EIEF
Einaudi Institute for Economics and Finance

EIEF Working Paper 03/11
January 2011

Option Value and Transitions in a Model of Postsecondary Education

by
Nicholas Trachter
(EIEF)

Option Value and Transitions in a Model of Postsecondary Education

Nicholas Trachter*

EIEF

nicholas.trachter@eief.it

January 5, 2011

Abstract

Option value arises in environments where an investment needs to be made under uncertainty. The decision to invest in postsecondary education is a perfect example. Students, as they learn about the uncertain educational outcomes, can drop out or transfer up to harder and more rewarding schools or even down to easier and less rewarding institutions, carrying a fraction of the accumulated human capital; here, academic 2-year colleges serve as a stepping stone towards more demanding environments as it provides a cheaper learning technology. A positive theory of postsecondary education is built and contrasted empirically. Using an estimated version of the model, it is found that option value explains a large share of the returns to postsecondary education. The elimination of academic 2-colleges, with freshmen enrollment of nearly 40% of that of 4-year colleges, would decrease total enrollment by 6% with very limited effect on the ex-ante returns to education.

KEYWORDS: Option value, learning, experimentation, stepping stone, dynamics, education.

*The online appendix of this paper can be found at my website: https://sites.google.com/site/trachter/. Earlier versions of this paper circulated as "Learning in Postsecondary Education: Academic 2-year Colleges as a Stepping Stone". I thank Fernando Alvarez and Rob Shimer for guidance and useful comments. I also benefited from comments by Gadi Barlevy, Marco Bassetto, Christian Broda, Jeremy Fox, Willie Fuchs, Aspen Gorry, James Heckman, Mariano Lanfranconi, Francesco Lippi, Derek Neal, Juan Pablo Nicolini, Ezra Oberfield, Marcelo Veracierto, Andy Zuppann and seminar participants at the Einaudi Institute for Economics and Finance (EIEF), Federal Reserve Bank at Chicago, Third European Economic Review for Talented Economists Clinic (EERTEC 2010), University of Chicago, Universitat Pompeu Fabra, Universidad Torcuato Di Tella, and Conference participants at LACEA/LAMES 2008 at Rio de Janeiro. All errors are my own.

1 Introduction

Option value is a feature of environments where an investment needs to be made under uncertainty, as in Pindyck (1982), Abel (1983) or Dixit (1989). When there is learning and capital is at least partially transferable, a stepping stone arises as in Jovanovic and Nyarko (1997). Notably, academic 2-year colleges constitute one of these stepping stones; as they provide a cheaper learning technology, they act as a safer path towards more rewarding and demanding environments, namely, 4-year colleges.

Evidence from the *National Longitudinal Study of the High School Class of 1972* (NLS-72) is compelling both in terms of transitions and returns. Dropout rates are high in both academic 2- and 4-year colleges. Graduation rates are 6 times higher in 4-year colleges and transfer rates are mostly one-sided from academic 2- to 4-year colleges. For students enrolling in academic 2-year colleges, the return to graduation is negligible while that to transferring is large. Academic 2-year colleges are cheaper, easier and some fraction of the accumulated credits can be transferred to a 4-year college. Students sort across institution types with respect to how uncertain is their likelihood of success.[1] For students with low expectations, academic 2-year colleges are an ideal training ground. In the educational ladder, 4-year colleges play the role of the step above academic 2-year colleges.

This paper produces a theory of post-secondary education where academic 2-year colleges act a stepping stone following from the availability of dropout and transfer options in both steps of the ladder. An estimated version of the model is then used to evaluate the value-added of each option to each step, how substitutable the steps are, and how does the provision of full insurance affects total enrollment and its composition across types of colleges.

The model incorporates academic 2-year colleges together with 4-year colleges and work. High-school graduates are uncertain about their ability to accumulate human capital. Pessimistic agents join the workforce, optimistic agents enroll in 4-year colleges and those with intermediate beliefs enroll in academic 2-year colleges. During their tenure as students,

[1] Similar to the likelihood of investing by firms facing uncertain demand as in Pindyck and Solimano (1993), Caballero and Pindyck (1996), or Leahy and Whited (1996).

agents are presented with exams, which govern the accumulation of credits and provide information that updates beliefs about ability, inducing dropouts and transfers.

Academic 2-year colleges are ideal for students with aspirations regarding graduation at 4-year colleges but with low expectations about their ability to accumulate human capital. Depending on the evolution of their beliefs and accumulation of credits, students can decide to transfer to 4-year colleges and carry with them a proportion of their stock of credits, implying that academic 2-year colleges act as a stepping-stone, similar to Jovanovic and Nyarko (1997) where workers move up in the work ladder once they acquired the necessary skills. Further, the model has features of bandit models as students learn about their innate ability to accumulate human capital, similar to Johnson (1978) and Miller (1984). Jovanovic and Nyarko (1997) evaluate the predictive power of bandit and stepping stone models in terms of job mobility and find that there is some evidence favoring a combination of both. Post-secondary education also presents features of bandit and stepping stone models.

The model relates to that one of Miao and Wang (2007), that studies the problem of an investor facing two investment projects, one risky and one riskless. The riskiness follows from uncertainty regarding the true quality of the project. For agents investing in the risky project, the arrival of information might induce re-balancing the investment portfolio towards the riskless one. The model developed here has, as a limiting case, a configuration similar to the model in Miao and Wang (2007). When grades provides only information and graduation is random with poisson arrival rate, the model can be casted in a very similar fashion to Miao and Wang (2007) with the only difference that in this model agents face 3 investment opportunities with two that are not absorbing, while in Miao and Wang (2007) they only face two with only one non absorbing. Even in this simplified setup, academic 2-year colleges can be used to learn in a cheaper and less risky environment.

This paper also relates to Groes, Kircher, and Manovskii (2010). They document, using a large panel of workers from Denmark, that for a given occupation, high and low wage earners are more likely to switch occupations, the first to more rewarding occupations and the second to occupations with lower mean wage. Later, they produce a theory of learning about the worker's own ability that explains the sorting and the transitions. In postsecondary education, and in the model developed here, the same pattern of sorting and transitions occurs.

An important feature of the model is its tractability, which allows for a clear characterization of the optimal policy that governs enrollment, dropout and transfer behavior. The model is parameterized using data from NLS-72 by assuming that observable measures of ability are correlated with the high-school graduate initial beliefs. The model is consistent with the following facts: (1) among those initially enrolled in academic 2-year colleges, more able agents are less likely to graduate, more likely to transfer, and less likely to dropout; (2) among those initially enrolled in 4-year colleges, more able agents are more likely to graduate and less likely to dropout or transfer; (3) there is a higher concentration of high ability students among transferees.

Using the parameterized version of the model, a counterfactual analysis of the value added of each option (dropout and transfer), the importance of the lack of full insurance, and the value added of academic 2-year colleges as a whole are performed. To do this not only changes in the enrollment pattern are analyzed but also how the return to enrollment, relative to joining the workforce directly after high school graduation, are affected by each of these options.

The decomposition of returns to enrollment shows that the dropout option accounts for 31% of the full average return to enrolling in an academic 2-year college while the transfer option accounts for 69%. No one would enroll in an academic 2-year college if these options were not available. For 4-year colleges, the dropout option accounts for 87% of the full average return to enrollment while the rest is explained by simple human capital accumulation that follows from enrollment until graduation. Full insurance would reduce enrollment in academic 2-year colleges from nearly 16% to 10%, while enrollment in 4-year colleges would rise from 27% to nearly 50%. The interaction of risk and option value proves to be an important force in post-secondary education.

Finally, the paper then turns to evaluate how close of a substitute are academic 2-year colleges for 4-year colleges to find a high degree of substitutability. In accordance with this high degree of substitutability, the welfare effect of the availability of academic 2-year colleges is at most moderate and is primarily driven by an increase in participation of nearly 7%. This result contrasts with the common wisdom that academic 2-year colleges are a key part of the educational system as they train the students at the margin.

2 Evidence on Option Value: Patterns and Returns of Postsecondary Education

This section presents statistics on postsecondary educational patterns and returns based on the NLS-72. The unit of analysis are high-school graduates that join the workforce directly after leaving school (with no spells of post-secondary education) or join a post-secondary institution with no discontinuities in their educational spells.[2] NLS-72 follows the educational histories of the senior class of 1972 up to 1980. A final wave in 1986 was performed to acquire long-run job market information.

The postsecondary educational system broadly consists of 4-year colleges, academic 2-year colleges and vocational/technical school. The goal of academic 2-year colleges is to prepare students for transferring to 4-year colleges while vocational schools offer specialized education through terminal programs that require either 2 or 3 years to complete.

Table 1 can be used to evaluate the dynamic pattern of postsecondary educational histories. Individuals are faced with an initial enrollment choice between 4-year colleges, academic 2-year colleges, vocational schools or joining the workforce. The first spell of education can end in three different ways. First, a student can dropout (D) and join the workforce. Second, a student can transfer (T) to a different type of educational institution.[3] Third, a student can graduate (G) and join the workforce.[4] The proportion of students that transfer more than once is negligible and therefore the analysis that follows is reduced to only account for

[2] Discontinuous spells are treated as educational histories that include periods of work.

[3] Within-type transfers (e.g. 4-year college to 4-year college) are not understood here as transfers.

[4] A student that transfer holding a degree is counted as a transfer.

students that transfer at most once. Students that transfer can end their second spell of education in two ways: obtaining a degree or becoming a dropout.

Table 1

Transitional dynamics

	share	j								
		vocational schools			academic 2-year colleges			4-year colleges		
		T to j	G at j	D at j	T to j	G at j	D at j	T to j	G at j	D at j
voc. school	9	-	6	88	3	22	78	3	22	78
ac. 2-year c.	15	4	14	86	-	5	59	32	56	44
4-year c.	25	3	11	89	2	11	89	-	54	41

The transitions are for first and second educational spells after high-school graduation. T stands for Transfer, G for graduation and D for drop out. Share: share of high-school graduates that enroll either at vocational school, academic 2-year colleges or 4-year colleges. Source: NLS-72.

Only half of the sample pursues higher education directly after high-school graduation. Among them, nearly 20% enroll in vocational school, around 30% enroll in academic 2-year colleges and the rest enroll in 4-year colleges.

Dynamics (i.e. dropout, graduation and transfer behavior) differ for students depending on their initial enrollment choice. Dropout rates are high in the three types of institutions but are higher in vocational schools and academic 2-year colleges than in 4-year colleges.[5] Transfer rates are important in academic 2-year colleges: approximately 32% of students that initially enroll in this type of institution eventually transfer to 4-year colleges; Only 4-year colleges graduate a large percent of their students. Note that the graduation rate at 4-year colleges is similar for those initially enrolled at 4-year colleges and for those that transferred from academic 2-year colleges. This fact favors the idea that the initial enrollment choice does not hinder the probability of graduation at 4-year colleges.[6]

Low enrollment and high attrition rates can be associated with the risk (possibly due to heterogeneity in returns) and costs attached to education. Costs include foregone earnings

[5]Dropout rates at vocational schools are inflated since vocational schools have students that enroll in particular classes such as Pottery, learning to use Excel, etc. Once they acquire the particular skill, these students leave the school and return to the workforce. These students don't get terminal degrees and so they are recorded as dropouts.

[6]Similar patterns are present in NELS:88 that correspond to a cohort that graduates from high-school in 1992.

(income stream that a student 'loses' by attending school) and direct costs of education that include tuition (and associated fees) and housing. Table 2 shows that 4-year colleges cost twice as much as academic 2-year colleges, providing one reason why students might enroll in academic 2-year colleges.

Table 2
Average Cost of Education

	Tuition	Tuition + R&B
vocational school	3803.84	5904.74
academic 2-year college	1811.36	2729.65
4-year college	3420.73	5038.8

R&B: Room and Board. Missing Values were imputed by running a cost regression and imputing missing values through observables. The values are measured in 1984 dollars. Source: NLS-72.

2.1 Returns to Education

Table 3 presents the Internal Rate of Return for the average student with a particular educational path relative to joining the workforce directly after high-school graduation. The Internal Rate of Return takes into account direct and indirect costs of education together with its benefits. Controlling for observable measures of ability, a mincer and growth regressions are used to back out the wages for different educational histories (benefit) relative to the foregone wage due to education (indirect cost).[7] The direct cost of education includes tuition, books and, sometimes, room and board.

Few agents join the workforce after graduating from academic 2-year colleges providing low confidence in the estimated coefficients in the mincer and growth regressions. In fact, Table 15 presents the results of a test that evaluates whether the estimated wage differential of graduating at academic 2-year colleges differs from the estimate for dropouts. The test can not reject the equality of the coefficients. With this in mind, Internal Rates of Return are calculated in two different ways. The first uses direct results from both mincer and growth regressions. The second calculates Internal Rates of Return by assuming that graduates at

[7]The regressions are in Appendix B.

academic 2-year colleges enjoy similar wage profiles than dropouts.

The results in Table 3 do not represent the true internal rate of return of each alternative since it relies on estimates of wage differentials that do not adjust for self-selection but nonetheless provide intuition regarding to the value attached to each of the possible educational paths.

Table 3

Internal Rates of Return

	Voc. school	Ac. 2-year C.	4-year C.
Graduation	9.37	-0.63 / 4.06	9.93
Dropout	2.07	6.4	8.55
Transfer to V.S.	-	1.44	0
Transfer to Ac. 2-year	3.71 / 4.74	-	3.37 / 3.46
Transfer to 4-year	8.49	8.84	-
Enrollment	2.75 / 2.78	6.63 / 6.87	8.93 / 8.94

The cost of education includes Room and Board for 4-year colleges only. All the values are in percentage points. Appendix B presents the details of the calculations and show how to obtain the wage estimates that are imputed using mincer and growth regressions. Statistics for the direct cost of education can be found in Table 2. The excluded group are agents that join the workforce directly. Finite lifetime of 47 years is assume through out the calculations. Returns to graduation at academic 2-year colleges are computed in two different ways: (1) using the parameters obtained by both mincer and growth regressions and (2) imposing a similar wage profile for graduates than for dropouts.

Table 3 shows that enrollment in 4-year colleges provides the highest return and is mostly driven by the return for graduates.[8] Among agents that enroll in academic 2-year colleges the results show that the best educational path is to eventually transfer to 4-year colleges rather than staying and eventually graduating, reinforcing the 'transfer function' associated with academic 2-year colleges. Finally, note that the internal rates of return conditional on initial enrollment choice are ordered: low in vocational school, average in academic 2-year colleges and high in 4-year college.

[8] The results are in line with empirical studies. See, for example, Belzil and Hansen (2002), or Cunha, Heckman, and Navarro (2005).

2.2 Sorting in Initial Enrollment

Students that enroll in 2-year colleges have observable measures of ability that lie between those of high-school graduates that join the workforce directly and those of students that enroll in 4-year colleges as noted by Grubb (1993) and Kane and Rouse (1999). Table 4 presents summary statistics for measures of ability affecting enrollment decisions tabulated by initial enrollment choice, extending the analysis of Grubb (1993) and Kane and Rouse (1999) by splitting 2-year colleges between vocational schools and academic 2-year colleges. From left to right the table shows that there is evidence of an ordered enrollment choice. For example, see the Rank in the senior year at high-school class. The rank decreases monotonically with the enrollment choice.

Table 4

Summary Statistics for Measures of Ability Conditional on Initial Enrollment Choice

		work	V.S.	Ac. 2-year	4-year C.
Male		0.492	0.4	0.54	0.524
		(0.5)	(0.49)	(0.5)	(0.5)
Black		0.108	0.126	0.085	0.105
		(0.31)	(0.33)	(0.28)	(0.3)
Socio. Status:	Low	0.415	0.294	0.192	0.166
		(0.49)	(0.45)	(0.39)	(0.37)
	Medium	0.506	0.582	0.55	0.407
		(0.5)	(0.49)	(0.49)	(0.49)
	High	0.078	0.122	0.253	0.426
		(0.27)	(0.33)	(0.43)	(0.49)
Education of Father:	<HS	0.516	0.391	0.289	0.215
		(0.5)	(0.48)	(0.45)	(0.41)
	HS	0.323	0.398	0.344	0.288
		(0.46)	(0.49)	(0.47)	(0.45)
	4-year C. (no degree)	0.108	0.145	0.211	0.19
		(0.31)	(0.35)	(0.41)	(0.39)
	4-year C. graduate	0.051	0.064	0.155	0.305
		(0.22)	(0.24)	(0.36)	(0.46)
Rank		0.495	0.435	0.395	0.271
		(0.31)	(0.28)	(0.27)	(0.24)

Rank=rank in high-school class. Socio-Status: Socioeconomic Status of Family at moment of high-school graduation. Source: NLS-72.

2.3 Academic 2-year colleges as a Stepping Stone in an Educational Ladder

Ladders have been associated with the growth of skill as discussed in Jovanovic and Nyarko (1997). Lower steps of the ladder are characterized as stepping stones because they provide a less risky environment to learn compared to higher steps. As agents acquire the necessary skills, they move upwards in the ladder. The process that starts after high-school graduation and culminates with 4-year college graduation is a ladder with two steps. The first step, the stepping stone, is academic 2-year colleges. The second step is 4-year colleges. But there also ways of going down. First, a student at 4-year college can decide to transfer to academic 2-year colleges. Second, a student can decide to dropout.

In contrast with the characterization of the ladder discussed above where agents move on once they acquire the necessary skills, this ladder also present features of bandit models as those discussed in Johnson (1978), Miller (1984) and Jovanovic and Nyarko (1997). Models of skill accumulation imply that agents should enroll first in the lower step of the ladder, as it provides a less riskier environment for learning and experimentation. Bandit models suggest that students should enroll in the harder step - the last step - since the learning technology provides more information about innate ability.

3 Model

The economy is populated by agents that, upon high-school graduation, decide whether to join the labor force or pursue a degree at a post-secondary educational institution. At $t = 0$ agents graduate from high-school endowed with asset level a_0. Agents differ in their ability to accumulate human capital at college, that can either be low or high. Let $\mu \in \{0, 1\}$ denote the ability level, with $\mu = 0$ denoting low ability. The ability level μ is not observable by the agent. Instead, a high-school graduate inherits a signal about her true type denoted by $p_0 \in [0, 1]$, where $p_0 = Pr(\mu = 1)$.

At any period in time an agent can either be working, studying at 4-year colleges (or C) or at academic 2-year colleges (or A). Let $i \in I = \{A, C\}$ denote the type of institution. The cost of education per period of schooling is denoted by τ^i, with $\tau^C > \tau^A$. A student graduates from institution i after accumulating T^i credits, with $T^C > T^A$. The evolution of credits is closely tied to signals that arrive during an agent's tenure as student that are labeled by η.

Work is assumed to be an absorbing state with constant wage function $h(GS, i, \mu)$, where the first argument accounts for the graduation status of the agent, the second for the type of institution where the agent graduated from (highest degree) and the third for her true ability level.[9] Further, the wage function $h(GS, i, \mu)$ is specified as follows:

$$h(GS, i, \mu) = \begin{cases} h^w & \text{if GS=0;} \\ h^i(\mu) & \text{if GS=1.} \end{cases}$$

with $h^i(1) \geq h^i(0) > h^w$ for all i and $h^C(\mu) \geq h^A(\mu) > h^w$ for all μ. That is, for any ability level, graduation at a 4-year college implies higher wage profiles than graduation at an academic 2-year college and, for any institution i, wage profiles of graduates are increasing in their ability level.

The evolution of the asset level a is given by

$$a_{t+1} = \begin{cases} (1+r)a_t - \tau^i - c_t & \text{if enrolled at } i; \\ (1+r)a_t + h(GS, i, \mu) - c_t & \text{if working.} \end{cases}$$

where no borrowing constraints are present.[10]

[9]In the current setup dropouts do not enjoy higher wage profiles. That is, increases in wages only occur after graduation. The model can be easily extended to include this feature by making the function $h(\cdot)$ to depend on amount of credits s but it will loose much of its tractability.

[10]The assumption of no borrowing constraints is consistent with Cameron and Heckman (2001), Cameron and Taber (2004) and Keane and Wolpin (2001), who found no evidence in favor of constraints for the NLSY and with Stinebrickner and Stinebrickner (2008) who found no evidence of credit constraints affecting dropout behavior of student for a panel designed for evaluating dropout behavior. Some studies, on the other hand, found evidence in favor of borrowing constraints (see, for example, Belley and Lochner (2008)).

During tenure as students agents receive signals in form of exams labeled as η. Let η denote the signal with PDF given by $f_i(\eta|\,\mu)$.

Assumption 1 *The ratio of densities* $\frac{f_i(\eta|\mu=1)}{f_i(\eta|\mu=0)}$ *satisfies the Monotone Likelihood Ratio Property (MLRP). That is, for any* $\eta_1 > \eta_0$, $\frac{f_i(\eta|\mu=1)}{f_i(\eta|\mu=0)}$ *is well defined (full support) and*

$$\frac{f_i(\eta_1|\,\mu=1)}{f_i(\eta_1|\,\mu=0)} \geq \frac{f_i(\eta_0|\,\mu=1)}{f_i(\eta_0|\,\mu=0)}$$

The assumption states that high ability students are prone to receiving better signals than low ability students.

The evolution of credits is a function of current signal η and amount of accumulated credits s,

$$s' = s + \Omega(\eta, s)$$

with

$$\Omega(\eta, s) = \begin{cases} \Omega(\eta) & \text{if } s < T^i; \\ 0 & \text{if } s \geq T^i. \end{cases} \tag{1}$$

with $\Omega'(\eta) \geq 0$ so that the evolution of credits is a non-decreasing function on the received signal.[11] Equation (1) states that, while the amount of current credits is less than the necessary amount for graduation, accumulation of credits is only a function of the received signal η.

Students are allowed to transfer and can carry with them part of the credits earned in the current institution. Let θ^i denote the operator that maps credits s in institution i to credits s in institution j, $j \neq i$. Formally,

$$\theta^i(s) : \mathbb{R}^+ \times I \to \left[0, T^j\right]$$

[11] Obtaining a C or an A in a particular subject provides the same accumulation of credits but a different re-evaluation of own's ability by Assumption 1.

where it is assumed that $\theta^i(s)$ is non-decreasing in credits s.

A high-school graduate, endowed with her prior p_0 and initial asset level a_0 chooses her consumption stream $\{C_t : t \geq 0\}$ and whether to enroll in, dropout or transfer in A or C, in order to maximize her time-separable expected discounted lifetime utility derived from consumption,

$$\mathbb{E}\left\{\sum_{0}^{\infty}\left(\frac{1}{1+r}\right)^{t}\left(\frac{e^{-\gamma c_t}-1}{-\gamma}\right)\middle| \mathcal{F}_0\right\}$$

where $\mathcal{F}_0 = \{p_0, a_0\}$ and γ is the coefficient of Constant Absolute Risk Aversion (CARA).

Let $V_i(a, s, p)$ denote the value for a student currently enrolled in institution type i with asset level a, amount of credits accumulated s and prior p. Also, let $W(a; h(GS, i, \mu))$ denote the value for a worker with asset level a and wage profile $h(GS, i, \mu)$. Finally, let $\Lambda(a_0, p_0)$ denote the value for a high-school graduate with asset level a_0 and prior p_0.

The value for a high-school graduate $\Lambda(a_0, p_0)$ equals

$$\Lambda(a_0, p_0) = \max\left(W(a_0; h^w), V_A(a_0, 0, p_0), V_C(a_0, 0, p_0)\right)$$

as the agent chooses whether to join the workforce or pursue higher education (either in academic 2-year colleges or 4-year colleges) by comparing the value of each alternative.

3.1 The problem of a worker

A worker with current asset level a and constant wage h faces the following problem:

$$W(a; h) = \max_{c,a'} \frac{e^{-\gamma c}-1}{-\gamma} + \frac{1}{1+r}W(a'; h) \tag{2}$$

where a' is

$$a' = (1+r)a + h - c$$

That is, the worker has to decide her consumption in the current period and the asset level for next period. The timing of the model is such that decisions (c and a') are made before the worker receives the payment for her work, h. The value for a worker with asset level a and wage profile h is

$$W(a;h) = -\frac{1+r}{\gamma r}e^{-\gamma(ra+h)} + \frac{1+r}{\gamma r} \tag{3}$$

where the details can be found in Appendix C.

Taking the limit when γ approaches 0 (using L'Hopital rule) provides the value for a risk neutral worker,

$$\lim_{\gamma \to 0} W(a;h) = \frac{1+r}{r}(ra+h)$$

that is, the value is just the present value of labor income in addition of the current wealth level.

3.2 The problem of a student

The student's beliefs are updated by the stream of information that arrive through the signal η. Let $p' = b(\eta;p)$ denote the posterior that depends on the prior p and the signal η. For a given institution i, Bayes' rule is

$$b(\eta;p) = \frac{1}{1 + \frac{f_i(\eta|\mu=0)}{f_i(\eta|\mu=1)}\frac{1-p}{p}}$$

The evaluation of expectations about future income flows depends on the likelihood of the signals. Any new signal can be produced by either $f_i(\eta|\mu = 1)$ or $f_i(\eta|\mu = 0)$ so that expectations about the governing pdf have to be accounted for. Define

$$H_i(\eta,p) \equiv pF_i(\eta|\,\mu = 1) + (1-p)F_i(\eta|\,\mu = 0)$$

as the CDF that accounts for this uncertainty.

The problem faced by a student in institution type i can be written as

$$V_i(a,s,p) = \max_{c,a'} \frac{e^{-\gamma c}-1}{-\gamma} + \frac{1}{1+r}\int_\eta \tilde{V}_i(a',s',p')H_i(d\eta,p) \tag{4}$$

with

$$\begin{cases} a' = (1+r)a - \tau^i - c \\ s' = s + \Omega(\eta) \\ p' = b(\eta;p) \end{cases}$$

The value $\int_\eta \tilde{V}_i(a',s',p')H_i(d\eta,p)$ accounts for the continuation value, where a student evaluates the different available options after updating the state of the problem. In any given period a student that accumulated s' credits faces alternatives. If $s' < T^i$ she can decide to stay in the current institution, transfer or drop. If $s' = T^i$ graduation is a fact and the options are reduced to graduation or transfer as dropping out is dominated.[12] Let $\mathbb{I} = 1$ if $s' < T^i$ and $= 0$ otherwise. $\tilde{V}_i(a',s',p')$ is equal to

$$\max\left\{W(a';h^w), \mathbb{I}V_i(a',s',p') + (1-\mathbb{I})[p'W(a';h_1^i) + (1-p')W(a';h_0^i)], V_{-i}(a',\theta^i(s'),p')\right\} \tag{5}$$

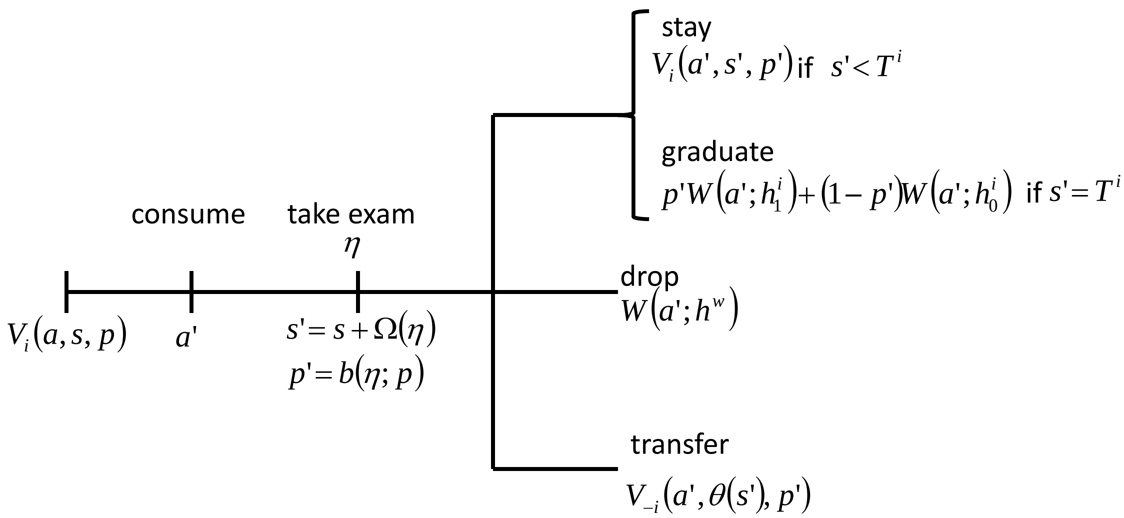
The timing of the problem, presented in Figure 1, is as follows. For a given an institutional choice i at the beginning of a period a student chooses her consumption and level of assets for the next period given her expectations about future income streams. Next, she receives the signal η, producing bayesian updating of prior $p' = b(\eta,p)$, and the amount of credits accumulated for next period s'. When the new period begins the student chooses whether to dropout or remain a student and whether to transfer to another institution.

The solution to the student's problem is

$$V_i(a,s,p) = -\frac{1+r}{\gamma r}e^{-\gamma(ra+v_i(s,p))} + \frac{1+r}{\gamma r} \tag{6}$$

[12] As $h_1^i > h_0^i > h^w$ and $W(a;h)$ increasing in wage h, it follows directly that $pW(a;h_1^i)+(1-p)W(a;h_0^i) > W(a;h^w)$ as $p \in [0,1]$. Then, the dropout option is strictly dominated by the graduation option.

Figure 1
Timeline



At the beginning of the period, $V_i(a,s,p)$ denotes the value of enrollment at institution type i with wealth level a, accumulated credits s and prior p. The student chooses her consumption level producing the accumulated wealth for next period a'. Later, she takes an exam, receiving grade η. This produces a reevaluation of beliefs $p' = b(\eta; p)$ and accumulation of credits $s' = s + \Omega(\eta)$. The student then chooses between staying in institution type i (or graduating if enough credits where accumulated), transferring or dropping out.

where $v_i(s,p)$ solves a recursive equation,

$$v_i(s,p) = \frac{\tilde{v}_i(s,p) - r\tau^i}{1+r} \tag{7}$$

with

$$\tilde{v}_i(s,p) = -\frac{1}{\gamma}\ln\left[\int_\eta -\max\left\{\begin{array}{c} -e^{-\gamma h^w}, -e^{-\gamma v_{-i}(\theta^i(s'),p')}, \\ -\left(\mathbb{I}\left(p'e^{-\gamma h_1^i} + (1-p')e^{-\gamma h_0^i}\right) + (1-\mathbb{I})e^{-\gamma v_i(s',p')}\right) \end{array}\right\} H_i(d\eta, p)\right] \tag{8}$$

being the consumption equivalent of the continuation value and where $p' = b(\eta, p)$ and $s' = s + \Omega(\eta)$. The details of the calculations can be found in Appendix D. Also, the value for a student $V(a,s,p)$ is both increasing and convex in credits s and prior p.[13]

[13] See Appendix E.

Taking the limit when γ approaches 0 produces the value under risk neutrality,

$$V_i(a,s,p) = \frac{1+r}{r}(ra + v_i(s,p))$$

with

$$v_i(s,p) = \frac{\lim_{\gamma\to 0} \tilde{v}_i(s,p) - r\tau^i}{1+r}$$

where

$$\lim_{\gamma\to 0} \tilde{v}_i(s,p) = \int_\eta \max \left\{ \begin{array}{c} h^w, v_{-i}(\theta^i(s+\Omega(\eta)), b(\eta,p)), \\ \mathbb{I}\left(b(\eta,p)h_1^i + (1-b(\eta,p))h_0^i\right) + (1-\mathbb{I})\, v_i(s+\Omega(\eta), b(\eta,p)) \end{array} \right\} H_i(d\eta,p)$$

The proof can be found in Appendix F.

3.3 Characterization of Solution

The model is built to analyze a particular pattern of education. That is, students with high priors enroll in 4-year colleges, average priors enroll in academic 2-year colleges and low priors join the workforce directly. The next assumption addresses this point.

Assumption 2 *The primitives of the model are such that*

$$\begin{cases} \frac{\tilde{v}_C(0,1)-r\tau^C}{1+r} & \geq & \frac{\tilde{v}_A(0,1)-r\tau^A}{1+r} & \geq & h^w \\ \frac{\tilde{v}_C(0,0)-r\tau^C}{1+r} & \leq & \frac{\tilde{v}_A(0,0)-r\tau^A}{1+r} & \leq & h^w \end{cases}$$

The assumption states that high-school graduates with low ability to accumulate human capital are better off joining the workforce and, in the eventuality of enrollment, they are better off in academic 2-year colleges than in 4-year colleges. The opposite idea applies for high ability agents. They are better off by pursuing higher education and the best enrollment choice for them are 4-year colleges.

Assumption 2 has an interesting interpretation. The existence of academic 2-year colleges in this model is driven by the learning mechanism and the option value.

The next proposition describes the optimal policy.

Proposition 1 *For a given institution i and amount of credits s, the optimal policy is a collection of dropout and transfer thresholds $\{p_d^i(s), p_t^i(s)\}_{s\in[0,T^i]}$ independent of the current wealth level a. The optimal policy for students enrolled in institution i is*

$$\text{Enrolled in } i \text{ today: } \begin{cases} \text{join workforce tomorrow} & \text{if } p < p_d^i(s) \\ \text{enrolled in } A \text{ tomorrow} & \text{if } p \in [p_d^i(s), p_t^i(s)] \\ \text{enrolled in } C \text{ tomorrow} & \text{if } p > p_t^i(s) \end{cases}$$

The proof is straightforward and therefore not presented. Independence of the optimal policy function with respect to current wealth level a is a result of the choice of CARA utility function and the dependence of the policy function on credits s follows as the amount of accumulated credits affect both the amount of credits that can be transferred and also the likelihood and time to graduation. Assumption 2 drives the optimal policy not only at time 0 but also as credits accumulate. Consider the case where T^i is large so that in terms of distance until graduation an agent with $s = 0$ and one with $s = 1$ are very similar. It follows that a similar condition holds for $s = 1$ but the difference in the value functions should decrease as students get closer to graduation.

The next proposition evaluates the interaction of accumulated credits s and the evolution of the thresholds for the case where the support of credits s is $\mathbb{R}^+$.

Proposition 2 *If (i) $F_i(\eta|\,\mu)$ is continuous and differentiable for all i and μ, (ii) $\Omega(\eta)$ is continuous and differentiable, and (iii) $\theta^i(s)$ is continuous, differentiable and concave with $\frac{\partial\theta^i(0)}{\partial s} < 1$, it is the case that*

$$\frac{\partial p_d^i}{\partial s} \leq 0 \quad , \quad \frac{\partial p_t^A}{\partial s} \geq 0 \quad , \quad \frac{\partial p_t^C}{\partial s} \leq 0$$

Proof. See Appendix G. ■

The proposition states that as credits accumulate, the likelihood of transferring or dropping out decreases as the terminal payoff at institution i is getting closer.

3.4 Returns to Enrollment

The lack of available assets to diversify the risk that comes from education (as income flows are unknown) imply that standard techniques to value the option to pursue post-secondary education can not be applied. To value the option and compute returns, define $\Sigma_i(p)$ as the value-added of enrollment in institution i relative to joining the workforce directly after high-school graduation for an individual with prior p.[14] $\Sigma_i(p)$ is the compensating variation of enrollment in institution i over the outside option and can be understood also as the maximum amount of units of consumption a high-school graduate is willing to forego in order to remain enrolled in i and not be forced to drop out (note that the option includes tuition),

$$V_i(a - \Sigma_i(p), 0, p) = W(a; h^w)$$

Solving the above equation yields an expression for the value-added by enrollment,

$$\Sigma_i(p) = \frac{v_i(0,p) - h^w}{r}$$

The intuition behind the formula for $\Sigma_i(p)$ has a clear interpretation. It is simply the difference between the risk-adjusted expected discounted flow of income due to enrollment and the discounted flow of income of the outside option.

The price of the option is given by the opportunity cost of becoming a student, $\frac{1+r}{r}h^w$. That is, the discounted income flow from joining the workforce directly after high-school graduation. The return to enrollment at institution i relative to joining the workforce $R_i(p)$

[14] Miao and Wang (2007) uses a similar approach to value an investment project where the income flow is uncertain and the risk is uninsurable.

is defined as

$$R_i(p) \equiv \frac{\Sigma_i(p)}{\frac{1+r}{r}h^w}$$

4 Parametrization

The model explores the dynamic interaction of academic 2-year colleges and 4-year colleges. The evidence obtained from NLS-72 shows that vocational school (excluded from the model analysis) can be merged with the workforce as little interaction occurs between vocational school and other types of institutions (see Table 1) and the sorting in initial enrollment, presented in Table 14 and Table 13, places vocational school below academic 2-year colleges. Version B in all of the tables accounts for the case where work and vocational school are merged. Further, as in the model, column 2 on the tables accounts for the cases where increases in wages only occur upon graduation.

The operator that maps credits s in institution i to credits that remain after transferring $\theta^i(s)$ is simplified to be of the multiplicative form,

$$\theta^i(s) = \begin{cases} \theta^i s & \text{if } \theta^i s < T^{-i} \\ T^{-i} & \text{if } \theta^i s \geq T^{-i} \end{cases}$$

The signal η plays two different roles in the model. First, it updates beliefs p as the signal conveys information regarding the likelihood of the true talent level of the student. Under this definition, the signal η accounts for grades in exams, in subjects, problem sets, overall experience as a student, etc. The second role of the signal η is to generate accumulation of credits through the function $\Omega_i(\eta)$, which suggests that the signal is closely tied to grades in subjects.

To simplify the model, think that the signal η is the mean of the grades in a quarter obtained by a student.[15] The set of possible values of η is simply the set of possible grades.

[15]It is possible to relax this assumption by choosing functional forms for the signal η that allows for a decomposition of the signal in two parts: one that accounts for grades in subjects and another that accounts for the rest.

For simplicity assume three possible grades: $\{F, N, E\}$. That is, a student can fail, get a neutral grade or excel in a particular exam. Let $q^i_\mu(\eta)$ denote the probability of each event. Further, assume that $q^A_1(F) = q^C_0(E) = 0$. That is, high ability students never fail an exam at academic 2-year colleges and low ability students never excel at 4-year colleges.

The function $\Omega(\eta)$, that maps grade η into credits s is chose to be as follows:

$$\Omega(\eta) = \begin{cases} 0 & \text{if } \eta = F \\ 1 & \text{if } \eta = N \\ 1 & \text{if } \eta = E \end{cases}$$

The time period is chosen to be a quarter, therefore $T^A = 8$ and $T^C = 16$ (a student needs to accumulate T^i quarters of accumulated credits at institution i to graduate). The risk-free interest rate r is set to be 0.45% which implies a yearly interest rate of 1.81% and a yearly discount factor of 0.9822. All the monetary values in the model are measured in logs and further standarized by the wage of agents with no degrees h^w, so that $h^w = 1$.[16] Academic 2-year colleges are located in every city and town while 4-year colleges are scarce. This way, the cost of education includes housing for 4-year colleges and does not include housing for academic 2-year colleges because students there can live with their parents. The standarized cost of education is then $\tau^A = 0.1152$ for academic 2-year colleges and $\tau^C = 0.3205$ for 4-year colleges (see Table 2). The risk aversion parameter, γ, is chosen to be equal to 8.[17]

Figure 2 plots the fraction of the initial population of academic 2-year colleges that drop, transfer or graduate for a given period. As seen in the figure, transfer occurs, on average, after the completion of the first year of education.[18] Then, θ^A is chosen to be $\frac{1}{2}$.

[16]The mean wage in 1985, in 1984 dollars, for agents with no degrees was 17740.63.

[17] The risk aversion parameter γ is hard to identify and the literature didn't spend much time estimating risk aversion parameters using CARA utility functions (for natural reasons). There is a whole string of literature in asset pricing that argues that the CRRA risk aversion parameter σ lies between 4 and 10. Using the definition of relative risk aversion it is possible to relate σ and γ, $\gamma c = \sigma$. Here, the consumption level c has a lower bound given by $ra + h^w \geq 1$ so $\gamma < 10$.

[18]Students transfer before obtaining a degree or completing the course-work at academic 2-year colleges. Only 12.5% of students that transfer from academic 2-year colleges to 4-year colleges in the NLS-72 sample holds a degree and usually transfer around one year later than those not holding a degree.

The evidence for students that transfer from 4-year colleges to academic 2-year colleges is less revealing as the fraction of students that transfer is very low (see Table 1). Figure 3 shows that students transfer during their first year of education. Among academic 2-year college graduates, those that started their educational career at 4-year colleges spend more time in school prior to graduation (4.5 years vs. 3.84 years). Then, $\theta^C = 0$.

Figure 2

Transitions in ac. 2-year colleges

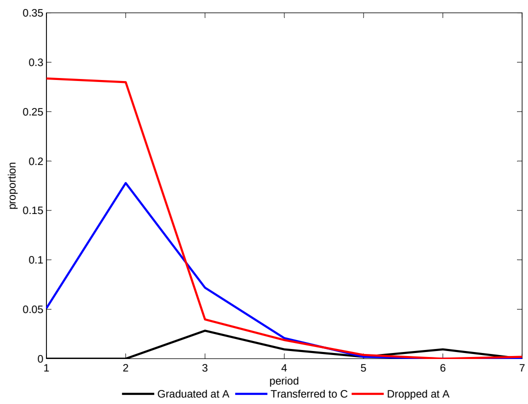
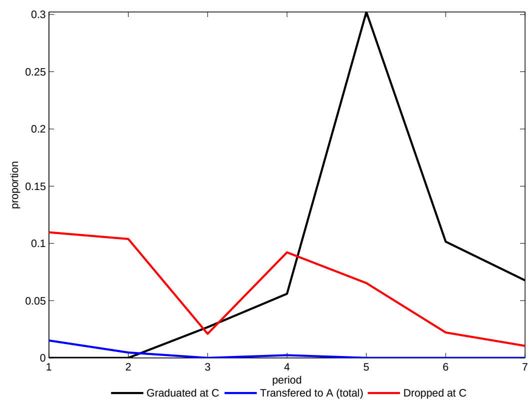


Figure 3

Transitions in 4-year colleges



The remaining parameters, $q^i_\mu(\eta)$ and $h^i(\mu)$ are estimated using a two-stage Simulated Method of Moments. In the first stage an ordered probit regression on the initial enrollment choice is used to produce estimates of the initial prior p_0. In the second stage, conditional on the distribution of priors obtained in the first stage, enrollment and transition choices are simulated for each individual on the sample to match an over-identified set of moments.

Let

$$p_0 = \left(1 + e^{-(X'\beta+\varepsilon)}\right)^{-1}, \, \varepsilon \sim N(0,1) \tag{9}$$

where X is a vector that includes all the observable characteristics of high-school graduates that are correlated with the ability level of the agent and β is the vector of factor loadings, identified by an ordered probit for the initial choice of agents. The estimates for β are presented in Table 13.

Students with high priors join 4-year colleges, with average priors join academic 2-year colleges and with low priors join the workforce. The thresholds are those of the optimal policy considered above for $s = 0$ as agents that graduate from high-school didn't acquire any credits yet. Further, monotonicity of p_0 as a function of $X'\beta + \varepsilon$ implies that β can be estimated by an ordered probit on the initial choice (Table 13 - Version B) and then upper and lower bounds for ε can be computed using $X'\hat{\beta}$ and the enrollment choice of the agent. Further, the cutoffs shown in Table 13 (version B) are monotonic transformations of the threshold for p.

Conditional indirect inference, through Simulated Method of Moments, is used to estimate the remaining 10 parameters (the 6 learning parameters, and the four wages). The chosen moments are: (1) proportion of students that join workforce from high-school, (2) proportion of students that enroll in academic 2-year colleges, (3) proportion of students that enroll in 4-year colleges, (4) proportion of students initially enrolled in academic 2-year colleges that dropped-out, (5) proportion of students initially enrolled in academic 2-year colleges that transfer to 4-year colleges, (6) proportion of students initially enrolled in academic 2-year colleges that graduate at 2-year colleges (highest degree), (7) proportion of students initially enrolled in 4-year colleges that dropped-out, (8) proportion of students initially enrolled in 4-year colleges that transfer to academic 2-year colleges, (9) proportion of students initially enrolled in 4-year colleges that graduate at 4-year colleges, (10) mean wage for academic 2-year college graduates after first spell of education, and (11) mean wage for 4-year college graduates after first spell of education.[19] The estimated parameters for the learning process are presented in Table 6 and in Table 5 for the wage distribution. Table 7 compares the moments for data with those generated by the model conditional on the distribution of priors implies by the data.

To produce the different moments it is necessary to compute the probability of dropout, transfer and graduation at both academic 2-year colleges and 4-year colleges for each level of the initial prior p_0. A simple way of computing these values is via the Kolmogorov Forward Equation. To fixate ideas see the next example. Let $Pr^A(D|p, s)$ denote the dropout probability at A with current prior p and accumulated credits s. Also, define $\omega_\eta^A = pq_\eta^A(1) +$

[19] Local identification was checked by computing the derivative of the moments and noting that they are different from zero.

$(1-p)q^A_\eta(0)$ as a set of weights. Finally, let p'_η denote the posterior for grade η and prior p. For $p \in [p^A_d(s), p^A_t(s)]$,

$$
\begin{aligned}
Pr^A(D|p,s) \quad = \quad & \sum_\eta \omega^A_\eta \left[\mathbb{I}\{p'_\eta > p^A_t(s')\} * 0 + \mathbb{I}\{p'_\eta < p^A_d(s')\} * 1 \right. \\
& \left. + \mathbb{I}\{p^A_d(s') \leq p'_\eta \leq p^A_t(s')\} * Pr^A(D|p'_\eta, s') \right]
\end{aligned}
$$

with terminal condition $Pr^A(D|p,T^A) = 0$ for $p \in [p^A_d(T^A), p^A_t(T^A)]$.

Using this recursive equation it is possible to back-out $Pr^A(D|p_0, 0)$. This can be easily extended to all the other probabilities.

Estimation of the model requires the next steps: (1) pick values for the learning parameters and wages, (2) solve the model numerically, (3) for every value of $X'\beta$ (that is, for every individual on sample - using the results of the ordered probit regression in Table 13 and the initial choice of each individual -) produce the censored distribution of initial priors p_0, (4) for every p_0, use the Kolmogorov Forward Equations to produce the probabilities of each different event, (5) using step (3) compute the probabilities of each different event and wages upon graduation for every individual on sample, (6) produce a set of moments, and (7) repeat the procedure until the convergence criterion is reached (objective is to minimize the squared difference of the simulated and true moments from data).

Table 5 shows that the wage differential to graduation at both academic 2- and 4-year colleges relative to joining the workforce without holding a degree. For high ability graduates the differential is almost three times bigger in 4-year colleges than in academic 2-year colleges suggesting that students with high expectations currently enrolled in academic 2-year colleges should be eager to transfer to 4-year colleges. As students can transfer credits, many enroll in academic 2-year colleges as individuals are risk-averse and therefore care about the high volatility of wages.

Table 6 shows the estimated learning parameters. Trivially, by assumption, failing an exam at academic 2-year colleges signals low ability while an excelling at 4-year colleges

Table 5

Estimated Wage Differentials by Conditional Simulated Method of Moments

		estimate
academic 2-year colleges		
low ability	$\frac{h^A(0)-h^w}{h^w}$	0.03
high ability	$\frac{h^A(1)-h^w}{h^w}$	0.08
4-year colleges		
low ability	$\frac{h^C(0)-h^w}{h^w}$	0.04
high ability	$\frac{h^C(1)-h^w}{h^w}$	0.22

signals high ability as the probability of failing an exam at academic 2-year colleges was set to be zero for high ability students and the probability of excelling in an exam at 4-year college was set to zero for low ability students. The likelihood of obtaining a neutral or an excellent at academic 2-year colleges is higher for high ability students and thus receiving these signals improves the expectations about the innate ability. Similar idea happens for a neutral grade at 4-year colleges while a fail lowers the expectations.

Table 6

Learning Parameters estimated by Conditional Simulated Method of Moments

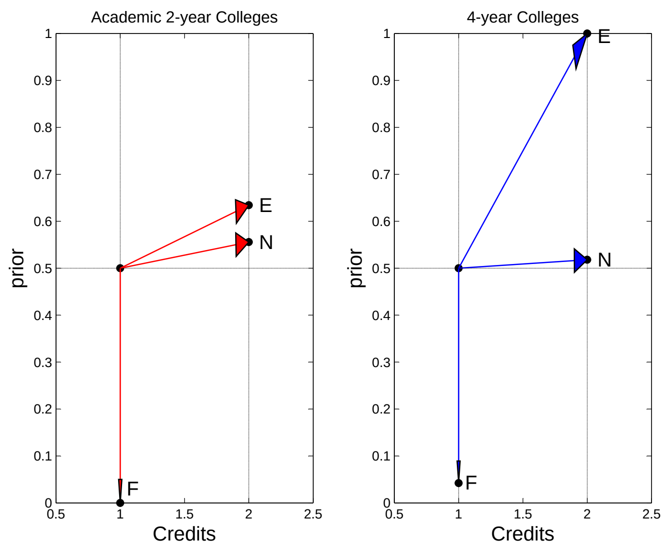
	academic 2-year colleges		4-year colleges	
	low ability $\mu = 0$	high ability $\mu = 1$	low ability $\mu = 0$	high ability $\mu = 1$
Fail	0.23	0[a]	0.2275	0.01
Neutral	0.695	0.87	0.7725	0.81
Excel	0.075	0.13	0[b]	0.18

[a] by assumption.
[b] by assumption.

Figure 4 presents the graphic analysis of the evolution of credits and beliefs as signals arrive. Receiving a fail in either academic 2- or 4-year college makes the posterior to fall sharply and the credits to remain constant. Obtaining a neutral or an excellent in both types of school increase by one the amount of credits and makes the posterior to be higher than the prior. A neutral grade at academic 2-year colleges has a higher effect on the belief than in 4-year colleges while an excellent provides much less information. These results show an important characteristic of post-secondary education: academic 2-year colleges are easy and thus provide more information in the left tail while 4-year colleges are hard and thus provide more information in the right tail. Further, note that the higher tuition at 4-year colleges relative to academic 2-year colleges is not only due to higher wages upon graduation. As seen in the figure, the variance of the signals is higher at 4-year colleges meaning that it provides more information.

Figure 4

Example of the Evolution of Prior and Credits as a result of the received grade



As initial priors are generated by using the empirical distribution of $X'\beta$ the estimation strategy is attempting to match not only the 11 aggregate moments but instead the marginal density of them. Table 7 presents the value of the eleven moments used in the estimation for both the NLS-72 and the simulated version of the model.

The evolution of thresholds as a function of accumulated credits are presented in Figure 5 and Figure 6. In a broad way, as discussed in Proposition 2, the inaction region in both academic 2- and 4-year colleges increases with credits. Opposed to Proposition 2 the inaction does not increase monotonically provided that signals are discrete and, as a result, the amount of credits s can only take a finite amount of values and thus the proof does not go through: the discreteness of the accumulated credits s together with the transfer function

Table 7
Moments in Data and Model

% of High-school graduates that	**Data**	**Model**
join Workforce	59.4	56.9
enroll in A	15.2	15.69
enroll in C	25.4	27.41
% of those initially enrolled in A that		
drop at A (1^{st} spell)	63.2	48.63
transfer from A to C (1^{st} spell)	32.2	43.79
graduate at A (1^{st} spell)	4.6	7.58
% of those initially enrolled in C that		
drop at C (1^{st} spell)	40.9	42.1
transfer from C to A (1^{st} spell)	2	0.11
graduate at C (1^{st} spell)	57.1	57.79
Mean Wage Differential for Graduates after (1^{st} spell)		
graduated from A	0.045	0.045
graduated from C	0.212	0.1482
Moments to Discipline Priors		
FOCs from ordered probit (for β)		YES
distribution of X		YES

To simulate the moments first it is necessary to produce initial priors. Let $p_0 = \left(1 + e^{-(X'\beta+\varepsilon)}\right)^{-1}$ where X is a vector of observable measures of ability and $\varepsilon \sim N(0,1)$. β is estimated using an Ordered Probit regression on the initial enrollment choice. Bounds for ε are also produced by looking at the choices of the individuals after high-school graduation.

$\theta^i(s)$ being increasing, imply that non-monotonicity can arise naturally in the framework.[20]

5 Fit of Model

The prior p_0 is positively correlated with the measure of ability $X'\beta$ obtained from the ordered probit regression (Table 13 and Table 14) so that the enrollment pattern generated

[20] To fully understand this results consider the case where $\theta^i(s) = 0$. Let $s_1 > s_0$ be two different amounts of credits. Further, let $p_t^A(s_0)$ denote the transfer threshold at s_0 at academic 2-year colleges. By definition, $v_A(s_0, p_t^A(s_0)) = v_C(0, p_t^A(s_0))$. As both functions are increasing and convex in p and $v_A(0,1) < v_C(0,1)$, the fact that $v_A(s_1, p_t^A(s_0)) > v_A(s_0, p_t^A(s_0)) = v_C(0, p_t^A(s_0))$ implies that $p_t^A(s_1) > p_t^A(s_0)$.

Figure 5

Regions at Academic 2-year Colleges

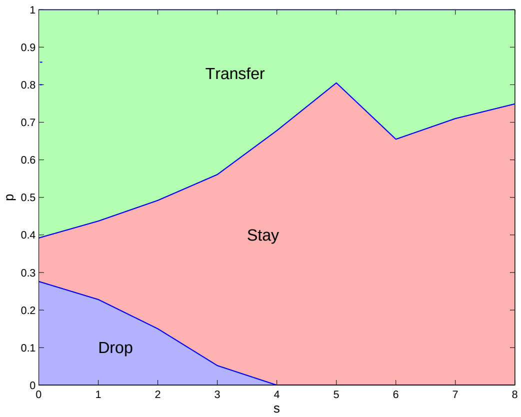
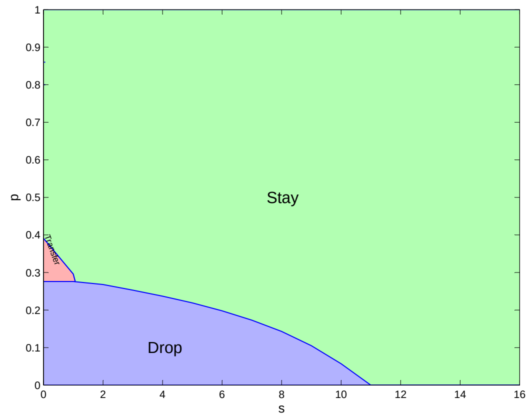


Figure 6

Regions at 4-year colleges



The dropout and transfer thresholds at both academic 2- and 4-year colleges are as a function of amount of accumulated credits s.

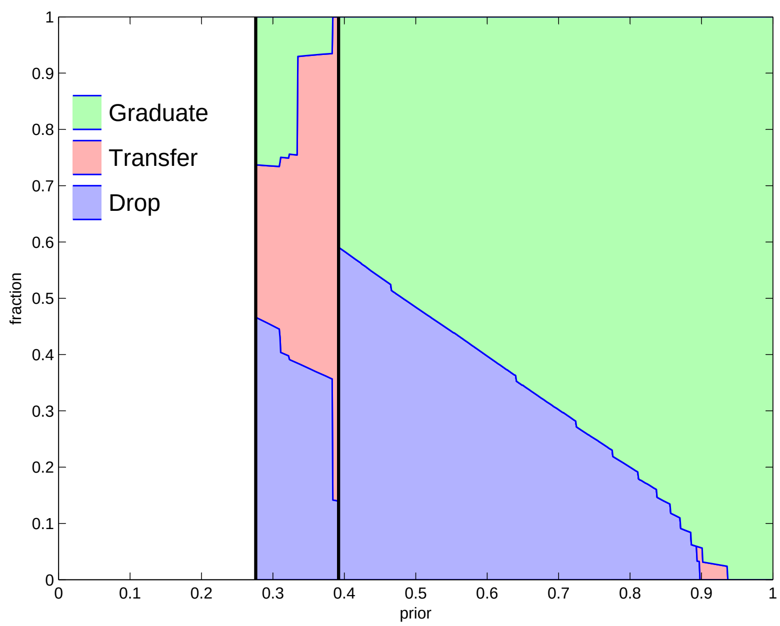
by the model fits the empirical distribution obtained from NLS-72.

The model also has predictions regarding the dropout, transfer and graduation behavior of students. In particular, conditional on the initial enrollment choice, the model produces probabilities of different educational patterns as a function of the initial prior p_0, as shown in Figure 7. The initial prior p_0 affects the decision making of the student and the dynamic pattern in two different ways. First, it affects the likelihood of different educational histories as the distance to different threshold values changes with the prior. Second, the value of the prior is related to the likelihood of different signals as $p_0 = \Pr[\mu = 1]$. In fact, the estimation of the model attempted to match the empirical marginal density of dropouts, transfers and graduation at both academic 2- and 4-year colleges.

Figure 7 has three different regions (the straight vertical lines separate the different regions). The first region, given for low values of the prior p_0, is for agents that join the workforce directly. The second region, the middle one, is for agents that enroll in academic 2-year colleges (average values for the prior) and the third region, the top one, is for agents that enroll in 4-year colleges. Conditional on the initial enrollment choice , Figure 7 presents the likelihood of each of the three possible events (i.e. drop, transfer or graduation) in the first spell of education for a student with a given initial prior p_0. The likelihood of graduation and dropping out in academic 2-year colleges are decreasing functions of the initial prior p_0 while the likelihood of transferring to 4-year colleges is increasing. For students that initially enroll in 4-year colleges, the likelihood of graduation increases with the prior while the likelihood of dropping out decreases with the prior. Another interesting aspect observed in Figure 7 is that students that transfer from 4-year colleges to 2-year colleges have above average priors (relative to students that enroll in 4-year colleges).[21]

Figure 7

Probability of Dropout, Transfer, and Graduation conditional on initial enrollment



The probabilities are computed according to the initial enrollment choice of an individual with initial prior p_0. The plot has three regions (from left to right). The first region is empty as individuals with low priors join the workforce. The second region is for individuals enrolling in academic 2-year colleges, while the third region is for individuals that enroll in 4-year colleges.

To evaluate the predictions of the model regarding the transition probabilities, Table 8 shows students by behavior (i.e. dropout, transfer, graduation) in the first spell of education and size of the measure of talent $X'\hat{\beta}$. For students initially enrolled in academic 2-year colleges, the pattern observed for dropout and transfer probabilities is similar to the

[21]The shape of the figure is robust to different calibrations of the model. In particular, calibrations that match more closely dropout and transfers (in detriment of enrollment moments).

model's predictions. A similar thing happens for the dropout and graduation probabilities for students initially enrolled in 4-year colleges. Graduation probability in academic 2-year colleges and transfer probability in 4-year colleges are less revealing due to the low number of students included in these cells. Still, the evidence in these two cases does not conflict with the model's predictions.

Table 8

Proportion of agents initially enrolled in i with particular history

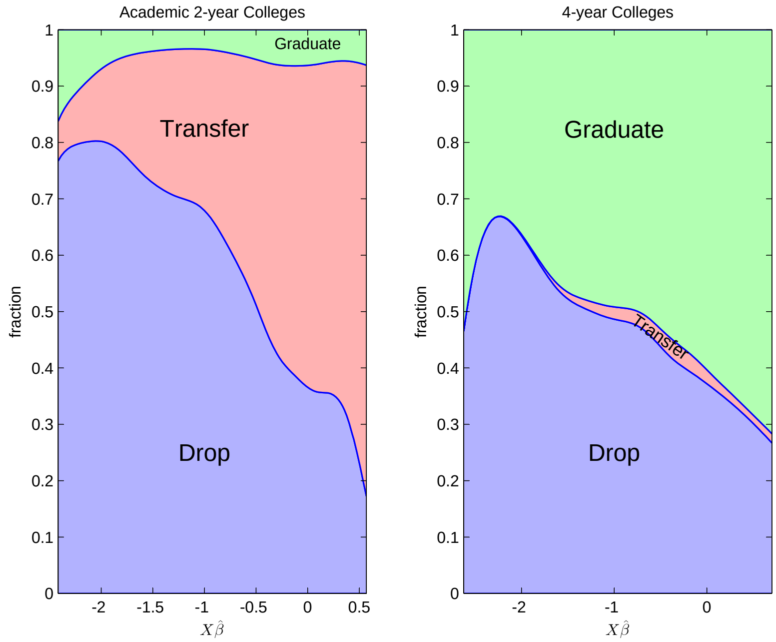
	Ac. 2-year C.			4-year C.		
	D	T	G	D	T	G
Low $X'\hat{\beta}$	77.3%	17.6%	5.1%	51.7%	1.7%	46.6%
Med $X'\hat{\beta}$	66.1%	29.4%	4.5%	45.4%	2.4%	52.2%
High $X'\hat{\beta}$	44.9%	50%	5.1%	35.6%	2.4%	62%
# of obs.	332	192	26	392	47	474

These proportions are conditional on the initial enrollment status of the student. D: Drop, T: Transfer, G: Graduate. $X'\hat{\beta}$: estimator of the observable measures of ability. Recall that $p_0 = \frac{1}{1+e^{-(X'\beta+\varepsilon)}}$ and that a student enrolls in 4-year college if her prior is high enough and in an academic 2-year college if it is average (so that she compares her initial prior with thresholds). Then, the enrollment problem can be estimated by using an Ordered Probit regression from where β is estimated.

An alternative, and more involved way of evaluating Figure 7 would be, conditional on the initial enrollment choice of agents, to do a nonparametrical estimation of the densities of each of the different educational histories. Weighting these densities accordingly, it is possible to produce the empirical counterpart of Figure 7. First, the density associated with a given history is weighted by its share on initial enrollment in a given institution. For a given measure of talent $X'\hat{\beta}$, now it is possible to compute the proportion of agents that eventually end their first spell of education either by becoming dropouts, transferring or by graduation. Figure 8 presents the results. Notice how the patterns in Figure 7 (model) are very similar to those in Figure 8 (data).

Figure 8

Density for the NLS-72 cohort of dropout, transfer, and graduation probabilities



As a function of ability measure $X'\hat{\beta}$ and conditional on initial enrollment choice. First, a conditional (on the initial enrollment choice) nonparametric estimation of each event was performed. Next, each density was weighted by their share in initial enrollment. Finally, for every level of $X'\hat{\beta}$, the proportion of each event was constructed.

6 Insurance and Option Value

High dropout and transfer rates are features commonly associated with risk and thus the availability of transfer and dropout options should be highly valued by agents as they provide lower bounds to the risk of the investment. In terms of risk, keeping the primitives unaltered the model is solved again letting γ tend to zero. This case maps to risk full insurance. Comparing the benchmark model (i.e. $\gamma = 8$) with the risk-neutral case provides insights regarding the interaction of risk and insurance with the optimal policy and returns in this economy.[22] A similar strategy is followed to evaluate the size of the option value. Still keeping the primitives unaltered, the model is solved two more times. The first time, the

[22]The analysis will abstract from Moral Hazard that can potentially arise from credit provision.

transfer option is discarded and the second time eliminates both the transfer and dropout options. The value-added of each option is then evaluated using a decomposition of returns.

6.1 Insurance

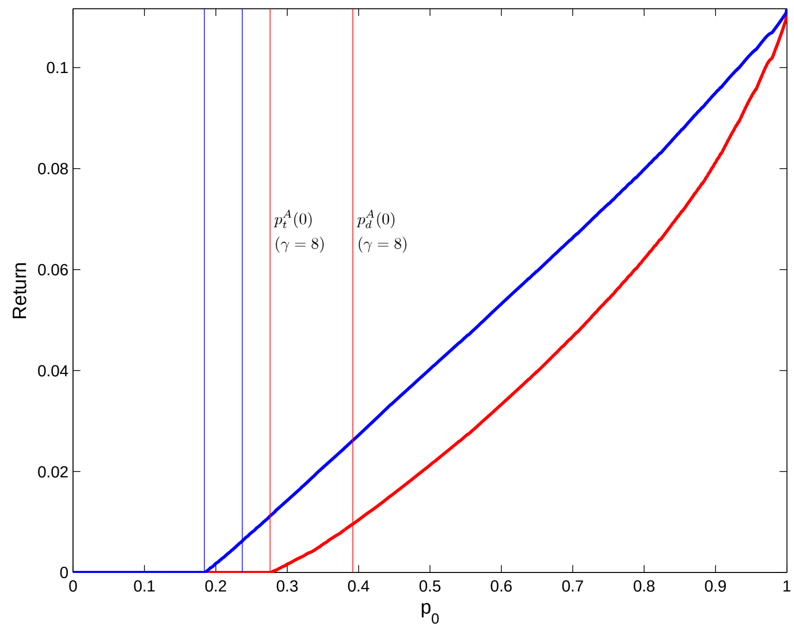
For a high-school graduate with any given prior p_0, Figure 9 presents the returns for the benchmark model ($\gamma = 8$) and the model where $\gamma \to 0$. The comparison is important since risk aversion is tightly connected to market completeness. The more complete the markets, the lower the value for γ. Figure 9 therefore compares the benchmark model with an economy where markets are complete. When risk aversion decreases, the enrollment thresholds shift to the left as the risk implied by education is discounted less heavily by agents. The fact that the shift to the left is stronger in the threshold between 4-year colleges and academic 2-year colleges than for the one between academic 2-year colleges and work is not casual: enrollment at 4-year colleges is more risky than enrollment at academic 2-year colleges (simple comparison of the ratio of wages). It follows that a decrease in risk aversion has a stronger effect on 4-year colleges than in academic 2-year colleges. Figure 9 also shows that risk aversion hinders the returns to education in an important way, and the effect is stronger the more uncertain the prior is.

Table 9 presents the mean return for the cross-section of agents that initially enroll at either academic 2-year colleges and 4-year colleges for both the benchmark and the risk-neutral models using the estimated distribution of priors for the NLS-72 data. The provision of insurance increases returns unambiguously for every prior p_0 but decreases measured returns in academic 2-year colleges through the compositional change that follows the provision of full insurance.

Providing insurance not only increases returns for every prior but also affects enrollment decisions (this can be seen in Figure 9 where the vertical dotted lines denote the enrollment thresholds). Table 10 computes the distribution of initial enrollment for both cases. As expected, full insurance increases total enrollment by nearly 40% and enrollment in 4-year colleges - where risk matters the most as the wedge in wages and cost of education are

Figure 9

The Effect of Full Insurance on Individual Returns



The full insurance case is when $\gamma \to 0$. The vertical lines define the indifference prior for enrollment between work and academic 2-year colleges and between academic 2-year colleges and 4-year colleges. Return: measures the return at high-school graduation of enrollment at institution i relative to joining the workforce.

Table 9

The Effect of Full Insurance on Average Measured Returns

	ac. 2-year college	4-year college
$\gamma = 8$	0.41	3.09
$\gamma \to 0$	0.29	3.42

All the numbers in the table are in percentage points. 10,000 samples of priors p_0 were generated using the measures of ability $X'\hat{\beta}$, the functional form $p_0 = \frac{1}{1+e^{-(X'\beta+\varepsilon)}}$ and the initial enrollment choice of each individual.

higher and time until graduation longer -by around 81%. Finally, the mass of students still enrolling in academic 2-year colleges with full insurance highlights the importance of the learning channel as a feature of academic 2-year colleges. Insurance affects the enrollment distribution both at the extensive and intensive level. At the extensive level, providing full insurance increases total enrollment. At the intensive level, the provision of full insurance affects the composition of enrollment as risk (through differential learning), tuition and wages upon graduation differ across types of institutions.

Table 10

The Effect of Full Insurance on the Initial Enrollment Distribution

	workforce	ac. 2-year college	4-year college
$\gamma = 8$	56.9	15.68	27.42
$\gamma \to 0$	39.82	10.42	49.76

All the numbers in the table are in percentage points. 10,000 samples of priors p_0 were generated using the measures of ability $X'\hat{\beta}$, the functional form $p_0 = \frac{1}{1+e^{-(X'\beta+\varepsilon)}}$ and the initial enrollment choice of each individual.

6.2 How much Option Value?

The model is solved again once more to evaluate the size of the option value, this time reducing the amount of options. First, the transfer option is discarded and therefore the only available alternative after the initial enrollment choice is to dropout. Second, the dropout option is discarded thus no action, other than consumption decisions, is possible during tenure as student. Let $R_i^{E+D+T}(p_0)$, $R_i^{E+D}(p_0)$, and $R_i^{E}(p_0)$ denote the value of enrollment at institution i: with both options available to the agent; with only the dropout option available; and with no dropout or transfer options available.[23]

Trivially,

$$R_i^{E+D+T}(p_0) = R_i^{E+D+T}(p_0) + R_i^{E+D}(p_0) - R_i^{E+D}(p_0) + R_i^{E}(p_0) - R_i^{E}(p_0)$$

Rearranging and dividing by $R_i^{E+D+T}(p_0)$ provides the following decomposition of returns,

$$1 = \frac{R_i^{E+D+T}(p_0) - R_i^{E+D}(p_0)}{R_i^{E+D+T}(p_0)} + \frac{R_i^{E+D}(p_0) - R_i^{E}(p_0)}{R_i^{E+D+T}(p_0)} + \frac{R_i^{E}(p_0)}{R_i^{E+D+T}(p_0)}$$

The first term in the right hand side is the value-added to total returns $R_i^{E+D+T}(p_0)$ by the transfer option, the second term provides the value-added by the dropout option and the third term accounts for the value accrued through enrollment. Figure 10 shows that

[23] An alternative decomposition would be first to discard the dropout option and later discard the transfer option.

returns at 4-year colleges are explained by the dropout option and by simply having the enrollment choice, in accordance with high graduation and dropout rates observed for 4-year college students at NLS-72. Also, Figure 10 accounts for the importance of the transfer option in explaining returns to academic 2-year college enrollment. As observed in NLS-72 (see Table 1), the value-added by the enrollment option that accounts for the simple human capital accumulation story has zero share of the returns.

Figure 10
Decomposition of Returns

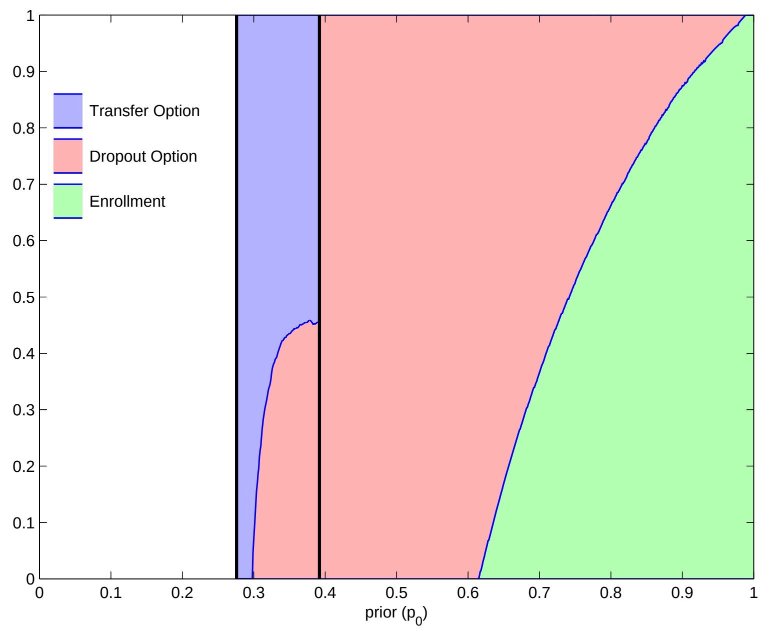


Table 11 produces the same decomposition this time for the mean return of the population distribution of priors p_0. The transfer option is very valuable in academic 2-year colleges, accounting for 69% of total value. The dropout option is valuable in both types of institutions but more so in 4-year colleges.

Table 11
The Averaged Value Added of each Option

	ac. 2-year C.		4-year C.	
	value-added	cumulative	value-added	cumulative
Enrollment	0	0	12.96	12.96
Dropout Option	30.81	30.81	87.04	100
Transfer Option	69.19	100	0	100

All the numbers in the table are in percentage points. 10,000 samples of priors p_0 were generated using the measures of ability $X'\hat{\beta}$, the functional form $p_0 = \frac{1}{1+e^{-(X'\beta+\varepsilon)}}$ and the initial enrollment choice of each individual.

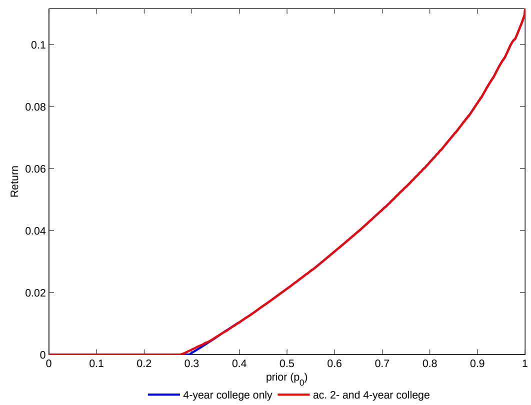
7 How close of a Substitute are Academic 2-year Colleges for 4-year Colleges?

The main purpose of enrollment in academic 2-year colleges is to learn in a less demanding and cheaper environment and eventually transfer to 4-year colleges with a fraction of the already accumulated credits. Still, agents have the choice to enroll at 4-year colleges. Given that a set of agents prefer to enroll in academic 2-year colleges, the question that arises is how much value academic 2-year colleges provide to these agents relative to initial enrollment at 4-year colleges. That is, how close of a substitute are these institutions? To answer this, Figure 11 evaluates how the elimination of academic 2-year colleges affects the return for an agent with prior p_0. For students that enroll in 4-year colleges in the benchmark model, the availability of academic 2-year colleges provides no value since the transfer option has no value for them. For students that enroll in academic 2-year colleges in the benchmark model what happens when academic 2-year colleges are eliminated is: (1) most students simply enroll in 4-year colleges and the difference in value is very small, and (2) the rest join the workforce. Overall, academic 2-year colleges are very close substitutes for 4-year colleges.

In order to characterize the welfare effect of academic 2-year colleges, an analysis of population aggregates is needed. Specifically, it is important to evaluate how participation in post-secondary education and welfare change due to the availability of academic 2-year colleges. Recall that $R_i(p)$ is a monotonous transformation of $v_i(0,p)$ that in turn is the

Figure 11

Value added by the availability of Academic 2-year colleges



certainty equivalent of enrollment at institution i abstracting from the wealth level a. As a result, $R_i(p)$ is a good measure of utility and therefore a comparison of returns provides a good approximation for welfare losses or gains. Table 12 compares total participation and measured returns for the benchmark model where academic 2-year colleges are available and the only enrollment option are 4-year colleges. The availability of academic 2-year colleges increases participation by nearly 7% and generates a drop in measured returns of around 6% mostly due to the compositional change in total participation.

Table 12

The Effect of Academic 2-year Colleges on Enrollment and Returns

	total participation	return
4-year colleges + ac. 2-year colleges	43.11%	2.12%
Only 4-year colleges	40.3%	2.25%

10,000 samples of priors p_0 were generated using the measures of ability $X'\hat{\beta}$, the functional form $p_0 = \frac{1}{1+e^{-(X'\beta+\varepsilon)}}$ and the initial enrollment choice of each individual.

8 Conclusion

This paper proposes a simple, highly tractable, model of post-secondary education, that incorporates together the work decision, academic 2- and 4-year colleges. In the model, students are allowed to drop out and transfer to more/less rewarding institutions as credits are accumulated and expectations are adjusted. The decision is not a trivial one. Higher rewards imply higher costs of education and more risk.

The parameterized version of the model is consistent with the data in terms of transitions, sorting, and returns. Therefore, it sheds some light on the role played by academic 2-year colleges. These institutions act as a stepping stone towards more rewarding, and demanding, environments, namely, 4-year colleges.

In line with the literature of investment under uncertainty, the drop out and transfer options, and the availability of academic 2-year colleges, have an attached option value. A novel feature of this paper is that presents structural estimates of the different option values.

The drop out option explains a large share of ex-ante returns on both academic 2- and 4-year colleges. In fact, very few students would pursue higher education if dropping out is not allowed. The transfer option is not valued at 4-year colleges but explains nearly 70% of ex-ante returns at academic 2-year colleges, consistent with the idea of a stepping stone.

Interestingly, academic 2-year colleges as a whole are not highly valued, even though there are 4 students enrolling in this institutions per 10 enrolling in a 4-year college. Their value comes mostly by a modest increase in participation from the marginal students. All the other, would simply enroll in a 4-year college given that they are still allowed to drop out. This raises the question on how the government should target its spending in postsecondary education.

References

Abel, A. B. (1983): "Optimal Investment under Uncertainty," *American Economic Review*, 73(1), 228–33.

Belley, P., and L. J. Lochner (2008): "The Changing Role of Family Income and Ability in Determining Educational Achievement," *Journal of Human Capital*, 1(1).

Belzil, C., and J. Hansen (2002): "Unobserved Ability and the Return to Schooling," *Econometrica*, 70(5).

Caballero, R. J., and R. S. Pindyck (1996): "Uncertainty, Investment, and Industry Evolution," NBER Working Papers 4160, National Bureau of Economic Research, Inc.

Cameron, S., and J. J. Heckman (2001): "The Dynamics of Educational Attainment for Blacks, Whites and Hispanics," *Journal of Political Economy*, 109(3).

Cameron, S., and C. Taber (2004): "Borrowing Constraints and the Returns to Schooling," *Journal of Political Economy*, 112.

Cunha, F., J. J. Heckman, and S. Navarro (2005): "Separating uncertainty from heterogeneoity in life cycle earnings," *Oxford Economics Papers*, 57(2).

Dixit, A. K. (1989): "Entry and Exit Decisions under Uncertainty," *Journal of Political Economy*, 97(3), 620–38.

Groes, F., P. Kircher, and I. Manovskii (2010): "The U-Shapes of Occupational Mobility," *working paper*.

Grubb, W. N. (1993): "The Varied Economic Returns to Postsecondary Education: New Evidence from the Class of 1972," *Journal of Human Resources*, 28(2).

Heckman, J. J., L. J. Lochner, and P. Todd (2006): *Earnings Functions, Rates of Return, and Treatment Effects: The Mincer Equation and Beyond*, vol. Handbook of the Economics of Education. Elsevier Science.

——— (2008): "Earnings Functions and Rates of Return," *Journal of Human Capital*, 2(1).

Johnson, W. R. (1978): "A Theory of Job Shopping," *Quarterly Journal of Economics*, May.

Jovanovic, B., and Y. Nyarko (1997): "Stepping-stone mobility," *Carnegie-Rochester Conference Series in Public Policy*, 46.

Kane, T. J., and C. E. Rouse (1995): "Labor-Market Returns to Two- and Four-Year College," *American Economic Review*, 85(3).

——— (1999): "The Community College: Educating Students at the Margin between College and Work," *Journal of Economic Perspectives*, 105(3).

Keane, M. P., and K. I. Wolpin (2001): "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, 42(4).

Leahy, J. V., and T. M. Whited (1996): "The Effect of Uncertainty on Investment: Some Stylized Facts," *Journal of Money, Credit and Banking*, 28(1), 64–83.

Lucas, R. E. J., and N. Stokey (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press.

Miao, J., and N. Wang (2007): "Experimentation under Uninsurable Idiosyncratic Risk: An Application to Entrepreneurial Survival," *unpublished*.

Miller, R. (1984): "Job Matching and Occupational Choice," *Journal of Political Economy*, 92(6).

Pindyck, R. S. (1982): "Adjustment Costs, Uncertainty, and the Behavior of the Firm," *American Economic Review*, 72(3), 415–27.

Pindyck, R. S., and A. Solimano (1993): "Economic Instability and Aggregate Investment," NBER Working Papers 4380, National Bureau of Economic Research, Inc.

Stinebrickner, T. R., and R. Stinebrickner (2008): "The Effect of Credit Constraints on the College Drop-out Decision: A Direct Approach Using a New Panel Study," *American Economic Review*.

Appendix

A Ordered Probit on Initial Enrollment Choice

Ordered returns to enrollment together with the evidence presented in Table 4 suggest that the initial enrollment choice is ordered as follows: work, vocational school, academic 2-year colleges, and 4-year colleges. Table 13 in Appendix A presents the results of an ordered probit regression of the initial enrollment choice on a vector X of observable measures of ability. Let β denote the vector of factor loadings. Relative to Kane and Rouse (1999), the analysis is extended here to consider vocational school and academic 2-year colleges as separate institutions. The reference column in Table 13 is Version A (Version B pools vocational school and work together).

The value $X'\beta$ is a composite measure of ability, consistently estimated by $X'\hat{\beta}$. To evaluate a measure of the degree of sorting in initial enrollment, Table 14 produces the mean and standard deviation (in the cross-section) of $X'\hat{\beta}$ across the different alternatives. See

Table 13
Ordered Probit Regression

Variable		Version A	Version B
Male		0.181	0.234
		(0.041)	(0.044)
Black		0.357	0.365
		(0.068)	(0.073)
Socio. Status:	Low	-0.896	-0.9
		(0.088)	(0.093)
	Medium	-0.596	-0.609
		(0.071)	(0.074)
Education of Father:	<HS	-0.363	-0.381
		(0.072)	(0.076)
	HS graduate	-0.137	-0.115
		(0.065)	(0.068)
	4-year graduate	0.324	0.352
		(0.083)	(0.086)
Rank		-1.301	-1.41
		(0.072)	(0.078)
Cut 1		-1.129	-0.893
		(0.077)	(0.079)
Cut 2		-0.859	-0.368
		(0.077)	(0.078)
Cut 3		-0.339	-
		(0.075)	-
# of observations		3462	3462

Ordered Probit estimation of the initial enrollment choice on observable measures of ability. Version B merges vocational school with work. Source: NLS-72.

Version A in the first row of the table. The measure of ability $X'\beta$ is unitless as it is just an ordinal representation of ability measures. Start with high-school graduates that join the workforce - labeled as work in the table - and move upwards across enrollment options. The mean value for the measure of ability increases monotonically with the enrollment options.

B Internal Rates of Return

B.1 Backing Out Wage Profiles

The typical Mincer regression evaluates the effect of educational histories on lifetime earnings by estimating a wage regression on years of education and work experience. There is a significant ongoing literature that accounts for non-linearities in years of education (see Grubb (1993), Heckman, Lochner, and Todd (2006), Heckman, Lochner, and Todd (2008)

Table 14
Measure of Ability $X'\hat{\beta}$

Version	work	vocational school	academic 2-year college	4-year college
A	-1.782	-1.609	-1.348	-0.972
	(0.567)	(0.571)	(0.604)	(0.649)
B		-1.426	-1.013	-0.633
		(0.572)	(0.609)	(0.654)

Constructed from Ordered Probit Estimation (see Table 13). In version B work and vocational school are collapsed together. Source: NLS-72.

and Kane and Rouse (1995)). The typical example in favor of non-linearities is the graduation premium or sheepskin effect. This literature has treated years of education (or amount of credits earned) in different type of institutions as perfect substitutes. Instead, it is now assumed that different educational histories affect lifetime earnings in different ways. Further, as has been already discussed in the literature, this analysis breaks the additive form (in the log version in the Mincer regression) of years of education and experience by estimating a growth equation.

Table 15 presents the results of the extended mincer regression, accounting for the different types of education and graduation premium. See Version A. Graduation in both vocational schools and 4-year colleges is associated with higher wages relative to dropping out. The same idea does not apply to academic 2-year colleges as the return to becoming a dropout is higher than the return from graduation. The low number of students that graduate at academic 2-year colleges raises questions about the significance of the difference. The significance test presented in Table 15 shows that there is not enough evidence to reject that the return to graduation is similar (or even higher) to the return to dropping out.

Table 16 presents the results of a growth regression where the dependent variable is the average growth rate of wages between 1979 and 1985. The growth rate for vocational school graduates, α_G^V is around half the growth rate of 4-year college graduates α_G^C. This fact, together with the results from Table 15 reads as follows: graduation at vocational schools provide a higher wage and 4-year colleges provide steeper profiles of wages, while graduation at academic 2-year colleges is dominated.

Backing out the initial wage upon graduation can not be done directly from the data because wage information prior to 1979 is scarce. To produce the initial wage, it is possible to combine the results of Table 15 and Table 16.

B.2 Computing Internal Rates of Return

Let w_0 denote the wage for an agent that joins the workforce at $t = 0$. Let L denote the lifetime of an agent, S_i the proportion of time spent at institution i, τ_i the flow cost of attendance, ϖ_i^D the increase in wages due to dropping out at instituion i, ϖ_i^G the graduation premium, α_i the increase in wages due to experience, G_i a dummy that accounts for

Table 15

Mincer Regression

	description	Version A	Version B (1)	Version B (2)
ϖ_C^D	drop at 4-year C.	0.24 (0.042)	0.22 (0.041)	-
ϖ_A^D	drop at Ac. 2-year C.	0.09 (0.045)	0.077 (0.044)	-
ϖ_V^D	drop at Voc. school	0.072 (0.046)	-	-
ϖ_C^G	graduation at 4-year C.	0.304 (0.04)	0.286 (0.04)	0.2122 (0.037)
ϖ_A^G	graduation at Ac. 2-year C.	0.015 (0.1)	0.0149 (0.104)	0.0445 (0.104)
ϖ_V^G	graduation at Voc. school	0.284 (0.15)	-	-
		Prob > F	Prob > F	
	Test: $\varpi_A^G = \varpi_A^D$	0.5698	0.5735	-
	Test: $\varpi_C^G = \varpi_V^G$	0.9092	-	-

Dependent Variable: log of wages in 1985. Independent Variables: graduation and dropout dummies for each type of institution and controls for ability measures. Version B collapse together vocational schools and workforce. Source: NLS-72.

graduation at institution i and D_i a dummy that accounts for dropping out at institution i. Further, assume that students transfer only once.[24]

The initial wage for a student with history $H = \{S, G, D\}_{i \in \{V,A,C\}}$ when joining the workforce is

$$w_0(H) = w_0 e^{\sum_i \varpi_i^D D_i + \varpi_i^G G_i}$$

The present value of costs $K(H, r)$ attached to history H is

$$K(H,r) = \int_0^{S_i} e^{-rs_i} \tau_i ds_i + \int_{S_i}^{S_i + S_{-i}} e^{-rs_{-i}} \tau_{-i} ds_{-i}$$

that can be reduced to

$$K(H,r) = \left(1 - e^{-rS_i}\right) \frac{\tau_i}{r} + e^{-rS_i} \left(1 - e^{-rS_{-i}}\right) \frac{\tau_{-i}}{r}$$

[24]Very few students transfer more than once in NLS-72. Also, the assumption makes the presentation of the methodology much easier.

Table 16
Growth Regression

	description	Version A	Version B	
			(1)	(2)
α_G^C	growth rate for C. grads.	0.0447 (0.0144)	0.0457 (0.011)	0.0507 (0.0143)
α_G^A	growth rate for A. grads.	-0.0036 (0.0351)	-0.0071 (0.018)	0.0046 (0.0358)
α_G^V	growth rate for V. grads.	0.0268 (0.0536)	-	-
α_D^C	growth rate for C. dropouts	0.024 (0.0165)	0.0236 (0.0164)	-
α_D^A	growth rate for A. dropouts	0.0103 (0.0196)	0.0091 (0.0194)	-
α_D^V	growth rate for V. dropouts	-0.0071 (0.0202)	-	-
α^0		-0.009 (0.0166)	-0.006 (0.0164)	0.0101 (0.0147)

Dependent Variable: average growth rate of wages between 1979 and 1985. Source: NLS-72

The internal Rate of Return is the interest rate $r(H)$ that solves,

$$\int_{S_i+S_{-i}}^{L} e^{-(r(H)-\alpha_i)t} w_0(H)dt - K(H, r(H)) = \int_0^L e^{-(r(H)-\alpha_0)t} w_0 dt - K(0, r(H))$$

The variables S_i, D_i and G_i can be obtained directly from inspection of the dynamic patterns of education. L is chosen so that agents are alive until they are 65 years old. Then, $L = 47$.

Table 15 shows the results of an extended Mincer regression using the log of wages in 1985 as dependent variable and years of education at a particular institution and graduation dummies as explanatory variables. The coefficients ϖ_i^D and ϖ_i^G can be obtained from that table (see Version A).

Table 16 present the results of a growth regression on the graduation status and type of every individual. The results of this table are interpreted here as estimates of α_i (see Version A).

The average time spent in each institution for a given educational path used in the calculations can be found in the Supplementary Material.

C Computing the Value for a Worker

Solving for c from the budget constraint and substituting back into equation (2) reduces the problem to a single variable problem. Further, it is straightforward to check that the conditions for unique solution to equation (2) are satisfied (see Lucas and Stokey (1989)).

The first order condition with respect to a' reads

$$e^{-\gamma((1+r)a-a'+h)} = \frac{1}{1+r}\frac{dW(a';h)}{da'}$$

Substituting back into equation (2) provides the maximized value function,

$$W(a;h) = \frac{1}{-\gamma(1+r)}\frac{dW(a';h)}{da'} + \frac{1}{1+r}W(a';h) + \frac{1}{\gamma}$$

This equation is satisfied for $W(a;h) = -\frac{1+r}{\gamma r}e^{-\gamma(ra+h)} + \frac{1+r}{\gamma r}$.

D Computing the Value for a Student

Solving for c from the budget constraint and substituting back into equation (4) reduces the problem to a single variable problem.

The lowest level for a' that can be chosen is 0 and the highest is $(1+r)a - \tau^i$. Define $\Gamma(a) = [0, (1+r)a - \tau^i]$ so that the choice variable a' belongs to the graph $\Gamma(a)$.

The next claim shows that there exists a unique solution to equation (4).

Claim 1 *$V_i(a, s, p)$ is single-valued.*

Proof. First note that the only choice variable is a, that p evolves stochastically and s' is a function of p and signals.

Signals η that arrive produce updating in the state p, which is thus stochastic. [25] Let P be such that $p \in P$. Trivially, $P = [0, 1]$ and thus P is compact. Also $s \in [0, T^i]$ so the set for s is compact. The union of compact sets is compact. Further, the transition from p to p' satisfies the feller property.

Next, note that $\Gamma(a)$ is non-empty, compact and continuous. Also, as $c > 0$, $\frac{e^{-\gamma c}-1}{-\gamma}$ is bounded.

Then, Theorem 9.6 of Lucas and Stokey (1989) is satisfied and thus the proposition holds. ∎

After substituting the first order condition into equation (4) provides the maximized value function,

$$V_i(a, s, p) = \frac{1}{-\gamma(1+r)}\frac{d\int_\eta \tilde{V}_i(a', s', p')H(d\eta, p)}{da'} + \frac{1}{1+r}\int_\eta \tilde{V}_i(a', s', p')H(d\eta, p) + \frac{1}{\gamma} \quad (10)$$

[25] In principle the function $p' = b(\eta, p)$ can be inverted to produce a stochastic process for the evolution of p.

Conject that $\int_\eta \tilde{V}_i(a', s', p')H(d\eta, p) = -\frac{1+r}{\gamma r}e^{-\gamma(ra'+\tilde{v}_i(s,p))} + \frac{1+r}{\gamma r}$ and substitute into equation (10) together with $a' = a - \frac{\tau^i}{1+r} - \frac{\tilde{v}_i(s,p)}{1+r}$ to obtain

$$V_i(a, s, p) = -\frac{1+r}{\gamma r}e^{-\gamma(ra+v_i(s,p))} + \frac{1+r}{\gamma r}$$

where $v_i(s, p)$ solves the recursive equation

$$v_i(s, p) = \frac{\tilde{v}_i(s, p) - r\tau^i}{1+r}$$

Further, applying the conjecture and using equation (5) reads,

$$\tilde{v}_i(s, p) = -\frac{1}{\gamma}\ln\left[\int_\eta - \max\left\{\begin{array}{c} -e^{-\gamma h^w}, -e^{-\gamma v_{-i}(\theta^i(s'),p')}, \\ -\left(\mathbb{I}\left(p'e^{-\gamma h_1^i} + (1-p')e^{-\gamma h_0^i}\right) + (1-\mathbb{I})e^{-\gamma v_i(s',p')}\right) \end{array}\right\} H_i(d\eta, p)\right]$$

with $s' = s + \Omega(\eta)$ and $p' = b(\eta, p)$.

Finally, note that the conjecture for $\tilde{v}_i$ holds as a result of the functional form of $V_i(a, s, p)$.

E Properties of Value Function of Students

The proof for p follows by induction. The ultimate goal in post-secondary education is graduation at 4-year colleges so start with a student that accumulated $s = T^C - 1$ credits. $\frac{V_C(a,T^C-1,p)}{dp} > 0$ as (1) the wage upon graduation is increasing in the agent's true ability level, (2) the prior p measures the probability of high ability, (3) the pdf of grades satisfy the Monotone Likelihood Ratio property and (4) $\Omega(\eta)$ non-decreasing. For a student enrolled in academic 2-year colleges with $s = T^A - 1$ the same proof applies but it is necessary to add that the continuation value (through the transfer option) is increasing in the prior p. For $s = T^C - 2$ and any institution i the proof follows as properties (2)-(4) still hold and the continuation values are increasing in p. Convexity follows directly as, (a) for any p the continuation value is bounded below by the dropout option, (b) the continuation value of transferring or remaining at institution i increasing in p, (c) the function max() being convex.

The proof for s is very similar and simpler so it is left as an exercise to the interested reader.

F Computing the Value for a Risk Neutral Student

See equation (8). For any given prior p, each of the exponentials inside the max{} as function of η are bounded as the set of attainable $b(\eta, p)$ being compact ($b(\eta, p) \in [0, 1]$) and the set of attainable payoff being $\left[e^{-\gamma h^w}, e^{-\gamma h_1^C}\right]$, also compact. It follows that each of the functions inside the max{} can be arbitrarily well approximated by a Taylor expansion (each of these functions is also differentiable). As an example, see that

$$e^{-\gamma v_{-i}(\theta^i(s+\Omega(\eta)),b(\eta,p))} \approx 1-\gamma v_{-i}(\theta^i(s+\Omega(\eta)),b(\eta,p))+O(-\gamma v_{-i}(\theta^i(s+\Omega(\eta)),b(\eta,p)))$$

where the term $O(-\gamma v_{-i}(\theta^i(s+\Omega(\eta)),b(\eta,p)))$ is a sequence of terms of order higher than one. Moreover, each of the functions $O(-\gamma v_{-i}(\theta^i(s+\Omega(\eta)),b(\eta,p)))$ is a convergent series and thus bounded.

The previous argument can be used to approximate the elements of $\max\{\}$. For example,

$$b(\eta,p)e^{-\gamma h_1^i}+(1-b(\eta,p))e^{-\gamma h_0^i} \approx 1-\gamma(b(\eta,p)h_1^i+(1-b(\eta,p))h_0^i)+O(\gamma,h_1^i,h_0^i,b(\eta,p))$$

where $O(\gamma,h_1^i,h_0^i,b(\eta,p)) = b(\eta,p)O(-\gamma h_1^i)+(1-b(\eta,p))O(-\gamma h_0^i)$ is bounded and convergent by composition of bounded and convergent series.

Rewrite the original expression as

$$\tilde{v}_i(s,p) = -\frac{1}{\gamma}\ln\left[\int_\eta \left(1-\max\left\{\begin{array}{c} \gamma\left(h^w-\frac{O(\gamma,h^w)}{\gamma}\right),\gamma\left(v_{-i}-\frac{O(-\gamma v_{-i})}{\gamma}\right), \\ \mathbb{I}\gamma\left(v_i-\frac{O(-\gamma v_i)}{\gamma}\right)+ \\ (1-\mathbb{I})\gamma\left(p'h_1^i+(1-p')h_0^i-\frac{O(\gamma,h_1^i,h_0^i,p')}{\gamma}\right) \end{array}\right\}\right)H_i(d\eta,p)\right]$$

where $p'=b(\eta,p)$ and the state of v_i and v_{-i} is omitted to ease notation.

The claim follows by taking the limit when γ approaches 0. L'Hopital is required as $\lim_{\gamma\to 0}\tilde{v}_i(s,p)=\frac{0}{0}$. The issue is whether the function $\max\{\}$ is differentiable with respect to γ and, in the affirmative case, how to characterize the derivative.

Recall that both v_i and v_{-i} depend of credits s and prior p. Define

$$\begin{array}{rcl} J_1(\gamma,p',s') & = & \gamma\left(h^w-\frac{O(\gamma,h^w)}{\gamma}\right) \\ J_2(\gamma,p',s') & = & \gamma\left(v_{-i}-\frac{O(-\gamma v_{-i})}{\gamma}\right) \\ J_3(\gamma,p',s') & = & \mathbb{I}\gamma\left(v_i-\frac{O(-\gamma v_i)}{\gamma}\right)+(1-\mathbb{I})\gamma\left(p'h_1^i+(1-p')h_0^i-\frac{O(\gamma,h_1^i,h_0^i,p')}{\gamma}\right) \end{array}$$

so that the previous expression for $\tilde{v}_i(s,p)$ can be written as

$$\tilde{v}_i(s,p) = -\frac{1}{\gamma}\ln\left[1-\int_\eta \max\left\{\begin{array}{c} J_1(\gamma,b(p,\eta),s+\Omega(\eta)), \\ J_2(\gamma,b(p,\eta),s+\Omega(\eta)), \\ J_3(\gamma,b(p,\eta),s+\Omega(\eta)) \end{array}\right\}H_i(d\eta,p)\right]$$

Claim 2

$$\int_\eta \max\left\{J_1(\gamma,b(p,\eta),s+\Omega(\eta)),J_2(\gamma,b(p,\eta),s+\Omega(\eta)),J_3(\gamma,b(p,\eta),s+\Omega(\eta))\right\}H_i(d\eta,p)$$

differentiable with respect to γ.

Proof. J_1, J_2 and J_3 are continuous functions and under the assumptions discussed in the paper, there is a unique threshold value for the signal η (that depends on s, p γ, and other parameters) that equates J_1 with J_2 and J_2 with J_3. Let $\eta^L(\gamma, p, s)$ and $\eta^H(\gamma, p, s)$ denote these thresholds. Note that, as J_1, J_2 and J_3 are differentiable with respect to γ, these thresholds are also differentiable by construction. It follows that the expression of the claim can be rewritten as

$$\int_{-\infty}^{\eta^L(\gamma,p,s)} J_1(\gamma, p', s') H_i(d\eta, p) + \int_{\eta^L(\gamma,p,s)}^{\eta^H(\gamma,p,s)} J_2(\gamma, p', s') H_i(d\eta, p) + \int_{\eta^H(\gamma,p,s)}^{\infty} J_3(\gamma, p', s') H_i(d\eta, p) \tag{11}$$

where $p' = b(p, \eta)$ and $s' = s + \Omega(\eta)$. that is differentiable with respect to γ. ■

Let $Q(\gamma, p, s)$ denote the object in equation (2).

Claim 3

$$\frac{dQ(\gamma,p,s)}{d\gamma} = \int_\eta \max\left\{\frac{\partial J_1(\gamma,p',s')}{\partial\gamma}, \frac{\partial J_2(\gamma,p',s')}{\partial\gamma}, \frac{\partial J_3(\gamma,p',s')}{\partial\gamma}\right\} H_i(d\eta, p)$$

Proof. Follows by applying Leibniz's rule to equation (11) and by noting that

$$\begin{aligned} J_1\left(\gamma, b(p,\eta^L), s+\Omega(\eta^L)\right)\tfrac{\partial\eta^L}{\partial\gamma} &= J_2\left(\gamma, b(p,\eta^L), s+\Omega(\eta^L)\right)\tfrac{\partial\eta^L}{\partial\gamma} \\ J_2\left(\gamma, b(p,\eta^H), s+\Omega(\eta^H)\right)\tfrac{\partial\eta^H}{\partial\gamma} &= J_3\left(\gamma, b(p,\eta^H), s+\Omega(\eta^H)\right)\tfrac{\partial\eta^H}{\partial\gamma} \end{aligned}$$

by construction of the thresholds η^L and η^H. ■

Now, the result follows by applying L'Hopital's rule for the case where $\gamma \to 0$ and by the results of Claim 3.

G Proof of Proposition 2

Pick two amount of accumulated credits s_1 and s_0. Without loss of generality assume that $s_1 > s_0$. Let $p_d^i(s_0) > 0$ be the dropout threshold associated with s_0 so that $V_i(a, s_0, p_d^i(s_0)) = W(a; h^w)$. As $V_i(a, s, p)$ increasing in credits s, $V_i(a, s_1, p_d^i(s_0)) > W(a; h^w)$. Finally, as $W(a; h^w)$ independent of p and $V_i(a, s, p)$ increasing in p, $p_d^i(s_1) < p_d^i(s_0)$. Note that if $p_d^i(s_0) = 0$, then the same argument implies that $p_d^i(s_1) = p_d^i(s_0) = 0$.

Next the proof for $p_t^A(s_1) > p_t^A(s_0)$ is provided (the proof for $p_t^C(s_1) < p_t^C(s_0)$ is almost identical). Let $p_t^A(s_0) < 1$ be the dropout threshold associated with s_0 so that

$$V_A(a, s_0, p_t^A(s_0)) = V_C(a, \theta^i(s_0), p_t^A(s_0))$$

Consider the case where the value V_i for $s = s_0$ do not include the transfer option. The next claim shows that, in this case, $V_A(a, s_1, p_t^A(s_0)) > V_C(a, \theta^A(s_1), p_t^A(s_0))$.

Claim 4 *$V_A(a, s_1, p_t^A(s_0)) - V_C(a, \theta^A(s_1), p_t^A(s_0)) > 0$.*

Proof. Let $s_1 = s_0 + \epsilon$ where $\epsilon \in (0, T^A - s_0]$. Let (to ease on notation) $\Upsilon_j^q = \{a, s_j, p_t^A(s_q)\}$ and $\Psi_j^q = \{a, \theta^A(s_j), p_t^A(s_q)\}$. Applying a Taylor Expansion of second order to both $V_A(a, s_1, p_t^A(s_0))$ and $V_C(a, \theta^A(s_1), p_t^A(s_0))$ around s_0 provides (the assumptions on the proposition guarantees that the value function is continuous and differentiable with respect to s),

$$V_A(\Upsilon_1^0) \approx V_A(\Upsilon_0^0) + \frac{\partial V_A(\Upsilon_0^0)}{\partial s}\epsilon + \frac{1}{2}\frac{\partial^2 V_A(\Upsilon_0^0)}{\partial s^2}\epsilon^2$$

and

$$V_C(\Psi_1^0) \approx V_C(\Psi_0^0) + \frac{\partial V_C(\Psi_0^0)}{\partial s}\frac{\partial \theta^A(s_0)}{\partial s}\epsilon + \frac{1}{2}\frac{\partial^2 V_C(\Psi_0^0)}{\partial s^2}\left[\frac{\partial \theta^A(s_0)}{\partial s}\right]^2\epsilon^2 + \frac{1}{2}\frac{\partial V_C(\Psi_0^0)}{\partial s}\frac{\partial^2 \theta^A(s_0)}{\partial s^2}\epsilon^2$$

Note that, by definition of $p_t^A(s_0)$, $V_A(\Upsilon_0^0) = V_C(\Psi_0^0)$. Moreover, by the optimality of the policy, $\frac{\partial V_A(\Upsilon_0^0)}{\partial s} = \frac{\partial V_C(\Psi_0^0)}{\partial s}$ and $\frac{\partial^2 V_A(\Upsilon_0^0)}{\partial s^2} = \frac{\partial^2 V_C(\Psi_0^0)}{\partial s^2}$ as when s is moved away from s_0 it is the case that both V_A and V_C coincide. Then, $V_A(a, s_1, p_t^A(s_0)) - V_C(a, \theta^A(s_1), p_t^A(s_0))$ reduces to

$$\begin{aligned} V_A(\Upsilon_1^0) - V_C(\Psi_1^0) \quad = \quad & \frac{\partial V_A(\Upsilon_0^0)}{\partial s}\left[1 - \frac{\partial \theta^A(s_0)}{\partial s}\right]\epsilon + \frac{1}{2}\frac{\partial^2 V_A(\Upsilon_0^0)}{\partial s^2}\left[1 - \left(\frac{\partial \theta^A(s_0)}{\partial s}\right)^2\right]\epsilon^2 \\ & -\frac{1}{2}\frac{\partial V_A(\Upsilon_0^0)}{\partial s}\frac{\partial^2 \theta^A(s_0)}{\partial s^2}\epsilon^2 \end{aligned}$$

Finally, as $V_A(a, s, p)$ increasing and convex in s, $\theta^A(s)$ increasing and concave and $\frac{\partial \theta^A(0)}{\partial s} < 1$, $V_A(\Upsilon_1^0) - V_C(\Psi_1^0) > 0$, which completes the proof. ■

Claim 5 *For any $p > \min\left\{p_d^A(s), p_d^C(s)\right\}$, the difference $V_A(a, s, p) - V_C(a, \theta^A(s), p)$ satisfy the single-crossing property.*

Proof. Follows from strict convexity of $V_A(a, s, p)$ and $V_C(a, s, p)$ as a function of p together with $V_C(a, s, 1) > V_A(a, s, 1)$ and $V_C(a, s, 0) < V_A(a, s, 0)$. ■

Assume now that $p_t^A(s_1) \leq p_t^A(s_0)$ (the proof follows by contradiction). Then, by the single crossing property $V_A(a, s_1, p_t^A(s_0)) - V_C(a, \theta^A(s_1), p_t^A(s_0)) < 0$ (Claim 5) which violates Claim 4. Then, it follows that $p_t^A(s_1) > p_t^A(s_0)$.

Note that the only case where $p_t^A(s_1) = p_t^A(s_0)$ is when both thresholds are inactive (that is, equal to 1). The proof in this case is trivial.