma(w_S-w_U)}
\end{aligned}
$$

We have that $h(B)=(\phi+r\ln B-\gamma r\,(w_S+a))\,B-\phi=0$. Also, because h is increasing for $B>1$ we have $e^{\gamma(w_s-w_u)}>B$ iff

$$
\begin{aligned}
(\phi+r\gamma\,(w_s-w_u)-\gamma r\,(w_S+a))\,e^{\gamma(w_s-w_u)}-\phi &> 0 \\
\phi\left(e^{\gamma(w_s-w_u)}-1\right)-r\gamma\,(w_U+w_a)\,e^{\gamma(w_s-w_u)} &> 0 \\
\phi\left(e^{\gamma(w_s-w_u)}-1\right)-r\gamma\,(w_U+w_a) &> 0
\end{aligned}
$$

Also, from (17),

$$
[\phi-\gamma r\,(w_u+a)]\,e^{\gamma(w_s-w_u)}-\phi>\lambda\left(B-e^{\gamma(w_s-w_u)}\right) \tag{31}
$$

Now suppose $e^{\gamma(w_s-w_u)}\leq B$.Then, we should have

$$
\begin{aligned}
\phi\left(e^{\gamma(w_s-w_u)}-1\right)-\gamma r\,(w_u+a) &\leq 0 \text{ and} \\
\lambda\left(B-e^{\gamma(w_s-w_u)}\right) &\geq 0
\end{aligned}
$$

which contradicts (31) above. Then, it follows that

$$
\frac{\partial p^*}{\partial \lambda}>0
$$

D.1.3 Graduation parameter

$$
\begin{aligned}
\frac{\partial p^*}{\partial \phi} &= \left(e^{\gamma(w_S-w_U)}-\lambda\frac{dB}{d\phi}-1\right)\left(\begin{array}{c}\frac{1}{(\phi+\lambda)e^{\gamma(w_S-w_U)}-\lambda B-\phi} \\ -\frac{(\phi+\lambda-\gamma r(w_U+a))e^{\gamma(w_S-w_U)}-\lambda B-\phi}{\left[(\phi+\lambda)e^{\gamma(w_S-w_U)}-\lambda B-\phi\right]^2}\end{array}\right) \\
&= \frac{\left(e^{\gamma(w_S-w_U)}-1-\lambda\frac{dB}{d\phi}\right)}{\left[(\phi+\lambda)\,e^{\gamma(w_S-w_U)}-\lambda B-\phi\right]^2}\gamma r\,(w_U+a)\,e^{\gamma(w_S-w_U)}>0
\end{aligned}
$$

D.1.4 Wage in the skilled sector

$$\begin{aligned}
\frac{\partial p^*}{\partial w_S} &= -\frac{\gamma r\left(w_U+a\right)\gamma e^{\gamma\left(w_S-w_U\right)}}{\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi} \\
&\quad +\frac{\left(\left(\phi+\lambda\right)\gamma e^{\gamma\left(w_S-w_U\right)}-\lambda\frac{dB}{dw_S}\right)}{\left[\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi\right]^2}\gamma r\left(w_U+a\right)e^{\gamma\left(w_S-w_U\right)} \\
&= \left(\frac{\left(\phi+\lambda\right)\gamma e^{\gamma\left(w_S-w_U\right)}-\lambda\frac{dB}{dw_S}}{\left[\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi\right]^2}-\frac{\gamma}{\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi}\right) \\
&\quad *\gamma r\left(w_U+a\right)e^{\gamma\left(w_S-w_U\right)} \\
&= \left(-\lambda\frac{dB}{dw_S}+\lambda\gamma B+\gamma\phi\right)\frac{\gamma r\left(w_U+a\right)e^{\gamma\left(w_S-w_U\right)}}{\left[\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi\right]^2} \\
&= \gamma\left(\lambda B\left(1-\frac{rB}{Br+\phi}\right)+\phi\right)\frac{\gamma r\left(w_U+a\right)e^{\gamma\left(w_S-w_U\right)}}{\left[\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi\right]^2}>0
\end{aligned}$$

D.1.5 Wage in the unskilled sector

$$\begin{aligned}
\frac{\partial p^*}{\partial w_U} &= \frac{\gamma r\left(w_U+a\right)\gamma e^{\gamma\left(w_S-w_U\right)}-\gamma r e^{\gamma\left(w_S-w_U\right)}}{\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi} \\
&\quad -\frac{\left(\phi+\lambda\right)\gamma e^{\gamma\left(w_S-w_U\right)}}{\left[\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi\right]^2}\gamma r\left(w_U+a\right)e^{\gamma\left(w_S-w_U\right)} \\
&= \left[\frac{\left(w_U+a\right)\gamma-1}{\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi}-\frac{\left(\phi+\lambda\right)\left(w_U+a\right)\gamma e^{\gamma\left(w_S-w_U\right)}}{\left[\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi\right]^2}\right]\gamma r e^{\gamma\left(w_S-w_U\right)} \\
&= -\left[\frac{\left(w_U+a\right)\gamma\left(\lambda B+\phi\right)}{\left[\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi\right]^2}+\frac{1}{\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi}\right]\gamma r e^{\gamma\left(w_S-w_U\right)}
\end{aligned}$$

Then,

$$\frac{\partial p^*}{\partial w_U}<0$$

D.1.6 Interest rate

$$\begin{aligned}
\frac{\partial p^*}{\partial r} &= -\left(\frac{1}{\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi}+\frac{\lambda r\frac{dB}{dr}}{\left[\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi\right]^2}\right) \\
&\quad *\gamma\left(w_U+a\right)e^{\gamma\left(w_S-w_U\right)}
\end{aligned}$$

Then,

$$\frac{\partial p^*}{\partial r}<0$$

D.1.7 Risk Aversion parameter

$$\begin{aligned}\frac{\partial p^*}{\partial \gamma} &= \frac{\gamma r\left(w_U+a\right)e^{\gamma\left(w_S-w_U\right)}}{\left[\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi\right]^2} \\ &\quad *\left(\left(\lambda B+\phi\right)\left(w_S-w_U\right)-\lambda\frac{dB}{d\gamma}-\frac{\left(\phi+\lambda\right)e^{\gamma\left(w_S-w_U\right)}-\lambda B-\phi}{\gamma}\right)\end{aligned}$$

E Deriving the probability of drop out $Q_i\left(p\right)$

For skilled students whose belief at time t is given by $p\left(t\right)=p$, the evolution of the dropout probability $Q\left(p\right)$ is given by

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

The fact that agents with prior 0 will never drop out implies that $Q_S\left(0\right)=0$. Rewritte the previous equation as

$$Q_S\left(p\right)=\left(1-\phi dt\right)\left(1-\lambda dt\right)Q_S\left(p\right)+\left(1-\phi dt\right)\left(1-\lambda dt\right)\left[Q_S\left(p+\lambda p\left(1-p\right)dt\right)-Q_S\left(p\right)\right]$$

disregarding the $\left(dt\right)^2$ terms,

$$\left(\phi+\lambda\right)dt Q_S\left(p\right)=\left(1-\left(\phi+\lambda\right)dt\right)\left[Q_S\left(p+\lambda p\left(1-p\right)dt\right)-Q_S\left(p\right)\right]$$

Set $dh=\lambda p\left(1-p\right)dt$ or $\frac{dh}{\lambda p(1-p)}=dt$ so,

$$\begin{aligned}\left(\phi+\lambda\right)\frac{dh}{\lambda p\left(1-p\right)}Q_S\left(p\right) &= \left(1-\frac{\left(\phi+\lambda\right)dh}{\lambda p\left(1-p\right)}\right)\left[Q_S\left(p+dh\right)-Q_S\left(p\right)\right] \\ \left(\phi+\lambda\right)Q_S\left(p\right) &= \lambda p\left(1-p\right)\left(1-\frac{\left(\phi+\lambda\right)dh}{\lambda p\left(1-p\right)}\right)\frac{Q_S\left(p+dh\right)-Q_S\left(p\right)}{dh}\end{aligned}$$

Taking the limit when $dh\rightarrow 0$,

$$\left(\phi+\lambda\right)Q_S\left(p\right)=\lambda p\left(1-p\right)Q_S'\left(p\right)$$

with terminal condition $Q\left(p^*\right)=1$. The solution to this differential equation is

$$Q_S\left(p\right)=\left(\frac{p}{p^*}\frac{1-p^*}{1-p}\right)^{\frac{\phi+\lambda}{\lambda}}$$

Similarly, the dropout rate of unskilled students are given by

$$Q_U\left(p\right)=\left(\frac{1-p^*}{1-p}\frac{p}{p^*}\right)^{\frac{\phi}{\lambda}}$$

F Approximation of $f(p)$

Recall that $\gamma f'(p)(\lambda p(1-p)) = \theta(p, f(p))$. We can rewrite this expression as

$$f'(p) = \frac{\theta(p, f(p))}{\gamma(\lambda p(1-p))} \tag{32}$$

This equation will prove useful in a moment. Also keep in mind that $f'(p^*)$ and $f(p^*)$ are known (by the VM and SP conditions).

Say we are interested in the value $f(p^* - \Delta)$.[15] With this in mind, we do a Taylor expansion of $f(p^* - \Delta)$ around $\Delta = 0$,

$$f(p^* - \Delta) \approx f(p^*) + f'(p^*)\Delta$$

If we also want the value of $f(p^* - 2\Delta)$ we can do a Taylor expansion of $f(p^* - 2\Delta)$ around $2\Delta = \Delta$,[16]

$$f(p^* - 2\Delta) \approx f(p^* - \Delta) + f'(p^* - \Delta)\Delta$$

We can apply to this equation the formula for $f'(p)$ that we have in (32),

$$f(p^* - 2\Delta) \approx f(p^* - \Delta) + \frac{\theta(p^* - \Delta, f(p^* - \Delta))}{\gamma(\lambda(p^* - \Delta)(1 - (p^* - \Delta)))}\Delta$$

This suggests that we can produce a recursive algorithm to approximate $f(p)$ for any $p \leq p^*$ in the following way:

$$f(p^* - (i+1)\Delta) \approx f(p^* - i\Delta) + \frac{\theta(p^* - i\Delta, f(p^* - i\Delta))}{\gamma(\lambda(p^* - i\Delta)(1 - (p^* - i\Delta)))}\Delta \tag{33}$$

G Data Appendix

The National Center for Education Statistics (NCES) provides a wide variety of data that can be easily gathered from their online site.[17] Table 25.2 in their website provides that 63.6% of males and 70.3% of females that obtained a high-school degree in 1996 attended some sort of college in 1997. The percentage of females in population is 51.5% so the total percentage of people attending college is 67.0505%. Further, 66.01% of the ones that attend college enroll into a 4-yr institution while the rest attend a 2-yr institution.

Table 310 provides information about the type of degree that a student enrolled in a college obtains. In particular, it follows the 1995-96 cohort and

[15] remember that it must be the case that $p \leq p^*$.

[16] we can also do the expansion around $2\Delta = 0$ but in that case the quality of the approximation will be reduced.

[17] http://nces.ed.gov/index.asp

check on them again in 2001 (so 6 years later). Next we transcribe some facts from that table:

	2-yr college	4-yr college
Bachelor	9.7%	58.4%
Associate and Certificate	22.8%	6.7%
Drop-out	45.2%	20.5%
Still enrolled	16.4%	14.4%

Students that obtain an Associate degree or a Vocational Certificate are considered in our model to work in the unskilled sector.[18] These degrees are obtained in 2-yr schools so we need to adjust the total percentage of people attending college and change the previous table in order to correct for this fact.

	4-yr college
College graduates	63.46%
Drop-out	21.46%
Still enrolled	22.14%

: 43.6The still enrolled students are tough to analyze. From the data we know that de drop-out rate in 4-yr college is something between 21.46% and 43.6%. Many of the still enrolled students are probably almost finishing college so we can consider them future drop-outs. For the moment, we will assum that half of them are drop-outs and half become bachelors. Then,

	4-yr college
College graduates	67.47%
Drop-out	32.53%

Our measure of college enrollment is

$$44.26\% + 0.097 * (22.7905\%) - 0.067 * (44.26\%) = 43.5\%$$

Wage data can also be obtained from the NCES website. Wages in 1995 at constant 2004 dollars are (see Table 20.1)

High-school graduate	26400
Some college+Assoc.	30200
College graduate	41100

where we had assumed that agents with an Associate Degree or Vocational Certificate earn the same wage than a college drop-out. Further, these are unskilled workers so,

Unskilled wage	28369
Skilled wage	41100

[18] even though they are probaly white-collar they are unskilled.

The expected time until completion of college education is on average 55 months (see Table 21.1 in the website). In years this mean that

$$\boxed{\text{Average Time until completion: 4.58 years}}$$

Table 319 on the website provides information about tuition. Average tuition, room, and board for 1997-2001 period is 12110 dollars. This number needs to be corrected for the fact that many students usually work while in school. Assuming that 40% of them work for 32 weeks, earning a weekly salary of 200 dollars, and facing a tax rate of 15%, the net cost of college is

$$\boxed{\text{Net cost of college per year: 9934 dollars}}$$

References

[Altonji (1993)] Altonji, Joseph G. (1993). "The Demand for and Return to Education when Education Outcomes are Uncertain", *Journal of Labor Economics*, Vol. 11, No. 1

[Belley-Lochner (2008)] Belley, Phillipe and Lochner, Lance (2008). "The Changing Role of Family Income and Ability in Determining Educational Achievement", *Journal of Human Capital*, Vol. 1, No. 1.

[Cameron-Heckman (2001)] Cameron, Stephen and Hecman, James J. (2001). "The Dynamics of Educational Attainment for Blacks, Whites and Hispanics". *Journal of Political Economy*, Vol. 109, No. 3.

[Cameron-Taber (2004)] Cameron, Stephen and Taber, Christopher (2004). "Borrowing Constraints and the Returns to Schooling". *Journal of Political Economy*, Vol. 112.

[Carneiro-Hansen-Heckman (2003)] Carneiro, Pedro, Hansen, Karsten T. and Heckman, James J. (2003). "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice". *International Economic Review*, Vol. 44, No. 2.

[Comay-Pollatschek (1973)] Comay Y., Melnik, A. and Pollatschek, M. A. (1973). "The Option Value of Education and the Optimal Path for Investment in

Human Capital", *International Economic Review*, Vol. 14, No. 2.

[Cunha-Heckman (2007)] Cunha, Flavio and Heckman, James J. (2007). "Indetifying and Estimating the Distribution of Ex Post and Ex Ante Returns to Schooling", *Labour Economics*, Vol. 14, No. 6.

[Dixit (1993)] Dixit, Avinash (1993). "The Art of Smooth Pasting", *Fundamentals of Pure and Applied Economics*, No. 55.

[Hansen-Sargent (2001)] Hansen, Lars P. and Sargent, Thomas J. (2001). "Robust Control and Model Uncertainty", *AER.*

[Heckman-Navarro (2006)] Heckman, James J. and Navarro, Salvador (2006). "Dynamic Discrete Choice and Dynamic Treatment Effects", *Journal of Econometrics.*

[Heckman-Stixrud-Urzua (2006)] Heckman, James J., Stixrud, Jora and Urzua, Sergio (2006). "The Effects of Cognitive and Noncognitive Abilities on Labor Markets Outcomes and Social Behavior", *Journal of Labor Economics*, Vol. 24, No. 3.

[Judd (2000)] Judd, Kenneth L. (2000). "Is Education as Good as Gold? A Portfolio Analysis of Human Capital Investment", *unpublished*, Hoover Institution.

[Keane-Wolpin (1997)] Keane, Michael P. and Wolping, Kenneth I. (1997). "The Career Decisions of Young Men" *Journal of Political Economy*, Vol. 105, No. 3.

[Keane-Wolpin (2001)] Keane, Michael P. and Wolping, Kenneth I. (2001). "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment" *International Economic Review*, Vol. 42, No. 4.

[Miao-Wang (2007)] Miao, Jianjun and Wang, Neng (2007). "Experimentation under Uninsurable Idiosyncratic Risk: An Application to Entrepreneurial Survival", *unpublished.*

[Navarro (2005)] Navarro, Salvador (2005). "Understanding Schooling: Using Observed Choices to Infer Agent's Information in a Dynamic Model of Schooling Choice when Consumption Allocation is Subject to Borrowing Constraints". *unpublished*, University of Chicago.

[Stange (2007)] Stange, Kevin (2007). "An Empirical Examination of the Option Value of College Enrollment". *mimeo*, University of California Berkeley