Learning and The Dynamics of Post-secondary Education: The Role of Academic 2-year Colleges (Preliminary and Incomplete)

Nicholas Trachter*

March 30, 2009

Abstract

I construct a model of educational choice where high-school graduates can decide whether to enroll in 2- or 4-year Colleges in an environment with uncertainty about future income flows. Uncertainty is linked to observable characteristics of high-school graduates. There will be sorting in the initial enrollment status. During tenure as students, agents are faced with exams that can induce updating of beliefs about future income streams. This updating generates that students decide to dropout or transfer to a different educational institution.

This paper has four main contributions to the literature. First, it provides a model of educational choice that incorporates both 2-year and 4-year where dropouts and transfers are endogenously determined. Second, it provides evidence suggesting that part of the gap between returns to College education for the marginal student and work found in many studies can be explained by incorporating 2-year Colleges. Third, it provides a link between heterogeneity and uncertainty that can potentially explain attrition rates. Fourth, it provides a set of testable predictions. Finally, it discusses a set of instruments

*The University of Chicago. Contact: trachter@uchicago.edu. I thank James Heckman, Derek Neal, Robert Shimer, and specially Fernando Alvarez for guidance and useful comments. I also benefited from the comments by Jeremy Fox, Ezra Oberfield, Andy Zuppann, Aspen Gorry and seminar participants at University of Chicago, Universidad Torcuato Di Tella and Conference participants at LACEA/LAMES 2008 at Rio de Janeiro. All errors are my own.

that can be used to provide insurance against the market incompleteness that arises from the uncertain environment.

1 Introduction

(CHANGE)

Many empirical studies analyzed labor-market returns to 4-year College graduates and its relation to the apparent low enrollment and graduation rates attached to 4-year Colleges. A few studies extended the analysis to labor-market returns to 2-year College graduates. Of the former, Altonji 1993 [1] introduces the notion of uncertain outcomes into the educational choice problem faced by high-school graduates. Among others, Cunha-Heckman 2007 [8], Carneiro-Hansen-Heckman [6], Keane-Wolpin 1997 [19] and Keane-Wolpin 2001 [20], introduce the notion of psychic costs. Psychic costs are modeled as an unobservable (for the econometrician) idiosyncratic component of utility function or income process and are estimated to have high predictive power in terms of College enrollment. An implicit result of Cunha-Heckman 2007 [8] is that there exists underinvestment in 4-year College education.[1] Of the later, work by Kane and Rouse [16, 17, 18] and Grubb [11] find that 2-year college graduates enjoy an increase in their wage profiles. All of these models are silent regarding the option value of dropouts.

Recent work by Heckman and Urzua [13] extends the analysis to allow high-school graduates (and GED takers) to decide between enrollment in 2-year or 4-year College in an environment where dropping out belongs to the set of strategies available to the student. They find that the option value of dropping out is more important in setups with learning about ability as opposed to configurations where learning about ability is not available. Further, they show that measures of ex-post returns tend to over-represent returns to education as option values are not present there. Opposed to it, they show that ex-ante returns, that account for the option value, seem more reasonable.

In this paper I construct a model of educational choice where high-school graduates can decide whether to enroll in 2- or 4-year Colleges in an environment with uncertainty about future income flows. In the model, uncertainty is linked to observable characteristics of high-school graduates and as so there will be sorting in the initial enrollment status. During tenure as students, agents will be faced with exams that might induce updating in beliefs about future

[1] Judd 2000 [15] perform a portfolio analysis comparaing 4-year Colleges with investment opportunities of same risk and liquidity. Judd's results also suggest underinvestment in education.

income streams. This updating will generate that students might decide to dropout or transfer to a different kind of educational institution. For previous work on the topic, Ozdagli-Trachter 2008 [26] analyze a simpler version of this model for the enrollment/dropout decision to 4-year Colleges. Miao-Wang 2007 [23] follows a similar strategy to understand entrepreneurial survival.

This paper has three main contributions to the literature. First, it provides a model of educational choice that incorporates both 2-year and 4-year where dropouts and transfers are endogenously determined. Uncertainty, and thus risk, will play a crucial role in the model. Then, oposed to many papers in the litarature, I work in a risk-averse environment. Second, the paper provides a link between heterogeneity and uncertainty that can potentially explain the attrition rates. Heckam-Urzua 2008 [13] works in an environment where the ex-ante expectation of uncertainty (given by θ in their paper) is 0 at time 0 for every agent. I model the relation between uncertainty and heterogeneity in a tighter way. As so, there will be a stronger link between initial heterogeneity (and its distribution) and the dynamic educational paths of each student. Third, the model provides a set of testable predictions

The paper is organized as follows. Section 2 presents evidence about the differences between 2-year and 4-year Colleges. Section 3 introduces the model's environment. In Section 4 I analyze models with no learning an discuss their limitations. First, I solve the model with no uncertainty in order to suggest that an environment with uncertainty is a necessary condition. Second, I introduce uncertainty and analyze its limitations. In Section 5 I introduce learning into the model in order to generate flows as observed in the data. Section 6 presents a set of testable predictions implied by the model. In section 7 I simulate the model. In section 8 I discuss the estimation strategy and attempt to estimate a simpler version of the model.

2 Understanding Post-secondary Education Dynamics and the role of Academic 2-year Colleges

In this section I present evidence of the role played by Academic 2-year Colleges on Post-secondary education. Figure 1 shows that the fraction of freshmen enrollment in 2-year

	Work	V.S.	A. 2-year	4-year
men	0.49	0.41	0.54	0.52
black	0.11	0.12	0.08	0.1
Father_LHS	0.516	0.38	0.3	0.21
Father_HS	0.324	0.4	0.34	0.29
Father_DR	0.11	0.15	0.21	0.19
Father_C	0.05	0.07	0.15	0.3
Socio_Low	0.41	0.3	0.2	0.17
Socio_Med	0.50	0.58	0.55	0.41
Socio_High	0.08	0.12	0.25	0.42
Rank	0.495	0.44	0.4	0.27

Table 1: **Sorting**. Proportion of initially enrolled in a given type of institution (or joining workforce) with a particular charactersitic. Father_LHS: maximum educational level of father less than high-school. Father_HS: maximum educational level of father is high-school degree. Father_DR: maximum educational level of father is taking some classes in College. Father_C: maximum educational level of father is College degree. Socio_Low, Socio_Med, Socio_High: socioeconomic status of family. Rank: ratio of rank in high-school class to total enrollment in class.

Colleges[2] is not a negligible fraction of total freshmen enrollment. In fact, as discussed in Betts-McFarland 1995 [3], 2-year Colleges absorved most of the increase in Post-secondary education during the 80's.

2-year Colleges combine academic and vocational courses.[3] Table 1 shows that there is sorting in terms of observable characteristics of ability into the different kind of academic institutions (Vocational School, Academic 2-year Colleges and 4-year Colleges) and work. The option value for Vocational School students that arises from the transfer option is small as it can be inferred from Appendix B, where it can be observed that very few students transfer from or into Vocational School. I will show that the value of this option seems to be important for students that enroll in Academic 2-year Colleges as it will be suggested by the transfer patterns and the Internal Rate of Returns.

The set of Post-secondary institutions represented in my dataset is a combination of 4-year colleges (from now on College) and Academic 2-year Colleges (from now on Junior College).

In October of 1972, of the full set of individuals available in my dataset[4], around 56% reported been working and with no participation in a post-secondary institution and 44%

[2] includes Vocational School

[3] there are some Colleges that specialize in a particular type of education.

[4] see Appendix A for a detailed description.

Figure 1: **Time series for the ratio of Freshmen Enrollment at Junior College to Freshmen Enrollment at College.** The dotted line plots the average.

were enrolled in a post-secondary institution. Of the set of individuals listed as students, 62.7% were enrolled in College. Table 2 and 3 present differences in aggregate patterns for students initially enrolled in College and Junior College. Table 2 shows that (1) graduation is much less likely at Junior College, (2) the transfer option seems to be important only for students initially enrolled in Junior College, and (3) many students dropout in both types of institutions (but more in Junior College). Table 3 provides some details for the dynamics of transferred students, among them we have that (1) graduation after transfer is more likely for students that transfer from Junior College to College, and (2) among students that transfer from Junior College to College, those with an Associate's degree never graduate in College. The proportion of students that become dropouts and those that transfer are decreasing functions of time (Figures 2 and 3).

Gross returns (change in wage relative to individuals that didn't attend school) also exhibit important differences. Table 4 presents the results of a regression using as dependent variable the log of wages and as independent variables time in a particular institution, dummies for graduation and variables that measure the agent's ability. As in Kane-Rouse 1995 [16], years of schooling (either in College or Junior College) increase returns and graduation in College is associated with an increase in wages. As in Grubb 1995 [11], graduation at Junior College has negative value.[5] The first column of Table 5 provides the estimates for the cross-section of agents, for College and Junior College dropouts and graduates. As expected, College graduates enjoy higher returns that Junior College graduates.

Total cost of education also differs between College and Junior College. The average cost of College in 1972 , in 1984 dollars, was 4784.4 (3725.8) . The average cost of Junior College in 1972, in 1984 dollars, was 2608.4 (4507).

The sorting discussed above in terms of observable measures of ability, together with the fact that transfers are very likely and graduation unlike and unprofitable, suggest that (a) Junior College might serve as a *stepping stone* towards more challenging educational environments, and (b) the Internal Rate of Return (IRR) for the marginal agent enrolled in College is not between College and Work but rather between College and Junior College. This argument implies that part of the average IRR for College enrollment is explained by the IRR for Junior College enrollment, potentially explaining the puzzle that arises from Cunha-Heckman 2007 [8], where the authors account for large returns for the marginal agent enrolled in College.

The IRR is the interest rate that leaves an investor indifferent between the benchmark investment (joining the workforce directly) and the different configurations of education.[6] In column 2 of Table 5 I present the IRR calculations. First of all, note that the IRR for College graduation (for students initially enrolled in College) is 3.2%. Cunha-Heckman 2007 [8] point out that, as returns to College education are big, there exists a puzzle as enrollment and graduation rates are low.[7] My analysis suggest that, once direct costs of education (tuition, room and board, etc) and foregone earnings are included into the picture, the return to College graduation does not seem to be unreasonably high. Second, the IRR for Junior College graduation is negative.

The fact that the IRR for Junior College graduation is negative raises the question of why so many agents enroll in Junior College. Is it just an irrational decision? is that they missjudge

[5] Grubb's findings apply for 2-year College graduates with academic degrees that match the sample of students enrolled in Junior College I am using here.

[6] see Appendix C for details.

[7] this analysis does not include the Junior College alternative.

	$q = J, \neg q = C$	$q = C, \neg q = J$
% of initial enrollment in q that graduated in q	9.8%	55.4%
% of initial enrollment in q that dropped out in q	61.2%	42.4%
% of initial enrollment in q that transferred to $\neg q$	32.2%	2.2%
% of transferred from q to $\neg q$ that graduated in q	10%	0%

Table 2: **Aggregate Dynamics for students initially enrolled in instituion q.**

their ability? or is it something else? the answer will be related to the stepping stone idea.

To answer this question I computed the IRR for for every agent that enroll in College or Junior College independently of their graduation status, meaning that for each agent I compute the IRR for any particular enrollment path (transfer, dropout, etc.). Table 6 presents the results. Not surprisingly, as transfers from College to Junior College and uncommon, the average gross return for College enrollment lies between the gross return for college graduation and dropouts. The average IRR for College enrollment is 3.04%, somewhat smaller that the IRR for graduation and dropouts, as students that transferred from College to Junior College enjoyed lower wage profiles. The average gross return to Junior College enrollment is much higher than the one for Junior College dropouts and graduates as many students transfer to College and everntually graduate. Further, the average IRR is 1.51% suggesting that the transfer option is an important part of the value of Junior College, and thus provides compeling evidence in favor of the stepping stone story.

Also, note that 50% of the average IRR for College enrollment can be explained by the average IRR for Junior College enrollment.

An important disclaimer has to be made with respect to the IRR calculations. First, I assumed that agents are forced to retire/die at age 75. If this value is increased, IRR of graduation would increase. Second, I assumed that the total cost of education includes room and board. Note that agents that do not enroll in College or Junior College has no costs (see Appendix C) even though they should be paying for rent. The idea is that for the IRR I present here I am assuming that, while young, workers can live in their parents home. If I modify this assumption the IRR will also go up. Finally, I assumed no growth in wages after graduation. Many studies have suggested that wages increase faster for College graduates and

	$q = J, \neg q = C$	$q = C, \neg q = J$
From q to $\neg q$ that graduated in $\neg q$	55.5%	10.5%
From q to $\neg q$ (with degree) that graduated in $\neg q$	0%	NA
From q to $\neg q$ (no degree) that graduated in $\neg q$	62.1%	10.5%

Table 3: **Aggregate Dynamics for transferred students.**

thus the IRR should also increase.

3 Environment

The economy is populated by agents that decide whether to join the labor force or pursue a degree at a post-secondary education institution. At $t = 0$ agents graduate from high-school. A high-school graduate skill level can either be low (unskilled) or high (skilled). Let μ denote this skill level, with $\mu \in \{0, 1\}$, with corresponding probabilities $1 - p$ and p. $\mu = 0$ means that the skill level is low while $\mu = 1$ means that the skill level is high. The skill level μ is not observable by the agent. Instead, a high-school graduate inherits a signal about her true type, denoted by $\vartheta \in [0, 1]$. Let $j_\mu(\vartheta)$ denote the density of signal ϑ conditional on the true type of the agent being μ, with $j_0 : \mathbb{R} \rightarrow [0, 1)$ and $j_1 : \mathbb{R} \rightarrow (0, 1]$. With the information at hand, a high-school graduate will generate her subjective belief about her own true type $p \in [0, 1]$, where $p = Pr(\mu = 1)$.

At any period in time an agent can be either working, studying at College, or studying at Junior College. Let $q \in \{C, J\}$ denote the type of institution, where C stands for 4-year College and J stands for Junior College, Community College or 2-year College. Tuition is modeled as a pay-as-you-go system with cost a^q per unit of time, with $a^C > a^J$. The graduation hazard rate is ϕ^q_μ. Upon graduation, all uncertainty is revealed. There is also a one-time up front cost of entering the post-secondary educational system, E.

Agents that do not enroll in College or Junior College (i.e. join the workforce directly) enjoy a constant wage profile w. Dropouts enjoy a wage profile $\underline{w}^q$, with $\underline{w}^C \geq \underline{w}^J \geq w$. Upon graduation, the wage profile is $\overline{w}^q_\mu$, with $\overline{w}^q_\mu \geq \underline{w}^q$ for any μ.

A particular high-school graduate, indexed by j, is endowed with a subjective initial prior

$\ln(wages)$	(1)	(2)
Years in C	0.054 (0.011)	-
Years in JC	0.03 (0.018)	-
Graduation at C	0.021 (0.058)	0.235 (0.03)
Graduation at JC	-0.06 (0.12)	0.058 (0.11)
R^2	0.2226	0.2165
F	14.37	14.34

Table 4: **Wage regression.** In the regression I also included a dummy for men, one for race, variables that account for socieconomic status of family, variables that account for maximum educational level of father, rank in high-school class, geography dummies and a constant.

p_0^j and an initial level of assets X^j. She chooses her consumption stream $\{c_t : t \geq 0\}$ and whether to enroll in, dropout or transfer from College or Junior College, in order to maximize her time-separable expected discounted lifetime utility derived from consumption

$$E\left\{\int_0^\infty e^{-rt}\left(\frac{e^{-\gamma c_t}}{-\gamma}\right)dt\middle|\mathcal{F}_0\right\}$$

where $\mathcal{F}_0 = (p_0^j, X_0^j)$ and γ is the coefficient of Constant Absolute Risk Aversion (CARA). The evolution of wealth is given by

$$\frac{dx}{dt} = \begin{cases} rx - a^q - c & if\, enrolled\, at\, institution\, q \\ rx + \zeta - c & if\, working\, with\, wage\, \zeta \end{cases}$$

where no borrowing constraints are present. Cameron-Heckman 2001 [4], Keane-Wolpin 2001 [20] and Cameron-Taber 2004 [5] find that there is no enough evidence in favor of existence of borrowing constraints using data from the NLSY.[8]

Let $V(x;\zeta)$ denote the value of being a worker with current wealth x and wage profile ζ. Let $W(x, p, q, S, TS)$ denote the value for an agent with current wealth x, prior p, currently enrolled in institution q, in educational stage $S \in \{1, 2\}$, and transfer status $TS \in \{0, 1\}$ as students are going to be allowed to transfer only once.[9] In order to avoid a continuous pattern

[8]Belley-Lochner 2008[2] and Lochner-Monge Naranjo 2007 [22] argue that credit constraints become binding in the 90's.

[9]this assumption is used in order to reduce the size of the state space (otherwise time should be part of the state).

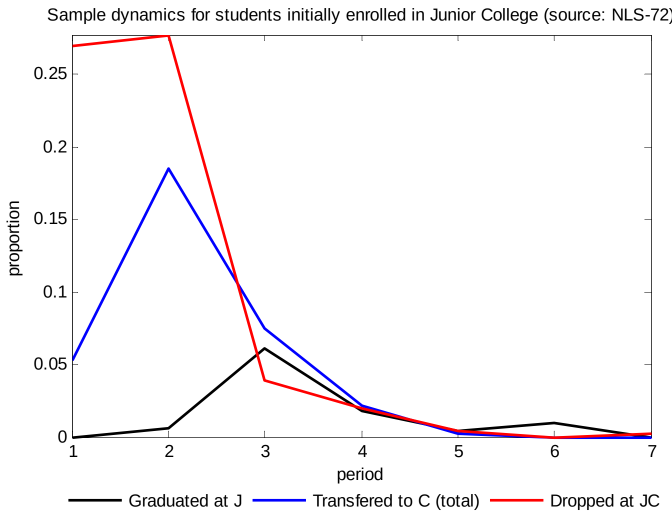

Figure 2: **Time series for students initially enrolled at Junior College.** The unit of the time period is schooling year (October to September).

of transfers some kind of friction needs to be included, other than keeping time as a state variable. The usual method of including a fixed cost of transfer will shut down completely the transfer option (with CARA instant utility there are no income effects). Allowing students to transfer only once is similar to provide to each high-school graduate two tickets - one for College and one for Junior College -.

4 Model with no Uncertainty

Next I solve two different problems. First, I analyze the problem of a worker. Next, I analyze the problem of a student with no uncertainty about her skill level. As no learning will be available while attending institution q, no dropout or transfer decision will be taken during tenure as student.

	Gross Return	IRR
C graduates (initially enrolled in C)	0.307 (0.05)	0.032 (0.008)
J graduates (initially enrolled in J)	0.058 (0.03)	−0.014 (0.01)
C dropouts (initially enrolled in C)	0.165 (0.09)	0.032 (0.009)
J dropouts (initially enrolled in J)	0.053 (0.02)	0.014 (0.007)

Table 5: **Gross Returns and Internal Rate of Return for Graduates and Dropouts.** Computing Gross Returns: pick an individual initially enrolled in institution q that also graduated at q. The gross return for this agent is just $\beta(years\,in\,q) * years\,in\,q + \beta(grad\,at\,q)$. Internal Rate of Return: see Appendix C.

4.1 Value at Work

The problem faced by a worker is given by

$$rV(x;\zeta) = \max_c \frac{e^{-\gamma c}}{-\gamma} + V_x(x;\zeta)\frac{dx}{dt} \tag{1}$$

where

$$\frac{dx}{dt} = rx + \zeta - c$$

Equation (1) states that the flow value of being a worker equals the instant utility derived from consumption in addition to the change in value due to the change in time (through wealth).

The solution to problem (1) is given by

$$V(x;\zeta) = -\frac{1}{\gamma r}e^{-\gamma c^V(x;\zeta)} \tag{2}$$

with consumption rule given by

$$c^V(x;\zeta) = rx + \zeta$$

so that the Permanent Income Hypothesis holds as workers consume the interest proceeds from their current wealth in addition of their period income [10].

	Gross Return	IRR
C	0.24 (0.1)	0.0304 (0.003)
J	0.116 (0.097)	0.0151 (0.009)

Table 6: **Gross Returns and Internal Rate of Return by initial enrollment status.**

4.2 No Uncertainty about Skill Level

Assumption 1. *The following inequalities hold*

$$\begin{cases} W(x - rE, 0, C, 0) < W(x - rE, 0J, 0) < V(x; w) \\ W(x - rE, 1, C, 0) > W(x - rE, 1, J, 0) > V(x; w) \end{cases}$$

This assumption provides: (a) low-skilled agents will join the workforce directly, (b) low-skilled students, if enrolled in an academic institution, are worse-off in College, (c) high-skilled agents will enroll in an academic institution and, (d) high-skilled agents are better-off by enrolling in College.

Corollary 2. *A low-skilled Junior College graduate ($\rho = 0$) will not pursue College education. In other words,*

$$V\left(x; \overline{w}_0^J\right) > W(x, 0, C, 0)$$

Proof. A low skilled Junior College graduate enjoys wage profile $\overline{w}_0^J > w$. By Assumption (1), and by the fact that $\frac{dV(x;\zeta)}{d\zeta} > 0$, $V\left(x; \overline{w}_0^J\right) > V(x; w) > W(x, 0, C, 0)$. □

Assumption 3. *A high-skilled Junior College graduate ($\rho = 1$) will not pursue College education. In other words,*

$$V\left(x; \overline{w}_1^J\right) \geq W(x, 1, C, 0)$$

This assumption implies that no Junior College graduate will enroll in College.[10]

[10]This simplyfying assumption follows from the fact that no Junior College graduate eventually graduated in College.

The value of being a student with skill level μ at institution q solves

$$\begin{aligned} rW(x,\mu,q,0) &= \max \frac{e^{-\gamma c}}{-\gamma} + W_x(x,\mu,q,0)\frac{dx}{dt} \\ &\quad +\phi_\rho^q\left(\max\left(W(x,\mu,\neg q,0), V\left(x;\overline{w}_\mu^q\right)\right) - W(x,\mu,q,0)\right) \end{aligned}$$

with

$$\frac{dx}{dt} = rx - a^q - c$$

where $\neg q$ stands for institution other than q. Corollary (2) and Assumption (3) provides

$$\max\left(W(x,\mu,\neg q,0), V\left(x;\overline{w}_\mu^q\right)\right) = V\left(x;\overline{w}_\mu^q\right)$$

so that the previous problem can be simplified to

$$\begin{aligned} rW(x,\mu,q,0) &= \max \frac{e^{-\gamma c}}{-\gamma} + W_x(x,\mu,q,0)\frac{dx}{dt} \\ &\quad +\phi_\mu^q\left(V\left(x;\overline{w}_\mu^q\right) - W(x,\mu,q,0)\right) \end{aligned} \tag{3}$$

with

$$\frac{dx}{dt} = rx - a^q - c$$

Equation 3 states that the flow value of being a student with current wealth x, skill level μ at intitution q with transfer status $TS = 0$ is equal to the instant utility derived from consumption plus the expected change in value due to graduation in addition to the change in value due to the change in time.

The solution to this problem is given by

$$W(x,\mu,q,0) = -\frac{A_\mu^q}{\gamma r}e^{-\gamma\left(rx+\overline{w}_\mu^q\right)} \tag{4}$$

where A_μ^q solves $\left(\phi_\mu^q - r\gamma\left(a^q + \overline{w}_\mu^q\right) + r\ln A_\mu^q\right)A_\mu^q = \phi_\mu^q$.

Appendix (E) provides the details of the solution.

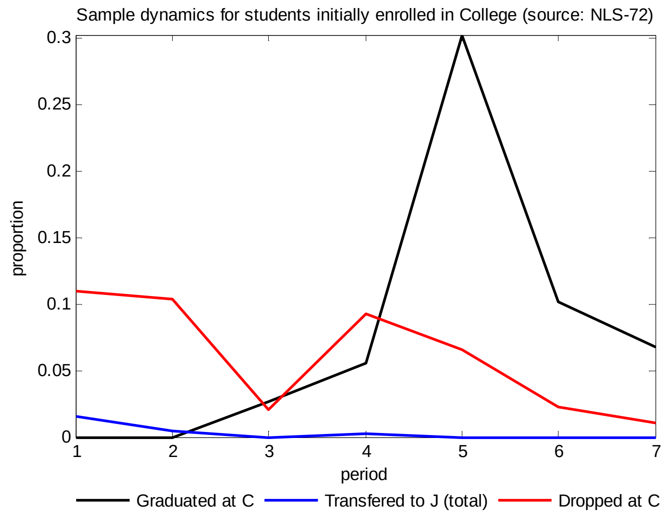

Figure 3: **Time series for students initially enrolled at College.** The unit of the time period is schooling year (October to September).

Proposition 4. If $\phi_\mu^q - r\gamma\left(a^q + \overline{w}_\mu^q\right) > 0$, A_μ^q exists, it is unique, and $A_\mu^q > 1$.

Proof. See Appendix (E). □

Note that current consumption is given by $rx + \overline{w}_\mu^q - \frac{\ln A_\mu^q}{\gamma}$ that satisfies Friedman's Permanent Income Hypothesis, as discussed in Wang [28] once we realize that the last term of this expression takes into account the time (and costs attached to it) until graduation.

Assumption 5. *Accumulated human capital can not be carried on when a student transfers.*

5 Models with Uncertainty

In this section I extend the model to setups where uncertainty about skill level is present. I will first describe the learning process that drives the updating of the prior p. Next I will discuss a model with learning but with no dropout or transfer option. Finally, I will extend the model to include these options.

so that students are allowed to dropout and/or transfer in an environment with uncertainty about their true skill level.

5.1 Learning Process

Learning at institution q is the result three different objects: (a) as graduation hazard is type dependent survival as student conveys information, (b) grades on exams, and (c) overall experience as student. I will cast the learning process in continuous time. As so, conditional on no graduation, the updating of prior $p(t)$ will be an additive function of the three different sources of information. Next I discuss these three different sources and how to construct the learning process.

At the end of period t a student is endowed with prior $p(t)$. Survival as student conveys imperfect information as high-skilled agents are more prone to graduate. For a small interval of time Δ bayesian updating implied by no graduation is given by

$$p(t+\Delta) = \frac{(1-\phi_1^q\Delta)\,p(t)}{(1-\phi_1^q\Delta)\,p(t)+(1-\phi_0^q\Delta)\,(1-p(t))}$$

substracting $p(t)$ from both sides, dividing by Δ and taking the limit when $\Delta \to 0$ yields $d\hat{p}$, the bayesian updating formula in continuous time due to no graduation,

$$d\hat{p} = (\phi_0^q - \phi_1^q)\,p\,(1-p)\,dt$$

Exams are presented to students every period and consists of two questions. The first question is a simple multiple choice question (from now on, Easy). The degree of difficulty of the question is low. As so, it provides more information when failing to answer correctly to the question. A high-skilled student (i.e. $\mu = 1$) fails the exam with probability zero, while a low-skilled student (i.e. $\mu = 0$) fails the exam with positive probability $\kappa\Delta$ per unit of time. The second question (from now on, Hard) is a question to develop. The complexity of this question is such that low-skilled students fail this question with certainty. High-skilled students fail this question if they didn't manage to study the particular topic before the exam. High-skilled students fail the question with probability $1-\lambda\Delta$ per unit of time (see Table 7). There are three possible outcomes on each exam: (1) answer incorrectly both questions and so fail (F), (2) answer correctly the Easy question and incorrectly the Hard question and so

	Easy	Hard
Low Skilled ($\rho = 0$)	$\kappa\Delta$	1
High Skilled ($\rho = 1$)	0	$1 - \lambda\Delta$

Table 7: **Exams in College.** Probability of answering incorrectly each question.

pass (P)[11], and (3) answering correctly both question and so excel in the exam (E). Table 7 depicts this information.

Conditional on survival as student, a P on an exams conveys imperfect information and thus requires bayesian updating. For a small interval of time Δ bayesian updating implied by a grade P is

$$p(t+\Delta) = \frac{(1-\lambda^q\Delta)\, p(t)}{(1-\lambda^q\Delta)\, p(t) + (1-\kappa^q\Delta)(1-p(t))}$$

substracting $p(t)$ from both sides, dividing by Δ and taking the limit when $\Delta \to 0$ yields $d\tilde{p}$, the bayesian updating formula in continuous time due to a P in an exam,

$$d\tilde{p} = (\kappa^q - \lambda^q)\, p(1-p)\, dt$$

Conditional on remaining a student and no fully informative signals arriving, the overall experience as a student produces information (problem sets, classes, etc.). Let E_t denote the stock of student experience,

$$E_t = \theta_\mu^q t + \sigma Z_t$$

where Z_t is a Wiener process and θ_μ^q is type and institution specific, with $\theta_1^q > \theta_0^q$. Standard techniques imply

$$dE_t = \theta_\mu^q dt + \sigma dZ_t$$

with $dZ_t \sim N\left(0, \sqrt{dt}\right)$. As the skill-level of the agent is not observable, the true value of θ_μ^q for a particular student is unknown and thus has to be infered.

The expected growth rate of the stock of experience is given by $p\theta_1^q + (1-p)\,\theta_0^q$. For a small interval Δ, the unexpected growth in the stock of grades is $E_{t+\Delta} - E_t - (p\theta_1^q + (1-p)\,\theta_0^q)\,\Delta$.

[11] the event 'answer incorrectly the Easy question and correctly the Hard question' is not possible.

After scaling by $\sigma\sqrt{t}$ and taking the limit as $\Delta \to 0$ yields a new process $d\bar{E}_t$ with

$$d\bar{E}_t = \frac{dE_t - \left(p\theta_1^q + (1-p)\,\theta_0^q\right)}{\sigma}$$

Lipster and Shiryayev 1977 [21] provides the updating of beliefs implies by changes in the stock of experience, $d\bar{p}$,

$$d\bar{p} = \frac{\left(\theta_1^q - \theta_0^q\right) d\bar{E}}{\sigma} p\,(1-p)$$

Now I turn to construct the full learning process dp. Conditional on no graduation the learning process is simply $dp = d\hat{p} + d\tilde{p} + d\bar{p}$, or

$$dp = p\,(1-p)\left[\left(\phi_0^q - \phi_1^q + \kappa^q - \lambda^q\right) dt + \frac{\left(\theta_1^q - \theta_0^q\right) d\bar{E}_t}{\sigma}\right]$$

Wang 2008 [29] produce a similar process for understanding consumption/saving behavior when income growth is unknown. The difference with this setup is that in his model income growth varies between two values and that there is no exogenous exit rates (here I have the graduation rate). Moscarini 2003 [24] uses a similar configuration of prior updating to test the quality of a match between a worker and a firm when a component of productivity is unobservable and unknown. In Moscarini's setup, $d\hat{p} + d\tilde{p} = 0$ and so dp would be a martingale.

5.2 The Problem of a Student

I divide this section in two parts. In the first one I setup the problem of a student when dropout and transfer options are not available. Then, I extend the problem to allow for dropout/transfer behavior.

5.2.1 No dropout and Transfer Options

Let $W^N(x,p,q,0)$ denote the value for a student enrolled at institution q with current wealth x, prior p and transfer status $TS = 0$. This value solves the following program:

$$
\begin{aligned}
rW^N(x,p,q,0) \quad = \quad & \max_c \frac{e^{-\gamma c}}{-\gamma} + p\phi_1^q \left(V(x;\bar{w}_1^q) - W^N(x,p,q,0)\right) \\
& +(1-p)\phi_0^q \left(V(x;\bar{w}_0^q) - W^N(x,p,q,0)\right) \\
& +p\lambda^q \left(W^N(x,1,q,0) - W^N(x,p,q,0)\right) \\
& +(1-p)\kappa^q \left(W^N(x,0,q,0) - W^N(x,p,q,0)\right) \\
& +W_x^N(x,p,q,0)\frac{dx}{dt} + W_p^N(x,p,q,0)\frac{dp^e}{dt} \\
& +W_{pp}^N(x,p,q,0)\,\Sigma(p)
\end{aligned} \tag{5}
$$

with

$$
\begin{cases}
\frac{dx}{dt} = rx - a^q - c \\
\frac{dp^e}{dt} = \left(\phi_0^q - \phi_1^q + \kappa^q - \lambda^q\right) p\,(1-p) \\
\Sigma(p) = \frac{1}{2\sigma^2}\left(\theta_1^q - \theta_0^q\right)^2 p^2 (1-p)^2
\end{cases}
$$

The first order condition with respect to consumption states that

$$
e^{-\gamma c} = W_p^N(x,p,q,0)
$$

Evaluating problem 5 at the optimum suggested by the first order condition and substituting the value function W^N with the conjecture of solution $W^N(x,p,q,0) = \frac{e^{-\gamma\left(rx + f^N(p,q,0)\right)}}{-\gamma r}$

yields

$$
\begin{aligned}
&(p(\phi_1^q+\lambda^q)+(1-p)\phi_0^q+(1-p)\kappa^q-\gamma r(a^q+f^N(p,q,0)) \\
&+\gamma f_p^N(p,q,0)\left(\phi_0^q-\phi_1^q+\kappa^q-\lambda^q\right)p(1-p))e^{-\gamma f^N(p,q,0)} \\
&= \\
&p\phi_1^q e^{-\gamma\bar{w}_1^q}+(1-p)\left(\phi_0^q+\kappa^q A_0^q\right)e^{-\gamma\bar{w}_0^q}+p\lambda^q\left(A_1^q e^{-\gamma\overline{w}_1^q}\right) \\
&-\gamma\Sigma(p)\left(f_{pp}^N(p,q,0)-\gamma\left[f_p^N(p,q,0)\right]^2\right)e^{-\gamma f^N(p,q,0)}
\end{aligned}
$$

Note that if this Ordinary Differential equation of second order is evaluated at $p=0$ or $p=1$, the equation provides the same implicit relation that A_μ^q needs to satisfy ($f^N(\mu,q,0)=\overline{w}_\mu^q-\frac{1}{\gamma}\ln A_\mu^q$).

5.2.2 Transfer and Dropout Options are available

Agents are only allowed to transfer once. As so, once they already transferred ($TS=1$) the dropout option is the only one that remains. Let $p_d^{q,1}$ denote the threshold such that students at institution q with priors lower than $p_d^{q,1}$ dropout and join workforce. At this stage (i.e. $TS=1$) the inaction zone is $\left[p_d^{q,1},1\right]$.

Decision rules when $TS=0$ are slightly more difficult and differ for J and C. Students enrolled in Junior College with very low priors might decide to drop from school and students with high priors possibly would decide to transfer. Let $p_d^{J,0}$ and p_t^J denote the thresholds that define the decision rules. The inaction zone for students enrolled in Junior College with $TS=0$ is then $\left[p_d^{J,0},p_t^J\right]$. Students enrolled in College with very low priors might decide to drop and students with slightly higher priors might transfer to Junior College. Let $p_d^{C,0}$ and p_t^C denote the thresholds that define the decision rules. The inaction zone for students enrolled in Junior College with $TS=0$ is then $\left[\max\left(p_d^{J,0},p_t^J\right),1\right]$.

In the inaction zone the problem faced by a student currently enrolled at institution q is

$$
\begin{aligned}
rW(x,p,q,0) \quad = \quad & \max_{c} \frac{e^{-\gamma c}}{-\gamma} + p\phi_1^q \left(V(x;\bar{w}_1^q) - W(x,p,q,0)\right) \\
& +(1-p)\phi_0^q \left(V(x;\bar{w}_0^q) - W(x,p,q,0)\right) \\
& +p\lambda^q \left(W(x,1,q,0) - W(x,p,q,0)\right) \\
& +(1-p)\kappa^q \left(V\left(x;\underline{w}^q\right) - W(x,p,q,0)\right) \\
& +W_x(x,p,q,0)\frac{dx}{dt} + W_p(x,p,q,0)\frac{dp^e}{dt} \\
& +W_{pp}\left(x,p,q,0\right)\Sigma\left(p\right)
\end{aligned}
\tag{6}
$$

with

$$
\begin{cases}
\frac{dx}{dt} = rx - a^q - c \\
\frac{dp^e}{dt} = \left(\phi_0^q - \phi_1^q + \kappa^q - \lambda^q\right) p\left(1-p\right) \\
\Sigma\left(p\right) = \frac{1}{2\sigma^2}\left(\theta_1^q - \theta_0^q\right)^2 p^2\left(1-p\right)^2
\end{cases}
$$

Evaluating problem 6 at the optimum suggested by the first order condition and substituting the value function W with the conjecture of solution $W\left(x,p,q,0\right) = \frac{e^{-\gamma(rx+f(p,q,0))}}{-\gamma r}$ yields

$$
\begin{gathered}
(p(\phi_1^q + \lambda^q) + (1-p)\phi_0^q + (1-p)\kappa^q - \gamma r(a^q + f(p,q,0)) \\
+\gamma f_p(p,q,0)\left(\phi_0^q - \phi_1^q + \kappa^q - \lambda^q\right) p(1-p))e^{-\gamma f(p,q,0)} \\
= \\
p\phi_1^q e^{-\gamma\bar{w}_1^q} + (1-p)\phi_0^q e^{-\gamma\bar{w}_0^q} + (1-p)\,\kappa^q e^{-\gamma\underline{w}^q} + p\lambda^q\left(A_1^q e^{-\gamma\bar{w}_1^q}\right) \\
-\gamma\Sigma\left(p\right)\left(f_{pp}^N\left(p,q,0\right) - \gamma\left[f_p^N\left(p,q,0\right)\right]^2\right)e^{-\gamma f(p,q,0)}
\end{gathered}
$$

5.3 Differences in learning between College and Junior College

College and Junior College not only differ in graduation rates, returns to education and tuition. Learning environments are different and as so they should provide different streams of information. Junior College classes are usually easier so grades in exams should only provide information for the lower tail of the talent distribution. This suggests that $\lambda^J < \lambda^C$. In particular, I am going to set $0 = \lambda^J < \lambda^C = \lambda$. Further, set $\kappa^J = \kappa^C = \kappa$. Finally, College students are generally more full-time (add quote) and so devote more time to study and class preparations. I understand this fact as $\theta_1^C - \theta_0^C > \theta_1^J - \theta_0^J$.

Assumption 6. *The following restriction on parameters holds,*

$$\phi_0^C - \phi_1^C + \kappa - \lambda < 0 < \phi_0^J - \phi_1^J + \kappa$$

Assumption 6 states that the expected change in priors is positive in Junior College and negative in College (conditional on no fully revealing signals arriving).

5.4 Boundary Conditions and Thresholds

There are three different sets of thresholds: dropout thresholds (denoted by $p^q_{d,TS}$), transfer thresholds (denoted by p^q_t), and enrollment thresholds (denoted by p^q_e and p^S_e). In this section I will discuss the different boundary conditions and how to obtain the different thresholds.

Assumption 7. *From now on*

$$\theta_1^C - \theta_0^C = \theta_1^J - \theta_0^J = 0$$

5.4.1 Dropout Thresholds

At the threshold $p^q_{d,TS}$ the value of enrollment and the value of working are equalized, producing the Value Matching Condition

$$W(x - rE^q, p^q_{d,TS}, q, TS) = V\left(x; \underline{w}^q\right)$$

Substituting the maximized value function for $W(x - rE^q, p^q_{d,TS}, q, TS)$ and $V\left(x; \underline{w}^q\right)$ produces

$$f(p^q_{d,TS}, q, TS) - rE^q = \underline{w}^q$$

which shows that the thresholds are independent of current wealth x. The optimality of the boundary $p^q_{d,TS}$ implies the Smooth Pasting Condition,

$$\frac{df(p^q_{d,TS}, q, TS)}{dp} = 0$$

Evaluating the ODEs for J and C (equation 6) at these values provides an expression for the thresholds,

$$p^J_{d,TS} = \frac{\gamma r\left(a^J + \underline{w}^J\right) - \phi^J_0\left(1 - e^{-\gamma\left(\overline{w}^J_0 - \underline{w}^J\right)}\right)}{\phi^J_1\left(1 - e^{-\gamma\left(\overline{w}^J_1 - \underline{w}^J\right)}\right) - \phi^J_0\left(1 - e^{-\gamma\left(\overline{w}^J_0 - \underline{w}^J\right)}\right)}$$

and

$$p^C_{d,TS} = \frac{\gamma r(a^C + \underline{w}^C) - \phi^C_0\left(1 - e^{-\gamma\left(\bar{w}^C_0 - \underline{w}^C\right)}\right)}{\phi^C_1\left(1 - e^{-\gamma\left(\bar{w}^C_1 - \underline{w}^C\right)}\right) + \lambda\left(1 - A^C_1 e^{-\gamma\left(\overline{w}^C_1 - \underline{w}^C\right)}\right) - \phi^C_0\left(1 - e^{-\gamma\left(\bar{w}^C_0 - \underline{w}^C\right)}\right)}$$

5.4.2 Transfer Thresholds

A student that transfer from institution q to $\neg q$ does it because it enjoys a higher value after the transfer. The indifference condition is

$$W(x, p^q_t, q, 0) = W(x, p^q_t, \neg q, 1)$$

which can be reduced to

$$f(p^q_t, q, 0) = f(p^q_t, \neg q, 1)$$

and also,

$$\frac{df(p^q_t, q, 0)}{dp} = \frac{df(p^q_t, \neg q, 1)}{dp}$$

5.4.3 Enrollment Thresholds

At the threshold p^q_e the value of enrollment and the value of working are equalized,

$$W(x - rE^q, p^q_e, q, 0) = V(x; w)$$

which can be reduced to

$$f(p_e^q, q, 0) - rE^q = w$$

As in the case with no learning, $p_e^S > p_e^C$ if $p_e^C > p_e^J$.

5.5 Solution of Model

Under some regularities conditions (among other, $\kappa + \phi_0^J - \phi_1^J > 0$ and $\kappa - \lambda + \phi_0^C - \phi_1^C < 0$), $p_e^J < p_e^C$, $p_t^C > p_{d,0}^C$. Then, the initial enrollment pattern will be given by

$$\begin{cases} Workforce & if\, p \in \left[0, p_e^J\right] \\ Junior\, College & if\, p \in \left(p_e^J, p_e^S\right) \\ College & if\, p \in \left[p_e^S, 1\right] \end{cases} \tag{7}$$

As the different value functions $W\left(x, f\left(\bullet\right)\right)$ are monotonous increasing functions on the corresponding $f\left(\bullet\right)$, a sufficient statistic for the optimal choice is current consumption net of interest proceeds (i.e. $f\left(\bullet\right) - rx$). Figure 4 depicts the sorting.

5.6 Value Added of Education

The value added by enrollment in insitution q, Σ^q, solves

$$W\left(x - \Sigma^q, p, q, 0\right) = V\left(x; w\right)$$

where Σ can also be interpreted as the maximum amount a high-school graduate with prior p is willing to pay for having the enrollment option in institution q. Solving the above equation yields,

$$\Sigma^q\left(p\right) = \frac{f\left(p, q, 0\right) - w}{r}$$

with $\frac{d\Sigma^q}{dp} > 0$ as $f\left(p, q, 0\right)$ increasing in p. Further, the return to enrollment at institution q is given by

$$R^q\left(p\right) = \frac{\Sigma^q\left(p\right)}{w/r} = \frac{f\left(p, q, 0\right) - w}{w}$$

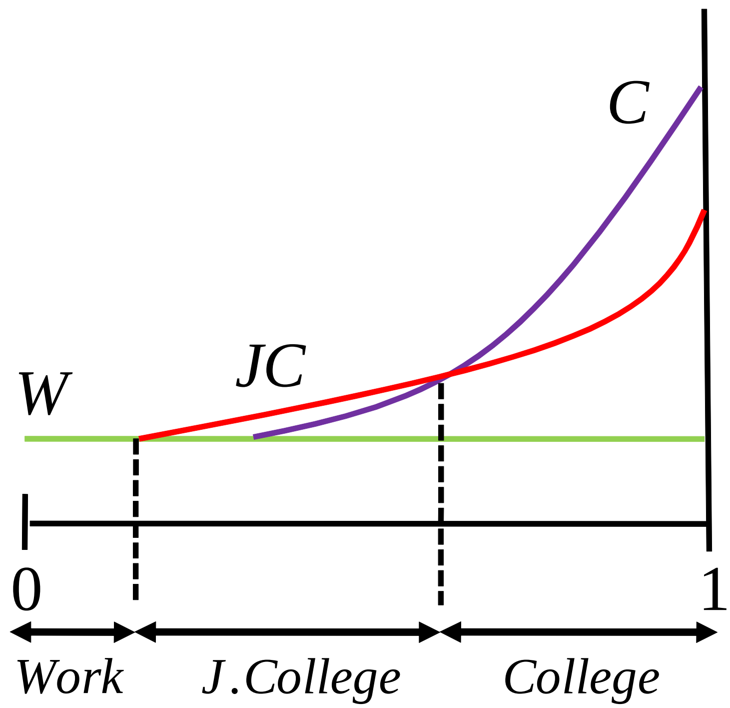

Figure 4: **Sorting on initial enrollment**

5.7 Lack of Insurance and Effect on Enrollment

Market incompleteness arises in this model not due to lack of access to credit but due to the inexistence of available assets to hedge the risk that follows from uncertainty from agent's true skill level and the different income profiles that follow from College or Junior College graduation. The inexistence of assets able to hedge the risk in the model implies lack of insurance and thus provides arguments for goverment intervention. This contrasts with many studies (add cites) that discusses goverment intervention through the implementation of subsidies in setups with complete markets and thus risk-neutral agents. In these models, due to market completeness the allocation can not be improved on.

Next I will show that the total enrollment in post-secondary institutions decreases when risk aversion γ increases. I will not be able to prove the statement for the general configuration of the model. Still, I will argue that the assumptions needed are not restrictive and thus the result should hold always.

Proposition 8. *If* $w + rE = \underline{w}^J$, $p_e^J = p_{d,TS}^J$.

Proof. $p^J_{d,TS}$ is the lowest value for the prior p for which an agent remains a student at Junior College. For $p < p^J_{d,TS}$ the agent is indifferent between remaining a student and dropping out. If $w + rE = \underline{w}^J$, the return for dropouts equals the foregone earnings of becoming a student. This fact implies that $p^J_e = p^J_{d,TS}$. □

The last proposition states that when the foregone earning due to enrollment in Junior College ($w+rE$) equals the returns for Junior College dropouts ($\underline{w}^J$), dropout and enrollment behavior are similar.

Under the restriction $w + rE = \underline{w}^J$,

$$\frac{dp^J_e}{d\gamma} = \frac{r\left(a^J + \underline{w}^J\right) - \phi^J_0\left(\overline{w}^J_0 - \underline{w}^J\right)e^{-\gamma\left(\overline{w}^J_0 - \underline{w}^J\right)}\left(1 - p^J_e\right) - p^J_e\phi^J_1\left(\overline{w}^J_1 - \underline{w}^J\right)e^{-\gamma\left(\overline{w}^J_1 - \underline{w}^J\right)}}{\phi^J_1\left(1 - e^{-\gamma\left(\overline{w}^J_1 - \underline{w}^J\right)}\right) - \phi^J_0\left(1 - e^{-\gamma\left(\overline{w}^J_0 - \underline{w}^J\right)}\right)} \tag{8}$$

Proposition 9. *If $w + rE = \underline{w}^J$ and $p^J_e \in (0, 1)$, $\frac{dp^J_e}{d\gamma} > 0$.*

Proof. I will build up the proof in steps.

1) $p^J_e \in (0, 1)$ iff

$$\begin{aligned} \Gamma_0(\gamma) &> 0 \\ \Gamma_1(\gamma) &< 0 \end{aligned}$$

where $\Gamma_\mu(\gamma) = \gamma r\left(a^J + \underline{w}^J\right) - \phi^J_\mu\left(1 - e^{-\gamma\left(w^J_\mu - \underline{w}^J\right)}\right)$, for $\mu \in \{0, 1\}$. The first and second derivatives of $\Gamma_\mu(\gamma)$ is given by

$$\begin{aligned} \Gamma'_\mu(\gamma) &= r\left(a^J + \underline{w}^J\right) - \phi^J_\mu\left(w^J_\mu - \underline{w}^J\right)e^{-\gamma\left(w^J_\mu - \underline{w}^J\right)} \\ \Gamma''_\mu(\gamma) &= \phi^J_\mu\left(w^J_\mu - \underline{w}^J\right)^2 e^{-\gamma\left(w^J_\mu - \underline{w}^J\right)} \end{aligned}$$

2) Properties of Γ_0 and Γ_1: $\Gamma_\mu(\gamma)$ strictly convex as $\Gamma''_\mu(\gamma) > 0$. Also, $\Gamma_\mu(0) = 0$, $\lim_{\gamma\to\infty}\Gamma_\mu(\gamma) = \infty$, $\Gamma'_\mu(0) = r\left(a^J + \underline{w}^J\right) - \phi^J_\mu\left(w^J_\mu - \underline{w}^J\right)$. If $\Gamma_1(\gamma) < 0$ holds then it has to be the case (given Γ_1 strictly convex) that $\Gamma'_1(0) < 0$. Finally, note that $\Gamma_0(\gamma) - \Gamma_1(\gamma) \geq 0$.

3) Let $\underline{\gamma} = \inf_\gamma : \Gamma_0(\gamma) > 0$ and $\overline{\gamma} = \sup_\gamma : \Gamma_1(\gamma) < 0$. The properties of the functions

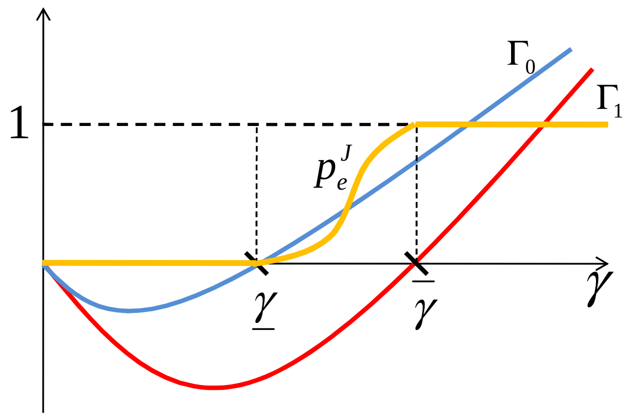

Figure 5: **Graphic Analysis of the proof of Proposition 9**.

$\Gamma_0$and Γ_1 discussed in step 2 implies that $\underline{\gamma} < \overline{\gamma}$. It follows that the set of admissible γ is given by $\left[\underline{\gamma}, \overline{\gamma}\right]$.

4) $\lim_{\gamma \downarrow \underline{\gamma}} \frac{dp_e^J}{d\gamma} > 0$ and $\lim_{\gamma \uparrow \overline{\gamma}} \frac{dp_e^J}{d\gamma} > 0$ as (a) p_e^J, $\frac{dp_e^J}{d\gamma}$ and $\frac{d^2 p_e^J}{d\gamma^2}$ continuous and differentiable functions of γ, and (b) otherwise contradicts the definition of $\underline{\gamma}$ and $\overline{\gamma}$.

5) Evaluating equation 8 at $p_e^J = 0$ and $p_e^J = 1$ yields

$$\left.\frac{dp_e^J}{d\gamma}\right|_{p_e^J=0} = \frac{r\left(a^J + \underline{w}^J\right) - \phi_0^J\left(\overline{w}_0^J - \underline{w}^J\right)e^{-\gamma\left(\overline{w}_0^J - \underline{w}^J\right)}}{\phi_1^J\left(1 - e^{-\gamma\left(\overline{w}_1^J - \underline{w}^J\right)}\right) - \phi_0^J\left(1 - e^{-\gamma\left(\overline{w}_0^J - \underline{w}^J\right)}\right)}$$

$$\left.\frac{dp_e^J}{d\gamma}\right|_{p_e^J=1} = \frac{r\left(a^J + \underline{w}^J\right) - \phi_1^J\left(\overline{w}_1^J - \underline{w}^J\right)e^{-\gamma\left(\overline{w}_1^J - \underline{w}^J\right)}}{\phi_1^J\left(1 - e^{-\gamma\left(\overline{w}_1^J - \underline{w}^J\right)}\right) - \phi_0^J\left(1 - e^{-\gamma\left(\overline{w}_0^J - \underline{w}^J\right)}\right)}$$

As p_e^J continuous and differentiable function of γ, $\left.\frac{dp_e^J}{d\gamma}\right|_{p_e^J=0} = \lim_{\gamma \downarrow \underline{\gamma}} \frac{dp_e^J}{d\gamma} > 0$ and $\left.\frac{dp_e^J}{d\gamma}\right|_{p_e^J=1} = \lim_{\gamma \uparrow \overline{\gamma}} \frac{dp_e^J}{d\gamma} > 0$.

6) Rewrite the derivative of p_e^J with respect to γ (equation 8) as

$$\frac{dp_e^J}{d\gamma} = p\left.\frac{dp_e^J}{d\gamma}\right|_{p_e^J=1} + (1-p)\left.\frac{dp_e^J}{d\gamma}\right|_{p_e^J=0} > 0$$

which completes the proof. □

Figure 5 presents the graphic analysis of the proof.

6 Model Predictions

I present a set of testable model predictions and how to link them to the data.

6.1 Sorting in Initial Enrollment with respect to Observables

High-school graduates with low priors join workforce, average priors enroll in Junior College and high priors enroll in College (see 7). p measures the belief of the agent of being high-skilled. If p provides sensible information about the agent's true skill level it has to be the case that $cov(p, \rho) > 0$. A natural approach is to link p with observed ability measures of an individual, Y, and unobserved measure, ϵ. In general terms, define ϑ as an operator that maps ability measures to the space of probabilities such that the prior for a given agent with measures of ability Y_i and ϵ_i is just $p_i = \varphi(Y_i, \epsilon_i)$. The next definition presents ϑ.

Definition 10. Given Y, vector of observable measures of ability and a measure of unobserved ability ϵ, let $\vartheta(Y, \epsilon)$ be an operator that maps measures of ability into probabilities,

$$\vartheta(Y, \epsilon) : \mathbb{R}^N \times \mathbb{R} \longrightarrow [0, 1]$$

Further, assume

$$\vartheta(Y, \epsilon) = \psi\left(Y'\beta + \epsilon\right)$$

with ψ invertible and increasing.

As ϵ is not observable I need to impose some structure on it. Let $\epsilon \sim N(0, 1)$.

A high-school graduate will enroll in College iff $p \geq p_e^S$ or $\psi(Y'\beta + \epsilon) \geq p_e^S$. As ψ increasing and invertible it is possible to rewrite this condition as

$$\epsilon \geq \psi^{-1}\left(p_e^S\right) - Y'\beta$$

which, given the normality of ϵ provides that the probability of enrollment at College is

$$\Pr(C) = 1 - \Phi\left(\psi^{-1}\left(p_e^S\right) - Y'\beta\right)$$

Following similar arguments,

$$\Pr(J) = \Phi\left(\psi^{-1}\left(p_e^J\right) - Y'\beta\right) - \Phi\left(\psi^{-1}\left(p_e^S\right) - Y'\beta\right)$$

and

$$\Pr(work) = \Phi\left(\psi^{-1}\left(p_e^J\right) - Y'\beta\right)$$

These calculations suggest that I can use an ordered probit regression to estimate β, $\psi^{-1}\left(p_e^S\right)$ and $\psi^{-1}\left(p_e^J\right)$. Let ORD_P=0 if agent joins workforce, =1 if enrolled in Junior College and =2 if enrolled in College. Table 8 shows the results of the ordered probit estimation. Every coefficient is significant and has the correct sign.

6.2 Probability of Dropout, Graduation and Transfer for Students Initially Enrolled in College

Let $\Pr(G|C;p)$ denote the unconditional graduation probability at College for students initially enrolled in College with prior p. Also, let $\Pr(G|C;p;\mu)$ denote the same object but for skill level μ. $\Pr(G|C;p)$ equals

$$\Pr(G|C;p) = p\Pr(G|C;p;1) + (1-p)\Pr(G|C;p;0)$$

As p is a monotonic increasing function on the aggregate measure of observed ability $Y'\beta$ ($\frac{dp}{d(Y'\beta)} = \psi'(Y'\beta+\epsilon) > 0$), the rate of change of this probability on observables is

$$\frac{d\Pr(G|C;p)}{dY'\beta} = \psi'\left(Y'\beta+\epsilon\right)\left(\Pr(G|C;p;1) - \Pr(G|C;p;0) + p\frac{d\Pr(G|C;p;1)}{dp} + (1-p)\frac{d\Pr(G|C;p;0)}{dp}\right) > 0$$

as $\Pr(G|C;p;1) - \Pr(G|C;p;0) > 0$ (because $\phi_1^C + \lambda > \phi_0^C$ and conditional on no fully revealing shock arriving the transfer probability is independent of the type of the agent) and $\frac{d\Pr(G|C;p;\mu)}{dp} > 0$ (see Appendix I.2). Then, the *unconditional graduation probability is increasing in in the observable measure of ability* $Y'\beta$ (Prediction 1-C).

Conditional on no graduation or fully informative signals arriving, the probability of reaching the transfer threshold is independent of the student's true skill level. The hazard rate (probability per unit of time of leaving this state) for a high-skilled student is $\phi_1^C + \lambda$ and for low-skilled students is $\phi_0^C + \kappa$. As $\phi_1^C + \lambda > \phi_0^C + \kappa$, the chance of survival on this state is decreasing on the agent's type. Then, the *unconditional transfer probability is decreasing in the observable measure of ability* $Y'\beta$ (Prediction 2-C).

Let $\Pr(D|C;p)$ denote the unconditional dropout probability at College for students initially enrolled in College. Also let $\Pr(D|C;p;\mu)$ be the conditional dropout probability for students with prior p and skill level μ. For a student with prior p^i this value equals

$$\Pr(D|C;p) = p\Pr(D|C;p;1) + (1-p)\Pr(D|C;p;0)$$

Given $p_t^C > p_d^{C,0}$ the only way a student will become a dropout is if the fully informative shock κ arrives, that only happens if true type is 0. Then, $\Pr(D|C;p;1) = 0$. It follows that

$$\frac{d\Pr(D|C;p)}{d(Y'\beta)} = \frac{dp}{d(Y'\beta)}\left((1-p)\frac{d\Pr(D|C;p;0)}{dp} - \Pr(D|C;p;0)\right)$$

where $\frac{dp}{d(Y'\beta)} = \psi'(Y'\beta+\epsilon) > 0$. Conditional on $\mu = 0$, $\frac{d\Pr(D|C;p;0)}{dp} > 0$ as there is more time to receive the fully revealing shock κ and thus inducing dropout behavior. Also, $\Pr(D|C;p;0) > 0$, and thus the sign of $\frac{d\Pr(D|C;p;0)}{d(Y'\beta)}$ is unknown.

It is the case that $\frac{d\Pr(D|C;0)}{d(Y'\beta)} = \frac{\psi'(Y'\beta+\epsilon)\kappa}{(\phi_1^C+\lambda)-(\phi_0^C+\kappa)}\frac{1}{p_t^C} > 0$ and $\frac{d\Pr(D|C;1)}{d(Y'\beta)} = \frac{\psi'(Y'\beta+\epsilon)\kappa}{-(\phi_0^C+\kappa)} < 0$. As $\frac{d\Pr(D|C;p)}{d(Y'\beta)}$ continuous on p, there is a region for p such that higher prior increases the dropout probability and other where it decreases this probability. There are two effects that push in opposite directions. First, conditional on true type being 0, higher p increases the dropout likelihood ("time effect"). Second, higher p signals high-skill and thus lower dropout probability ("type effect"). In the first region the time effect dominates. In the second region the type effect dominates. The average for the population enrolled in college is proportional to

$$\psi'\left(Y'\beta+\epsilon\right)\int_{p_t^C}^{1}\frac{d\Pr(D|C;p)}{dp}F(dp)$$

ORD_P	β	Std. Error
men	0.205	0.046
black	0.402	0.077
Father_LHS	-0.747	0.102
Father_HS	-0.472	0.094
Father_DR	-0.349	0.09
Socio_LOW	-0.926	0.097
Socio_MED	-0.625	0.078
Rank	-1.477	0.081
R^2	0.1508	
/cut1	-1.37	0.071
/cut2	-0.826	0.073

Table 8: **Estimation of initial prior p_0.** Ordered probit regression where ORD_P=0 if join workforce, =1 if enrolled in Junior College and =2 if enrolled in College. Father_LHS: if father didn't finish high-school. Father_HS: maximum education level is graduating from high-school. Father_DR: if maximum education level was completing some courses in College. Socio_LOW: if socioeconomic status of family is low. Socio_MED: if socioeconomic status is medium. Rank: ratio of rank in class to total class population.

	C	J
$\Pr(G)$	0.124 (0.025)	0.015 (0.004)
$\Pr(T)$	-0.001 (0.0004)	0.256 (0.034)
$\Pr(D)$	-0.123 (0.025)	-0.271 (0.036)

Table 9: **Marginal Effects.**

where $F(p)$ denotes the distribution of priors. If this value is positive the time effect dominates. If negative the opposite happens.

Then, the *unconditional dropout probability is* $\begin{matrix} \textit{decreasing} \\ \textit{increasing} \end{matrix}$ *in the observable measure of ability* $Y'\beta$ *if* $\int_{p_t^C}^{1} \frac{d\Pr(D|C;p)}{dp} F(dp) \begin{matrix} <0 \\ >0 \end{matrix}$. (Prediction 3-C)

To test the predictions I estimated how each probability changes with the measure of ability $Y'\beta$. Table 9 presents the results of the estimation. The estimates have the correct sign and are statistically different from zero.

6.3 Probability of Dropout, Graduation and Transfer for Students Initially Enrolled in Junior College

The graduation hazard rate for high-skilled students is higher to the one of low-skilled students. Then, $\Pr(G|J;p;1) > \Pr(G|J;p;0)$. In Appendix H.2 I show that $\frac{d\Pr(G|J;p;\mu)}{dp} < 0$, implying that

$$\frac{d\Pr(G|J;p)}{dY'\beta} = \psi'\left(Y'\beta + \epsilon\right)\left(\Pr(G|J;p;1) - \Pr(G|J;p;0) + p\frac{d\Pr(G|J;p;1)}{dp} + (1-p)\frac{d\Pr(G|J;p;0)}{dp}\right)$$

can't be signed.

Conditional on no graduation or fully informative signals, the probability of reaching the transfer thresholds is independent of the student's true skill level. As $\kappa + \phi_0^J - \phi_1^J > 0$, the *unconditional transfer probability is increasing in the observable measure of ability $Y'\beta$.* (Prediction 1-J).

The dropout argument is similar to the one for College. The only difference is that there is no signal that fully reveals high ability. Still, low-skilled agents can receive a signal that fully reveals their type and thus induce dropout. The *unconditional dropout probability is decreasing in the observable measure of ability $Y'\beta$.* (Prediction 3-J). Note that the unconditional dropout probability can be written as

$$\Pr(D|J;p) = p\Pr(D|J;p;1) + (1-p)\Pr(D|J;p;0) = (1-p)\Pr(D|J;p;0)$$

which is decreasing in p (see Appendix H.1).

Table 9 presents the results of the estimation. The estimates of interest (first and third row) have the correct sign and are statistically different from zero.

6.4 Proportion that dropout (from the original cohort) in institution q if initially enrolled at institution q decreases through time

First, the probability of hitting dropout thresholds is zero. Next, note that the only students that might become dropouts before transferring are low-skilled students. Further, note that at every period some students graduate and others transfer as they hit the transfer threshold. It follows that the pool of students that can potentially become dropouts shrinks through time.

7 Simulation of Model

In order to simulate the model first I need to solve it completely, which is done by numerical methods. An easy approach is, similar to the one used to characterize the tresholds, to start by solving the problem of students that already transferred ($TS = 1$) and then solve the problem of students at time 0. At any stage the functions $f()$ need to satisfy the corresponding ordinary differential equations and the value matching and smooth pasting conditions. Instead of discretizing the differential equations I will follow a different approach. In a nutshell, I will combine a collocation approach (evaluating at Chebyshev's nodes) with Value Matching and Smooth Pasting Conditions. A treatment of Collocation methods and Chebyshev's nodes can be found in Numerical Methods in Economics, by K. Judd [14]. In Appendix J I provide an example of this approach.

I solve the model for the following set of parameters chosen to roughly match characteristics of the data:

- Wage parameters: $w = 1$, $\underline{w}^J = 1$, $\underline{w}^C = 1$, $\hat{w}_0^J = 1.055$, $\hat{w}_1^J = 1.12$, $\hat{w}_0^C = 1.07$, $\hat{w}_1^C = 1.35$.
- Graduation parameters: $\phi_0^J = 0.4$, $\phi_1^J = 0.5556$, $\phi_0^C = 0.1538$, $\phi_1^C = 0.2632$.
- Learning parameters: $\lambda = 0.35$, $\kappa = 0.3$.
- Cost parameters: $E = 0$, $a^J = 0.1$, $a^C = 0.26$.[12]
- other: $\gamma = 4$, $r = 0.03$.

The optimal tresholds for junior college are

$$p_d^{J,1} = p_e^J = 0.3991$$

[12] I discarded from the schooling cost room and board.

and

$$p_t^J = 0.6527$$

For college I obtain

$$p_d^{C,1} = 0.4032$$

and

$$p_t^C = 0.4303$$

Finally,

$$p_e^S = 0.6516$$

Figure 6 shows how the different alternative relate to Value Matching and Smooth Pasting (and its relation with the tresholds). Figure 7 computes and compares the model with and without learning. Note that: (1) the learning channel that operates through the dropout and transfer options has important value, (2) value is higher with learning ,and (3) learning greatly affect the initial enrollment pattern.

Add:

1) Explain computational method to solve the model (in a nutshell: Collocation Methods at Chebyshev's nodes)

2) Simulate the model for an arbitrary set of parameters.

3) Compare 3 models: Full Model (with Learning + uncertainty) - Model with uncertainty - Model with certainty

4) Discuss possible bias in estimates of returns if Junior College not included.

5) Show some nice plots and compare ex-ante vs. ex-post returns

8 Proposed method for model estimation

Even though the model as it is will be hard to be estimated (hard to handle model) I propose a way of estimating the model.

1) Say that there are 6 parameters to be estimated (the others we calibrate them from

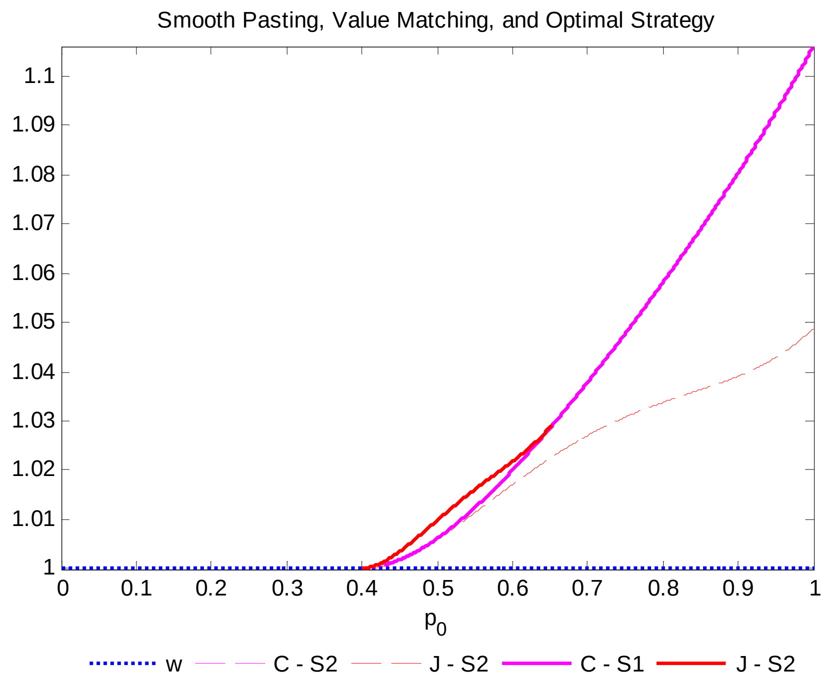

Figure 6: **Smooth Pasting and Value Matching of different alternatives**.

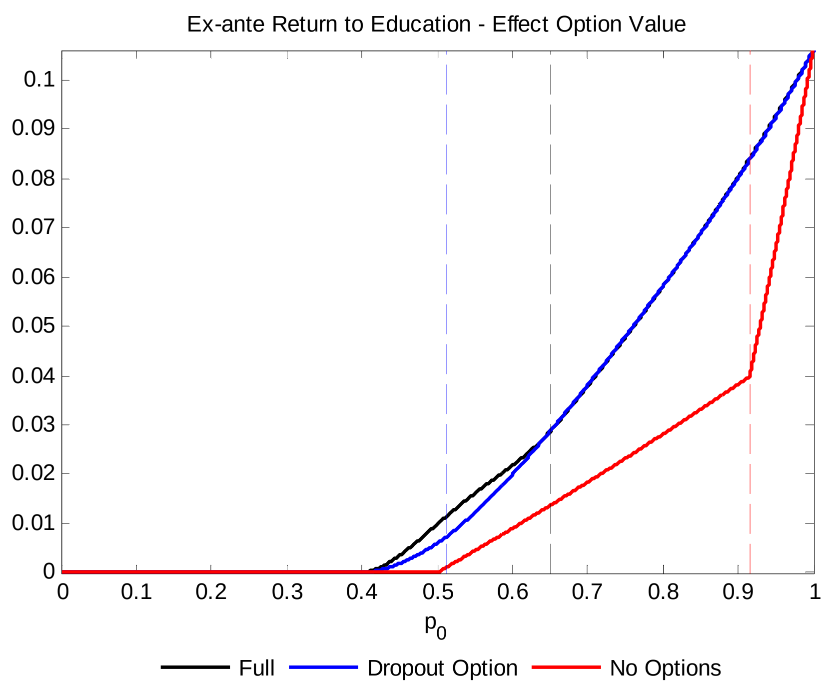

Figure 7: **Returns and Enrollment Pattern. Effect of learning and option values**.

data).

2) It is possible to compute by hand the probability of dropout, transfer and expected time until graduation for students in College and Junior College for any given skill level.

3) We can then use these objects to compute ex-ante unconditional probabilities and expected time until graduation.

4.1) In this was we constructed six moments as function of the 6 unknown parameters. We can attempt to match them to the moments gathered from data using Simulated Method of Moments.

References

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

[2] Belley, P. and Lochner, L.: *The Changing Role of Family Income and Ability in Determining Educational Achievement*. Journal of Human Capital, Vol. 1, No. 1, 2008.

[3] Betts, J. R. and McFarland, L. L.: *Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments*. Journal of Human Resources, Vol. 30, No. 4, 1995.

[4] Cameron, S. and Heckman, J. J.: *The Dynamics of Educational Attainment for Blacks, Whites and Hispanics*. Journal of Political Economy, Vol. 109, No. 3, 2001.

[5] Cameron, S. and Taber, C.: *Borrowing Constraints and the Returns to Schooling*. Journal of Political Economy, Vol. 112, 2004.

[6] Carneiro, P, Hansen, K. T. and Heckman, J. J.: *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, 2003.

[7] Comay, Y, Melnik, A. and Pollatschek, M. A.: *The Option Value of Education and the Optimal Path for Investment in Human Capital*. International Economic Review, Vol. 14, No. 2, 1973.

[8] Cunha, F. and Heckman, J. J.: *Indetifying and Estimating the Distribution of Ex Post and Ex Ante Returns to Schooling*. Labour Economics, Vol. 14, No. 6, 2007.

[9] Dixit, A.: *The Art of Smooth Pasting*. Fundamentals of Pure and Applied Economics, No. 55, 1993.

[10] Friedman, M.: *A Theory of Consumption Function*. Princeton University Press, 1957.

[11] Grubb, W. N.: *The Varied Economic Returns to Postsecondary Education: New Evidence from the Class of 1972*. Journal of Human Resources, Vol. 28, No. 2, 1993.

[12] Heckman, J. J. and Navarro, S.: *Dynamic Discrete Choice and Dynamic Treatment Effects*. Journal of Econometrics, 2006.

[13] Heckman, J. J. and Urzua, S.: *The Option Value of Educational Choices and the Rate of Return to Educational Choice*. Cowles Foundation Structural Conference, Yale, 2008.

[14] Judd, K. L.: *Numerical Methods in Economics*. MIT press, 1998.

[15] Judd, K. L.: *Is Education as Good as Gold? A Portfolio Analysis of Human Capital Investment*. unpublished, Hoover Institution, 2000.

[16] Kane, T. J. and Rouse, C. E.: *Labor-Market Returns to Two- and Four-Year College*. American Economic Review, Vol. 85, No. 3, 1995.

[17] Kane, T. J. and Rouse, C. E.: *Comment on W. Norton Grubb: "The Varied Economic Returns to Postsecondary Education: New Evidence from the Class of 1972"*. Journal of Human Resources, Vol. 30, No. 1, 1995.

[18] Kane, T. J. and Rouse, C. E.: *The Community College: Educating Students at the Margin between College and Work*. Journal of Economic Perspectives, Vol. 13, No. 1, 1999.

[19] Keane, M. P. and Wolping, K. I.: *The Career Decisions of Young Men.* Journal of Political Economy, Vol. 105, No. 3, 1997.

[20] Keane, M. P. and Wolping, K. I.: *The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment.* International Economic Review, Vol. 42, No. 4, 2001.

[21] Lipster, R. T., and Shiryayev, A. N.: *Statistics of Random Processes.* Volume I, II, New York: Springer-Verlag.

[22] Lochner, L., and Monge-Naranjo, A.: *Human Capital Formation with Endogenous Credit Constraints: Goverment Student Loan Programs and Private Lenders.* working paper, 2007.

[23] Miao, J., and Wang, N.: *Experimentation under Uninsurable Idiosincratic Risk: An Application to Entrepreneurial Survival.* unpublished, 2007.

[24] Moscarini, G.: *Skill and Luck in the Theory of Turnover.* working paper, Yale University, 2003.

[25] Navarro, S.: *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, 2005.

[26] Ozdagli, A. K. and Trachter, Nicholas: *College Enrollment, Dropouts and Option Value of Education.* Unpublished, University of Chicago, 2008.

[27] Stange, K.: *An Empirical Examination of the Option Value of College Enrollment.* unpublished, University of California, Berkeley, 2007.

[28] Wang, N.: *Generalizing the permanent-income hypothesis: Revisiting Friedman's conjecture on consumption.* Journal of Monetary Economics, 53, 2006.

[29] Wang, N.: *Optimal Consumption and Asset Allocation with Unknown Income Growth.* firthcoming Journal of Monetary Economics, 2009.

Appendix

A Dataset

I am currently working with the National Longitudinal Study of 1972 (NLS-72) that follows a cohort of high-school seniors in spring of 1972 up to 1979 and a final wave in 1986.

For constructing my sample I eliminated individuals with the following charactersitics:

- high-school graduates that join a postsecondary institution after October of 1972.
- individuals enrolled in a postsecondary institution but had at least one period outside the educational sector.
- individuals with missing values for the observable characteristics.
- individuals enrolled in Vocational or Technical schools (2-year colleges).
- individuals with Professional, Master or PhD. degrees.
- individuals that transferred more than once.[13]

As observable characteristics of the agents I picked:

- Location of High School (State)
- Race
- Gender
- Socioeconomic Status of Family (Low, Medium or High)
- Maximum educational level of father (high-school dropout, high-school graduate, college dropout, college graduate)
- rank in high-school class

[13] the amount of individuals that transferred more than once is negligible.

	%
initially enrolled in 4-year College that transferred to V.S.	3.4
initially enrolled in Academic 2-year College that transferred to V.S.	4.6
initially enrolled in V.S. that transferred to 4-year College	4%
initially enrolled in V.S. that transferred to Academic 2-year College	4%

Table 10: **Interaction of Vocational School with Academic 2-year Colleges and 4-year Colleges.**

With respect to the cost of postsecondary education I am using the total cost of attendance (tuition+room+board+etc.). Several missing values are present in the dataset. Instead of eliminating these observations I regressed the cost of attendance on observables in addition of state dummies in order to produce estimates for the missing values.

I also have variables that keep track of type of employment (part-time, full-time, no work) for students.

My final dataset contains 3122 observations.

B Vocational School

2-year colleges are a combination of two distinct types of courses: academic courses and vocational courses. Vocational courses provide specifical training for a particular profession while academic courses are similar to the courses taught in 4-year colleges in the first 2 years. Grubb 1993 [11] observes that wage differentials (relative to high-school graduation) are much higher for students that graduate from vocational school than from 2-year colleges with academic degrees. I performed a similar analysis to get the same relation.

The interaction of Vocational or Technical Schools with Academic 2-year Colleges or 4-year Colleges is small, as it can be seen in table 10.

C Computing Internal Rate of Returns

Let T^{MAX} denote the retirement year of an agent and $Q_0(r;w)$ the value for a worker with wage w since $t=0$ up to retirement,

$$Q_0(r;w) = \int_0^{T^{MAX}} e^{-rt} w dt$$

where $r>0$ denote the risk-free interest rate. Let $Q_S(r;w_{q,T,G},c_q,T,G)$ denote the value for a student for T years, graduation status $G \in \{0,1\}$ in institution q (with flow cost c_q) and wage upon termination $w_{q,T,G}$,

$$Q_S(r;w_{q,T,G},c_q,T,G) = \int_0^{T} e^{-rt} c_q dt + \int_T^{T^{MAX}} e^{-rt} w_{q,T,G} dt$$

The Internal Rate of Return is the interest rate $\hat{r}$ such that

$$Q_S(\hat{r};w_{q,T,G},c_q,T,G) - Q_0(\hat{r};w) = 0$$

For the calculation in this paper I am assuming that an agent retires/dies when she is 75 byears old so that $T^{MAX} = 57$.

D Value at Work

The FOC of problem (1) is given by

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

which implies that in the optimum $\frac{e^{-\gamma c}}{-\gamma} = -\frac{1}{\gamma} V_x(x;\zeta)$ and $c = -\frac{1}{\gamma}\ln V_x(x;\zeta)$. Substituting these two equations back into problem (1) provides

$$rV(x;\zeta) = -\frac{1}{\gamma}V_x(x;\zeta) + V_x(x;\zeta)\left(rx + \zeta + \frac{1}{\gamma}\ln V_x(x;\zeta)\right) \tag{9}$$

Let's guess that the solution to this equation is given by $V(x;\zeta) = -\frac{1}{\gamma r}e^{-\gamma(rx+\zeta)}$. It can be checked that the conjecture solves (9). Finally, note that the FOC states that the optimal consumption rule is given by $c = -\frac{1}{\gamma}\ln V_x(x;\zeta) = rx+\zeta$ as desired.

E Value at institution q (no uncertainty)

The FOC of (3) is

$$e^{-\gamma c} = W_x\left(x,\mu,q,0\right)$$

Conject that $W\left(x,\mu,q,0\right) = -\frac{A^q_\mu}{\gamma r}e^{-\gamma\left(rx+\overline{w}^q_\mu\right)}$. Apply the conjecture together with the FOC and (2) to (3) to get

$$\left(\phi^q_\mu - r\gamma\left(a^q + \overline{w}^q_\mu\right) + r\ln A^q_\mu\right)A^q_\mu = \phi^q_\mu \tag{10}$$

Next I derive some properties of A^q_μ. First note that $\phi^q_\mu - r\gamma\left(a^q+\overline{w}^q_\mu\right) + r\ln A^q_\mu > 0$, or no A^q_μ exists. Define $g(A^q_\mu) \equiv \left(\phi^q_\mu - r\gamma\left(a^q+\overline{w}^q_\mu\right) + r\ln A^q_\mu\right)A^q_\mu - \phi^q_\mu$. We have that $\lim_{A^q_\mu\to 0^+} g(A^q_\mu) = -\phi^q_\mu < 0$, $\lim_{A^q_\mu\to\infty} g(A^q_\mu) = \infty$, $g'(A^q_\mu) = \phi^q_\mu + r - r\gamma\left(a^q+\overline{w}^q_\mu\right) + r\ln A^q_\mu$, and $g''(A^q_\mu) = \frac{r}{A^q_\mu} > 0$ as $A^q_\mu > 0$. These conditions imply that A^q_μ exists, it is unique and $A^q_\mu \in [0,\infty)$. Further, as $g(1) = -r\gamma\left(a^q+\overline{w}^q_\mu\right) < 0$ so that $A^q_\mu > 1$, then $A^q_\mu \in (1,\infty)$.

F Value for a Junior College Student

The FOC of problem (**??**) is

$$e^{-\gamma c} = W_x(x,p,J,TS) \tag{11}$$

applying (11) to (**??**) together with the conjecture of solution $W(x,p,J,TS) = -\frac{1}{\gamma r}e^{-\gamma(rx+f(p,J,TS))}$, and (2) provides

$$(p\phi_1^J + (1-p)\phi_0^J + (1-p)\kappa - \gamma r(a^J + f(p,J,TS)) + \gamma f_p(p,J,TS)\left(\phi_0^J - \phi_1^J + \kappa\right)p(1-p))e^{-\gamma f(p,J,TS)}$$

$$=$$

$$p\phi_1^J e^{-\gamma \bar{w}_1^J} + (1-p)\left(\phi_0^J e^{-\gamma \bar{w}_0^J} + \kappa e^{-\gamma \underline{w}^J}\right)$$

G Value for a College Student

The FOC of problem (**??**) is

$$e^{-\gamma c} = W_x(x,p,C,TS) \tag{12}$$

applying (**??**) to (**??**) together with the conjecture of solution $W(x,p,C,TS) = -\frac{1}{\gamma r}e^{-\gamma(rx+f(p,C,TS))}$, and (2) provides

$$(p(\phi_1^C + \lambda) + (1-p)\phi_0^C + (1-p)\kappa - \gamma r(a^C + f(p,C,TS)) + \gamma f_p(p,C,TS)\left(\phi_0^C - \phi_1^C + \kappa - \lambda\right)p(1-p))e^{-\gamma f(p$$

$$=$$

$$p\phi_1^C e^{-\gamma \bar{w}_1^C} + (1-p)\left(\phi_0^C e^{-\gamma \bar{w}_0^C} + \kappa e^{-\gamma \underline{w}^C}\right) + p\lambda\left(A_1^C e^{-\gamma \overline{w}_1^C}\right)$$

H Probabilities at Junior College

H.1 Dropout probabilities at Junior College

Let the probability for a given student of becoming a dropout during spell at institution q and transfer status TS be given by $D_{\mu}^{q,TS}(p)$, where μ denotes the student's true skill level and p her current belief. I will first obtain the conditions for students currently enrolled in Junior College.

Proposition 11. *As long as $\phi_0^J - \phi_1^J + \kappa \geq 0$, high-skilled students (i.e. $\rho = 1$) will never drop out of Junior College independently of their transfer (TS).*

Proof. First note that as $\phi_0^J - \phi_1^J + \kappa \geq 0$, belief updating is such that $dp/dt \geq 0$. Second, if the student is enrolled at time t in Junior College (relative to be a worker) it has to be the case that $W(x,p,J,TS) > V(x;w)$. Third, as we already established that $dp/dt \geq 0$ and that the optimal policy is independent of wealth x, $W\left(x+\frac{dx}{dt}, p+\frac{dp}{dt}, J, TS\right) > V\left(x+\frac{dx}{dt}; w\right)$. $\square$

Proposition 12. *For low-skilled students at institution J: (1) $D_0^{J,1}(p) = D_0^{J,1}$ so that the probability of dropout is independent of current belief p, (2) $D_0^{J,1} = \frac{\kappa}{\kappa+\phi_0^J}$ and, (3) $D_0^{J,0}(p)$ solves the ODE given by*

$$\left(\phi_0^J+\kappa\right) D_0^{J,0}(p) = \kappa + \frac{dD_0^{J,0}(p)}{dp} p(1-p)\left(\phi_0^J - \phi_1^J + \kappa\right) \tag{13}$$

for $p \leq p_t^J$ with terminal condition $D_0^{J,0}\left(p_t^J\right) = 0$.

(1) as $\phi_0^J - \phi_1^J + \kappa \geq 0$, and given that students in this case already transferred, the only reason for becoming a drop out is by receiving the fully informative signal κ, which is only dependent on the agent's true skill level. Then, $D_0^{J,1}(p) = D_0^{J,1}$. (2) In a discrete time setup with time period given by Δ we have that $D_0^{J,1} = \left(1-\phi_0^J\Delta\right)\left(\kappa\Delta + (1-\kappa\Delta) D_0^{J,1}\right)$. Discarding the terms of order higher than Δ and solving for $D_0^{J,1}$ provides $D_0^{J,1} = \frac{\kappa}{\kappa+\phi_0^J}$. (3) Setting the problem in discrete time provides

$$D_0^{J,0}(p) = \left(1-\phi_0^J\Delta\right)\left(\kappa\Delta + (1-\kappa\Delta) D_0^{J,0}(p+dp)\right)$$

discarding the terms of order higher than Δ and by approximating $D_0^{J,0}(p+dp)$ by $D_0^{J,0}(p) + \frac{dD_0^{J,0}(p)}{dp} dp$ provides

$$\left(\phi_0^J+\kappa\right)\Delta D_0^{J,0}(p) = \kappa\Delta + \left(1-\left(\phi_0^J+\kappa\right)\Delta\right)\frac{dD_0^{J,0}(p)}{dp} dp$$

Taking the limit when $\Delta \to 0^+$,

$$\left(\phi_0^J+\kappa\right) D_0^{J,0}(p) = \kappa + \frac{dD_0^{J,0}(p)}{dp} p(1-p)\left(\phi_0^J - \phi_1^J + \kappa\right)$$

With respect to the terminal condition we have that if the student transfers to College the probability of dropout in Junior College is clearly zero and if they dropout the probability is one.

Proposition 13. *The solution to 13 is given by*

$$D_0^{J,0}(p) = \frac{\kappa}{\kappa + \phi_0^J} - \frac{\kappa}{\kappa + \phi_0^J}\left(\frac{p_t^J}{1-p_t^J}\frac{1-p}{p}\right)^{-\frac{\kappa+\phi_0^J}{\phi_0^J-\phi_1^J+\kappa}} \tag{14}$$

$D_0^{J,0}(p)$ *has the following properties:(a)* $D_0^{J,0}\left(p_t^J\right) = 0$, *(b)* $\lim_{p\to 0^+} D_0^{J,0}(p) = \frac{\kappa}{\kappa+\phi_0^J}$ *and, (c)* $\frac{dD_0^{J,0}(p)}{dp} < 0$.

Proof. To check whether 14 solves 13 follows from evaluating this last expression and checking that the condition holds. (a) follows directly as the solution needs to satisfy the boundary condition $D_0^{J,2,0}\left(p_2^{T,J}\right) = 0$. (b) as $\phi_0^J - \phi_1^J + \kappa \geq 0$ we have that $\lim_{p\to 0^+}\left(\frac{p_2^{T,J}}{p}\frac{1-p}{1-p_2^{T,J}}\right)^{-\frac{\kappa+\phi_0^J}{\phi_0^J-\phi_1^J+\kappa}} = 0$ so that $\lim_{p\to 0^+} D_0^{J,2,0}(p) = \frac{\kappa}{\kappa+\phi_0^J}$. (c) Follows from differentiation wrt. to p. □

H.2 Graduation probability at Junior College

Let $G_\mu^J(p)$ denote the graduation probability at J for students enrolled in J. Using similar arguments that for the dropout probability

$$G_0^J(p) = \frac{\phi_0^J}{\phi_0^J + \kappa} - \frac{\phi_0^J}{\phi_0^J + \kappa}\left(\frac{p}{p_t^J}\frac{1-p_t^J}{1-p}\right)^{\frac{\phi_0^J+\kappa}{\phi_0^J-\phi_1^J+\kappa}}$$

and

$$G_1^J(p) = 1 - \left(\frac{p}{p_t^J}\frac{1-p_t^J}{1-p}\right)^{\frac{\phi_1^J}{\phi_0^J-\phi_1^J+\kappa}}$$

where I used that $G_\mu^J\left(p_t^J\right) = 0$.

Note that $\frac{dG_\mu^J(p)}{dp} < 0$ given that $\phi_0^J - \phi_1^J + \kappa > 0$.

H.3 Transfer probability at Junior College

Add

I Probabilities at College

I.1 Dropout probability at College

Proposition 14. *For low-skilled students and transfer status $TS = 0$ the dropout probabilities are given by*

$$D_0^{C,0}(p) = \frac{\kappa}{\kappa + \phi_0^C}\left(1 - \left(\frac{p}{p_t^C}\frac{1-p_t^C}{1-p}\right)^{-\frac{\kappa+\phi_0^C}{\phi_1^C+\lambda-\phi_0^C-\kappa}}\right)$$

where $D_0^{C,0}\left(p_t^C\right) = 0$.

I.2 Graduation probability at College

Let $\Pr(G|C;p;\mu)$ denote the dropout probability for skill level μ. For $\mu = 0$, this value solves the ODE given by

$$\left(\phi_0^C + \kappa\right)\Pr(G|C;p;0) = \phi_0^C + \frac{d\Pr(G|C;p;0)}{dp}\left(\phi_0^C - \phi_1^C + \kappa - \lambda\right)p(1-p)$$

with terminal condition $\Pr\left(G|C;p_t^C;0\right) = 0$. The solution to this ODE is

$$\Pr(G|C;p;0) = \frac{\phi_0^C}{\phi_0^C + \kappa}\left(1 - \left(\frac{1-p}{1-p_t^C}\frac{p_t^C}{p}\right)^{-\frac{\phi_0^C+\kappa}{\phi_0^C-\phi_1^C+\kappa-\lambda}}\right)$$

with

$$\frac{d\Pr(G|C;p;0)}{dp} = -\frac{\phi_0^C}{\phi_0^C - \phi_1^C + \kappa - \lambda}\left(\frac{1}{1-p} + \frac{1}{p}\right)\left(\frac{1-p}{1-p_t^C}\frac{p_t^C}{p}\right)^{-\frac{\phi_0^C+\kappa}{\phi_0^C-\phi_1^C+\kappa-\lambda}} > 0$$

I.3 Transfer probability at College

add

J Example of Computational Method

In this section I provide an example on how to approach the solution to the model with learning using collocation methods.

I am interested in obtaining $f(p,J,1)$. Assuming that $f(p,J,1)$ is a differentiable function,

$$f(p,J,1) \approx \sum_{i=0}^{I} s_i^{J,1} p^i$$

At the dropout treshold, $f(p,J,1)$ needs to satisfy value matching and smooth pasting,

$$\begin{aligned} f\left(p_d^{J,1},J,1\right) &= \underline{w}^J \\ f_p\left(p_d^{J,1},J,1\right) &= 0 \end{aligned}$$

Also, for $p=1$ it has to be the case that

$$f(1,J,1) = \hat{w}_1^J - \frac{\ln\left(A_1^J\right)}{\gamma}$$

providing three equations.

Next, note that the ODE given by equation 13 needs to holds for every $p \in \left[p_d^{J,1},1\right]$. This means that equation 13, evaluated at different values for p produces more equations. The collocation approach suggests that we should "collocate" the approximation for $f(p,J,1)$ at these points. Finally, the way to choose the collocation nodes is not trivial. As the ODE given by equation 13 is nonlinear, the best possible nodes to pick are the ones given by the Chebyshev's nodes.