

# Chutes and Ladders in Post-Secondary Education

## Academic 2-year colleges as a Stepping Stone\*

Nicholas Trachter

University of Chicago

[trachter@uchicago.edu](mailto:trachter@uchicago.edu)

October 7, 2009

### Abstract

The high returns to graduating from a 4-year college conflicts with low enrollment and graduation rates. Using evidence from two panels of post-secondary education, this paper argues that incorporating 2-year colleges into a model of post-secondary educational choice together with 4-year colleges and work can reconcile this apparent conflict. The ex-post return to graduating from an academic 2-year college is low, but there is a moderate return to dropping out and a large return to transferring to a 4-year college: academic 2-year colleges act as a *stepping stone* in which agents learn about themselves in a cheaper and less demanding environment.

In the model, agents are initially uncertain about their innate ability to accumulate human capital. Pessimistic agents join the workforce, optimistic agents enroll in 4-year colleges and those in the middle enroll in academic 2-year colleges. Exams govern the accumulation of credits and produce information that can be used to update beliefs about ability, inducing dropouts and transfers. The model is consistent with the following stylized facts which are documented for both data sets: (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) the composition of those that transfer is weighted toward higher ability students.

A decomposition of returns show that the dropout and transfer option account for 90% of the full return to enrolling in an academic 2-year college while the dropout option explain 70% of the full return to enrolling in 4-year colleges. Full insurance would reduce enrollment in academic 2-year colleges from 17% to 5%, while enrollment in 4-year colleges would rise from 25% to 42%. The interaction of risk and option value is an important force in post-secondary education.

---

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

# 1 Introduction

The high returns to graduating from a 4-year college conflicts with low enrollment and graduation rates. Heckman et al. (2008a) evaluates the Internal Rate of Return<sup>1</sup> of the 4-year college investment option relative to work, to find that since 1960 Internal Rates of Return had been around 10% or higher depending on the cohort and different specifications of labor markets and taxes.<sup>2</sup> Judd (2000) combines CAPM techniques with the indivisibility of human capital to compare the return to 4-year college graduation with assets of similar risk to find an excess of return to the college investment option. Cunha et al. (2005), using data from NLSY/1979, extend the analysis to evaluate the Internal Rate of Return for the marginal student, the agent with the lowest observable measures of ability that enrolls in 4-year college, to find an unexplained wedge in returns.<sup>3</sup>

Using evidence from two panels of post-secondary education, NLS-72 and NLS-92<sup>4</sup>, this paper argues that incorporating 2-year colleges, or community colleges, into a model of post-secondary educational choice together with 4-year colleges and work can reconcile this apparent conflict. Estimating a mincer regression that allows for heterogeneity across types of institutions, graduation premium and by relaxing the assumption of linearity on years of education and experience, Internal Rates of Return are recomputed to find that the wedge in returns is explained by 2-year colleges.

The ex-post return to graduating from an academic 2-year college<sup>5</sup> is low, but there is a moderate return to dropping out and a large return to transferring to a 4-year college: academic 2-year colleges act as a *stepping stone* in which agents learn about themselves in a cheaper and less demanding

---

<sup>1</sup>The rate of return that makes the discounted value of two investment decisions to be equalized.

<sup>2</sup>Belzil and Hansen (2002) find Internal Rates of Return of 7% for the cohort of NLSY/1979.

<sup>3</sup>Can also be found in the Handbook of Economics of Education (see Heckman et al. (2006)).

<sup>4</sup>NLS-92 is part of NELS:88. NLS-92 refers to the study of post-secondary patterns of high-school graduates at 1992.

<sup>5</sup>2-year colleges are, broadly speaking, a combination of academic 2-year colleges and vocational school. The former has as a goal transferring students to 4-year colleges while the later main goal is to produce a labor-force by providing terminal programs.

environment. In the model, 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 college. During tenure as students, agents are presented with exams, which govern the accumulation of credits and produce information that can be used to update beliefs about ability, inducing dropouts and transfers. The sequential process of education was pointed out first in [Comay et al. \(1973\)](#). [Altonji \(1993\)](#) computes Internal Rates of Return for a simple sequential model where agents are uncertain about future income flows and thus evaluation of expectations induce dropout behavior. [Heckman and Urzua \(2008\)](#) and [Stange \(2007\)](#) estimate models of educational choice where students are allowed to drop out.

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 college and carry with them a proportion of their stock of credits, implying that academic 2-year colleges acts as a *stepping-stone* towards more demanding environments, namely, 4-year colleges. Further, the model has features of bandit models as students learn about their innate ability to accumulate human capital. [Jovanovic and Nyarko \(1997\)](#) evaluates the predictive power of both models in terms of job mobility and find that there is some evidence favoring a combination of both. The same it rue with the educational ladder where 4-year college plays the role of the step above academic 2-year colleges.

An important aspect of the model is its tractability that 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 initial belief of high-school graduates, to evaluate the model's predictions. The model is consistent

with the following stylized facts which are documented for both data sets: (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) the composition of those that transfer is weighted toward higher ability students.

A decomposition of returns show that the dropout and transfer option account for 90% of the full return to enrolling in an academic 2-year college while the dropout option explain 70% of the full return to enrolling in 4-year colleges. Full insurance would reduce enrollment in academic 2-year colleges from 17% to 5%, while enrollment in 4-year colleges would rise from 25% to 42%. The interaction of risk and option value is an important force in post-secondary education.

### **How has the literature reconciled the low enrollment and graduation rates at 4-year colleges with high rates of return?**

To reconcile low enrollment with high returns to 4-year college graduation, [Cunha et al. \(2005\)](#) and [Carneiro et al. \(2003\)](#) argue in favor of non-pecuniary costs of education.<sup>6</sup> Their findings show that these costs play a large role of enrollment decisions. These costs are viewed in a broad way and can be understood as a combination of tastes for college-going, tastes for studying, etc. In their model, agents are assumed to be risk neutral so risk aversion is also part of the non-pecuniary costs of education.

To explain high college dropout rates [Heckman and Urzua \(2008\)](#) and [Stange \(2007\)](#) extend the model to allow for a sequential revelation of information and where students are allowed to dropout as a result of an optimal re-evaluation of expectations. In their setups, students learn about their own ability and non-pecuniary costs of education. Estimates show that learning about

---

<sup>6</sup>Also known as psychic costs.

non-pecuniary costs is important in explaining dropout behavior. [Stinebrickner and Stinebrickner \(2008\)](#), using a panel designed specifically for analyzing dropout behavior, show strong evidence against non-pecuniary costs of education explaining dropouts. The paper shows that bad grades explain both dropouts and claims made by students about disliking college.

## 2 Patterns of Postsecondary Education for the Class of 1972

This section presents statistics on Postsecondary educational patterns and returns based on the *National Longitudinal Study of 1972* or NLS-72.<sup>7</sup> The unit of analysis are high-school seniors that join the workforce directly (with no spells of Postsecondary education) or join a Postsecondary institution with no discontinuities in their educational spells.<sup>8</sup> 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.

2-year colleges originated in the late 19th century when W.R. Harper, founding president of The University of Chicago, ideated a plan to teach students lower division "preparatory" material in order to increase participation in higher education without compromising existing 4-year colleges.<sup>9</sup>

Between WWI and WWII there was an unsatisfied demand for technified workers and 2-year colleges

---

<sup>7</sup>The choice of NLS-72 over other data sets is not an arbitrary one. *High school and Beyond*, or HS&B, follows a cohort from 1982 to 1990. *National Education Longitudinal Study of 1988*, or NELLS:88, follows a cohort from 1992 to 2000. From now on, this data set will be described as NLS-92. Relative to these data sets, NLS-72 presents longer horizon wage information (13 years vs. 8 years after high-school graduation in the newer data sets). Also, the design of the questionnaire of NLS-72 included questions regarding the type of 2-year college education at any point in time (broadly speaking, 2-year colleges are a combination of academic 2-year colleges and vocational school), while these questions were not available in the newer data sets. Last, NLS-72 has a more detailed analysis of the cost structure of post-secondary education. An alternative is to use *National Longitudinal Survey of the Youth*, or NLSY, that presents better life-cycle earnings information but that requires a lot of data mining (in particular, there is no easy way to disentangle vocational school from academic 2-year colleges). Further, many Community colleges have extended their scope to offer both types of programs making increasingly difficult to distinguish. [Section 7](#) collapses together vocational school and academic 2-year colleges and compares dynamic and enrollment patterns for both NLS-72 and NLS-92.

<sup>8</sup>Discontinuous spells are treated as educational histories that include periods of work.

<sup>9</sup>The Joliet Junior college was the first 2-year college in the U.S. and still functions in the Chicago area.

started to expand their scope to prepare a labor-force by providing specialized terminal programs. Is in this era where the distinction between academic-year colleges and vocational school arises. In this paper, academic 2-year colleges are understood as institutions where the transfer function is the main goal (even though they also provide terminal degrees) while vocational schools are understood as institutions where the main goal is to produce a labor force. Kane and Rouse (1999) presents a more detailed analysis and description of the history of 2-year colleges.

	share	j								
		voc. school			ac. 2-year c.			4-year c.		
		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

Table 1: **Transitional dynamics for first and second educational spells after high-school graduation. Source: NLS-72** 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.

Table 1 presents the postsecondary educational histories for the Senior Class of 1972. Individuals are faced with an initial enrollment choice between 4-year colleges, academic 2-year colleges, vocational school 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 kind of educational institution.<sup>10</sup> Third, a student can graduate ( $G$ ) and join the workforce.<sup>11</sup> The proportion of students that transfer more than once is negligible and so here the analysis is reduced to students that transfer only once. Students that transfer can end their second spell of education in two ways: obtaining a degree or becoming a dropout.

Only half of the senior class of 1972 pursue higher education. Among them, nearly 20% enroll in vocational school, around 30% enroll in academic 2-year colleges and the rest enroll in 4-year

<sup>10</sup>Within transfers (for example: 4-year college to 4-year college) are not understood here as transfers.

<sup>11</sup>Graduates that transfer are viewed as transfers here.

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 school and academic 2-year colleges than in 4-year colleges.<sup>12</sup> Transfer rates are important in academic 2-year colleges as around 32% of students that initially enroll in this type of institution eventually transfer to 4-year colleges. Finally, only 4-year colleges graduate a large percent of their students. Also, note that the graduation rate at 4-year college 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. [Section 7](#) contrasts NLS-72 with NLS-92 to find similar patterns.

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 (income stream that a student 'lose' by being at school) and direct costs of education that include tuition (and associated fees) and housing ([Table 2](#)). As seen in the table, 4-year colleges cost twice as much as academic 2-year colleges, providing one reason for why students might enroll in academic 2-year colleges.

## 2.1 Returns to Education

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 large ongoing literature that accounts for non-linearities in years of education (see [Grubb \(1993\)](#), [Heckman et](#)

---

<sup>12</sup>Dropout rates at vocational school are inflated as vocational school also includes 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.

	Tuition	Tuition + R&B
4-year college	3420.73 (2510.25)	5038.8 (4010.8)
academic 2-year college	1811.36 (1355.1)	2729.65 (4665.35)
vocational school	3803.84 (6131.9)	5904.74 (9615.25)

Table 2: **Differences in Cost of Education (NLS-72)**. Missing Values were imputed by running a Cost regression and imputing missing values through observables. The values are measured in 1984 dollars.

al. (2006), Heckman et al. (2008b) 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 here it is assumed that different educational histories affect lifetime earnings in different ways. Further, as has been already discussed in the literature, the analysis here breaks the additive form (in the log version in the Mincer regression) of years of education and experience by estimating a growth equation.

Table 3 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 school and 4-year colleges provides higher wages relative to dropping out. The same idea does not apply in academic 2-year colleges as the return to becoming a dropout is higher than the return from graduation.

Table 4 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,  $\alpha_G^V$  is around half the growth rate of 4-year college graduates  $\alpha_G^C$ . This fact, together with the results from Table 3 reads as follows: graduation at vocational school provides a higher wage and 4-year college provide steeper profiles of wages, while graduation at academic 2-year colleges is dominated.

description		Version A	Version B	
			(1)	(2)
$\varpi_C^D$	drop at 4-year C.	0.24 (0.042)	0.22 (0.041)	-
$\varpi_A^D$	drop at Ac. 2-year C.	0.09 (0.045)	0.077 (0.044)	-
$\varpi_V^D$	drop at Voc. school	0.072 (0.046)	-	-
$\varpi_C^G$	graduation at 4-year C.	0.304 (0.04)	0.286 (0.04)	0.2122 (0.037)
$\varpi_A^G$	graduation at Ac. 2-year C.	0.015 (0.1)	0.149 (0.104)	0.0445 (0.104)
$\varpi_V^G$	graduation at Voc. school	0.284 (0.15)	-	-

Table 3: **Mincer Regression (NLS-72)**. Dependent Variable: log of wages in 1985. Independent Variables: graduation and dropout dummies for each type of institution and controls for characteristics. Version B collapse together vocational school and workforce.

Using the results of [Table 3](#), [Table 4](#), the cost of education ([Table 2](#)) and assuming a finite lifetime of 47 years (retirement or death at age 65), [Table 5](#) presents the Internal Rate of Return (IRR) for the average student with a particular educational paths relative to joining the workforce directly after high-school graduation (see [Section F](#) for the details of the calculations). [Table 18](#), [Table 19](#) and [Table 20](#) (see [Section H](#) in the Appendix) present the average time spent in each institution for a given educational path (that is, conditioning on initial enrollment status, dropout and transfer behavior).

Inspection of [Table 5](#) provide that enrollment in 4-year colleges provides the highest return and is mostly driven by the return for graduates.<sup>13</sup> 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

<sup>13</sup>The results are in line with [Cunha et al. \(2005\)](#), [Heckman et al. \(2008a\)](#), and [Belzil and Hansen \(2002\)](#), among others.

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

Table 4: **Growth Regression (NLS-72)**. Dependent Variable: average growth rate of wages between 1979 and 1985.

	Voc. school	Ac. 2-year C.	4-year C.
Graduation	8.07	1.84	9.12
Dropout	2.37	5.31	7.73
Transfer to V.S.	-	2.48	1.52
Transfer to Ac. 2-year	3.76	-	3.4
Transfer to 4-year	8.01	8.24	-
Enrollment	2.93	5.96	8.2

Table 5: **Internal Rates of Return (NLS-72)**. The cost of education includes R&B for 4-year colleges (hybrid case). All the values are in percentage points. [Section F](#) presents the details of the calculations. The excluded group are agents that join the workforce directly.

2-year colleges. Finally, note that IRR conditional on initial enrollment choice are ordered: low in vocational school, average in academic 2-year colleges and high in 4-year college.

See the return to graduation at 4-year colleges. This value is similar to what has been found by [Heckman et al. \(2008a\)](#), [Cunha et al. \(2005\)](#) and [Belzil and Hansen \(2002\)](#) and suggest that returns are too high if compared to low enrollment and graduation rates. If we extend the definition of

4-year college to include dropout and transfer the mean return decrease from 9.12% to 8.2%. Now, note that 73% of the mean return to 4-year college enrollment can be explained by the return for those that initially enroll in academic 2-year college. How to explain the wedge between 4-year colleges and academic 2-year colleges? To answer this question, a model of educational choice will be evaluated to show that this wedge in measured returns is possible even though ex-ante the marginal student is indifferent between both options.

## 2.2 Sorting in Initial Enrollment

Students that enroll in 2-year colleges had 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 6 present 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 school and academic 2-year colleges. Moving from left to right through the table shows that there is some evidence of an ordered enrollment choice. For example, see the rank in high-school class (i.e. Rank). The rank decreases monotonically with the enrollment choice. Same result can be found for Socioeconomic status of family and educational level of father.

Ordered returns to enrollment (see Table 5) together with the evidence presented in Table 6 suggest that the initial enrollment choice is ordered as follows: work, vocational school, academic 2-year colleges, and 4-year colleges. Table 21 (see Section H in the Appendix) presents the results of an ordered probit regression of the initial enrollment choice on a vector  $X$  of observable measures of ability. Let  $\beta$  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

		Full	work	V.S.	Ac. 2-year	4-year C.
Male		0.5 (0.5)	0.492 (0.5)	0.4 (0.49)	0.54 (0.5)	0.524 (0.5)
Black		0.105 (0.31)	0.108 (0.31)	0.126 (0.33)	0.085 (0.28)	0.105 (0.3)
Socio. Status:	Low	0.307 (0.46)	0.415 (0.49)	0.294 (0.45)	0.192 (0.39)	0.166 (0.37)
	Medium	0.495 (0.5)	0.506 (0.5)	0.582 (0.49)	0.55 (0.49)	0.407 (0.49)
	High	0.197 (0.39)	0.078 (0.27)	0.122 (0.33)	0.253 (0.43)	0.426 (0.49)
Education of Father:	<HS	0.394 (0.48)	0.516 (0.5)	0.391 (0.48)	0.289 (0.45)	0.215 (0.41)
	HS	0.324 (0.47)	0.323 (0.46)	0.398 (0.49)	0.344 (0.47)	0.288 (0.45)
	4-year C. (no degree)	0.148 (0.35)	0.108 (0.31)	0.145 (0.35)	0.211 (0.41)	0.19 (0.39)
	4-year C. graduate	0.132 (0.34)	0.051 (0.22)	0.064 (0.24)	0.155 (0.36)	0.305 (0.46)
Rank		0.418 (0.3)	0.495 (0.31)	0.435 (0.28)	0.395 (0.27)	0.271 (0.24)

Table 6: **Summary Statistics for measures of ability (NLS-72)**. Rank=rank in high-school class. Socio-Status: Socioeconomic Status of Family at moment of high-school graduation.

institutions. The reference column in [Table 21](#) 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 sorting in initial enrollment, [Table 7](#) produces the mean and standard deviation (in the cross-section) of  $X'\hat{\beta}$  across the different alternatives. See Version A in the first row of the table. The measure of ability  $X'\beta$  is unit less as it is just an ordinal representation of ability measures. Start with high-school graduates that join the workforce - labeled as work in table - and move upwards across enrollment options. The mean value for the measure of ability increase monotonically with the enrollment options.

Version	work	Voc. school	Ac. 2-year C.	4-year C.
A	-1.782 (0.567)	-1.609 (0.571)	-1.348 (0.604)	-0.972 (0.649)
B		-1.426 (0.572)	-1.013 (0.609)	-0.633 (0.654)

Table 7: **Evidence on Sorting: Measure of Ability (NLS-72)**. Constructed from Ordered Probit Estimation (see Table 21).

### 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, are academic 2-year colleges. The second step is 4-year colleges.

In contrast with the characterization of a ladder discussed above where agents move on once their 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 (usually associated with the pure stepping stone story) imply that agents should enroll first in the lower step of the ladder, as it provides a less riskier environment for learning and experimentation (through the lower cost of education and the shorter time to graduation). Bandit models suggest that students should enroll in the harder step - the last step - as the learning technology provides more information about innate ability (classes in 4-year colleges are harder than in academic 2-year colleges).

### 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, that can either be low or high. Let  $\mu$  denote the ability level, with  $\mu \in \{0, 1\}$ , where  $\mu = 0$  denotes low ability. The ability level  $\mu$  is not observable by the agent. Instead, a high-school graduate inherits a signal about her true type, denoted by  $\vartheta \in [0, 1]$ . Let  $j_\mu(\vartheta)$  denote the density of signal  $\vartheta$  conditional on the true ability level of the agent being  $\mu$ , with  $j_0 \rightarrow [0, 1)$  and  $j_1 \rightarrow (0, 1]$ . With the information at hand, a high-school graduate generate a subjective belief about her own true ability level  $p_0 \in [0, 1]$ , where  $p_0 = Pr(\mu = 1)$ .

At any period in time an agent can either be working, studying at 4-year college (from now on college or plain  $C$ ) or at academic 2-year colleges (from now on academic or plain  $A$ ). Let  $i \in \{A, C\}$  denote the type of institution. The cost of education per period of schooling (includes tuition, room and board, fees, books, etc.) is denoted by  $\tau^i$ , with  $\tau^C > \tau^A$ . An student graduates from institution  $i$  after she accumulates  $T^i$  credits, with  $T^C > T^A$ . The evolution of credits is closely tied to signals that arrive during tenure as student by the agent. Let  $\eta$  denote the signal with PDF given by  $f_i(\eta|\mu)$ .

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

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

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  $\eta$  and credits  $s$ ,

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

with

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

with  $\Omega'(\eta) \geq 0$  so that the evolution of credits is a non-decreasing function on the received signal.<sup>14</sup>

The functional form chosen for  $\Omega(\eta, s)$  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  $\eta$ .

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

$$\theta^i(s) : s \times i \rightarrow [0, T^{-i}]$$

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,

$$\mathbb{E} \left\{ \sum_0^{\infty} \left( \frac{1}{1+r} \right)^t \left( \frac{e^{-\gamma c_t} - 1}{-\gamma} \right) \middle| F_0 \right\}$$

---

<sup>14</sup>Obtaining a C or an A in a particular subject provides the same accumulation of credits but a different re-evaluation of own ability.

where  $F_0 = \{p_0, a_0\}$  and  $\gamma$  is the coefficient of Constant Absolute Risk Aversion (CARA).

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 institution where the agent graduated (highest degree) and the third for her true ability level.<sup>15</sup> Further, the 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  $\mu$ . That is, for any talent level, graduation at 4-year college implies higher wage profiles than graduation at academic 2-year colleges 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. The assumption of no borrowing constraints is consistent with [Cameron and Heckman \(2001\)](#), [Cameron and Taber \(2004\)](#) and [Keane and Wolpin \(2001\)](#), that found no evidence in favor of constraints for the NLSY.<sup>16</sup>

Let  $V_i(a, s, p)$  denote the value for a student currently enrolled in  $i$  with asset level  $a$ , amount

---

<sup>15</sup>In the current setup dropouts do not enjoy higher wage profiles. That is, increase 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  $s$ .

<sup>16</sup>Recent studies found evidence in favor of borrowing constraints (see [Belley and Lochner \(2008\)](#) and [Lochner and Monge-Naranjo \(2007\)](#)).

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

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

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 wage  $h$  faces the following problem:

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

where  $a'$  is

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

That is, the worker has to decide her consumption in the current 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 next proposition summarizes the solution to this simple problem.

**Proposition 1** *The value for a worker with asset level  $a$  and wage profile  $h$  is*

$$W(a; h) = -\frac{1+r}{\gamma r} e^{-\gamma(ra+h)} + \frac{1+r}{\gamma r} \quad (2)$$

**Proof.** See Section A of the Appendix. ■

One of the goals of the paper is to evaluate the effect of insurance on the allocation. Noting that  $\gamma = 0$  provides the same allocation as full insurance, the next corollary presents the solution to the worker's problem under risk neutrality.

**Corollary 1** *The value for a risk neutral worker is linear in assets and wage profile. That is,*

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

**Proof.** Follows directly by applying l'hospital rule to equation (2). ■

### 3.2 The problem of a student

The arrival of information through the signal  $\eta$  generates updating of beliefs by the student. Let  $p' = b(\eta; p)$  denote the posterior that depends on the prior  $p$  and the signal that arrived  $\eta$ . 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 signals. Any signal that arrive 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

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

as the CDF that accounts for this uncertainty.

**Lemma 1** *For a given prior  $p$ ,  $H_i(\eta, p)$  is a well-defined CDF.*

**Proof.** Straightforward as  $F_i(\eta | \mu)$  is a CDF and  $p \in [0, 1]$ . ■

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

$$V_i(a, s, p) = \max_{c, a'} \frac{e^{-\gamma c} - 1}{-\gamma} + \frac{1}{1 + r} \int_{\eta} \tilde{V}_i(a', s', p') H_i(d\eta, p) \quad (3)$$

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. 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 so the options are reduced to graduation, drop or transfer. Let  $\mathbb{I} = 1$  if  $s' < T^i$  and  $= 0$  otherwise.  $\tilde{V}_i(a', s', p')$  is equal to

$$\max \{W(a'; h^w), \mathbb{I}V_i(a', s', p') + (1 - \mathbb{I})[p'W(a'; h_1^i) + (1 - p')W(a'; h_0^i)], V_{-i}(a', \theta^i(s'), p')\} \quad (4)$$

**Lemma 2** *A student currently enrolled in institution  $i$  with accumulated credits  $T^i$  will never drop of the current institution.*

**Proof.** 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. ■

The timing of the problem is the following. Given an institutional choice  $i$ , a given period can be decomposed into two subperiods. In the first subperiod, a student chooses her consumption and level of assets for next period given her expectations about future income streams. In the second subperiod, the student receives the signal  $\eta$ , 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 next proposition summarizes the solution to the problem.

**Proposition 2** *The value for a student enrolled in institution  $i$  with asset level  $a$ , schooling years  $s$  and prior  $p$  is*

$$V_i(a, s, p) = -\frac{1+r}{\gamma r} e^{-\gamma(ra+v_i(s,p))} + \frac{1+r}{\gamma r} \quad (5)$$

where  $v_i(s, p)$  solves

$$v_i(s, p) = \frac{\tilde{v}_i(s, p) - r\tau^i}{1+r} \quad (6)$$

and  $\tilde{v}_i(s, p)$  solves the recursive equation:

$$\tilde{v}_i(s, p) = -\frac{1}{\gamma} \ln \left[ \int_{\eta} -\max \left\{ \begin{array}{l} -e^{-\gamma h^w}, -e^{-\gamma v_{-i}(\theta^i(s'), p')}, \\ -\left( \mathbb{I} \left( p' e^{-\gamma h_1^i} + (1-p') e^{-\gamma h_0^i} \right) + (1-\mathbb{I}) e^{-\gamma v_i(s', p')} \right) \end{array} \right\} H_i(d\eta, p) \right] \quad (7)$$

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

**Proof.** See Appendix (Section B). ■

The value  $\tilde{v}_i(s, p)$  is the consumption equivalent of the continuation value.

**Proposition 3**  $V_i(a, s, p)$  increasing and convex in  $p$  and  $s$ .

**Proof.** See Section D in the Appendix. ■

The next lemma characterizes the solution when  $\gamma \rightarrow 0$ .

**Lemma 3** The value of initial enrollment at institution  $i$  for a risk neutral agent is given by

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

with

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

where

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

**Proof.** See Appendix (Section C). ■

### 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 of transferring.

**Proposition 4** *For any amount of credits  $s$ , the optimal policy is independent of the asset level  $a$ .*

*Further, the optimal policy is a collection of dropout and transfer thresholds,*

$$\{p_d^i(s), p_t^i(s)\}_{s \in [0, T^i]}$$

**Proof.** The optimal policy is independent on the asset level  $a$  as every value function shares the common term where  $a$  appears,  $e^{-\gamma r a}$ . For any amount of credits  $s$ , a student compares the value of continuation with the alternatives (transfer and drop). As the value functions are linear in  $e^{-\gamma r a}$ , the optimal policy that arise from the comparison of value functions is independent of asset level  $a$ . Finally, as the amount of accumulated credits  $s$  affect both the amount of credits that can

be transferred and also the likelihood and time to graduation the optimal policy is a function of accumulated credits  $s$ . ■

**Assumption 2** drives the optimal policy not only at time 0 but also as credits accumulate. Consider the case  $T^i$  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.

**Proposition 5** *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 next proposition evaluates the interaction of accumulated credits  $s$  and the evolution of the thresholds.

**Proposition 6** *If  $\theta^i(s)$  concave and  $\frac{\partial \theta^i(0)}{\partial s} < 1$ , for any two level of accumulated credits  $s_1$  and  $s_0$  with  $s_1 > s_0$  it is the case that*

$$\begin{cases} p_d^i(s_1) \leq p_d^i(s_0) \\ p_t^A(s_1) \geq p_t^A(s_0) \\ p_t^C(s_1) \leq p_t^C(s_0) \end{cases}$$

**Proof.** See **Section E** in the Appendix. ■

The proposition states that as credits accumulate, the likelihood of transferring or dropping out decreases as the terminal pay-off 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 postsecondary education can not be applied. To value the option and compute returns, define  $\Sigma_i(p)$  as the value-added (or payoff) of enrollment in institution  $i$  relative to joining the workforce directly after high-school graduation for an agent with prior  $p$ .<sup>17</sup>  $\Sigma_i(p)$  it 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

---

<sup>17</sup>Miao and Wang (2007) uses a similar approach to value an investment project where the income flow is uncertain and the risk is uninsurable.

return to enrollment at institution  $i$  relative to joining the workforce  $R_i(p)$  is defined as

$$R_i(p) \equiv \frac{\Sigma_i(p)}{\frac{1+x}{r} h^w} \quad (8)$$

### 3.5 Exams: Experimentation and Evolution of Credits

The signal  $\eta$  play two different roles in the model. First, the arrival of information through signal  $\eta$  produces updating of belief  $p$  as the signal conveys information regarding the likelihood of the true talent level of the student. Under this definition, the signal  $\eta$  accounts for grades in exams, in subjects, problem sets, overall experience as a student, etc. The second role of the signal  $\eta$  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 it will be assumed here that the signal  $\eta$  is merely the mean of the grades in a quarter obtained by a student.<sup>18</sup> The set of possible values of  $\eta$  is simply the set of possible grades. For simplicity assume three possible grades:  $\{F, N, E\}$ . That is, a student can fail, pass or excel in a particular period. Let  $q_\mu^i(\eta)$  denote the probability of each event.

## 4 Parametrization

The model explores the interaction of academic 2-year colleges and 4-year colleges. The evidence obtained from NLS-72 shows that vocational school (that is excluded from the model analysis) can be collapsed together with the workforce as little interaction occurs between vocational school and other types of institutions (see [Table 1](#)) and the initial sorting analysis (see [Table 21](#) and [Table 7](#))

---

<sup>18</sup>It is possible to relax this assumption by choosing functional forms for the signal  $\eta$  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.

places vocational school below academic 2-year colleges. Version B in all of the tables is for the case where work and vocational school are collapsed together. Further, (2) on the tables is for the cases where increase in wages only occur upon graduation (as in the model).

The operator  $\theta^i(s)$  is simplified to be of the multiplicative form,

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

Parametrization of the model requires to chose values and functional forms for different objects: risk-free rate  $r$  (1 parameter), risk aversion parameter  $\gamma$  (1 parameter), length of education  $T^A$  and  $T^C$  (2 parameters), wage structure (5 parameters), cost of education  $\tau^A$  and  $\tau^C$  (2 parameters), transfer of credits  $\theta^A$  and  $\theta^C$  (2 parameters), evolution of credits  $\Omega(\eta)$  and density function for  $\eta$ . Exams are simplified to have only three, mutually exclusive, grades: fail (F), neutral (N) and excellent (E) with corresponding probabilities given by  $q_\mu^i(\eta)$  (8 parameters). Further, the evolution of credits is chosen to be as follows:  $\Omega(F) = 0$ ,  $\Omega(N) = \Omega(E) = 1$ . Overall, 21 parameters have to be chosen.

The time period is chosen to be a quarter, and so  $T^A = 8$  and  $T^C = 16$  (so 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% that 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 standardized by  $h^w$ , so that  $h^w = 1$ .<sup>19</sup> The mincer regression shown in [Table 3](#) suggests that the average return for 4-year college graduation is around 22% and around 5% for academic 2-year college graduation. The wage structure is chosen to be as follows:  $h^A(0) = 1.03$ ,  $h^A(1) = 1.08$ ,  $h^C(0) = 1.04$  and

---

<sup>19</sup>The mean wage in 1985, in 1984 dollars, for agents with no degrees was 17740.63 (11618.55).

$h^C(1) = 1.22$ . academic 2-year colleges are located in every city and town while 4-year colleges are scarce. As so, the cost of education includes housing for 4-year colleges and do not include housing for academic 2-year colleges as students there can live with their parents. The standardized 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  $\gamma$  is hard to identify and the literature didn't spend much time estimating risk aversion parameters using CARA utility functions. There is a whole string of literature in asset pricing starting with [Mehra and Prescott \(1985\)](#) that argues that the CRRA risk aversion parameter  $\sigma$  lies between 4 and 10. Using the definition of relative risk aversion it is possible to relate  $\sigma$  and  $\gamma$ ,

$$\gamma c = \sigma$$

In the model presented here consumption level  $c$  has a lower bound given by  $ra + h^w \geq 1$  so  $\gamma < 10$ . Here  $\gamma$  is chosen to be equal to 8.

[Figure 1](#) 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 occur, on average, after the completion of the first year of education.<sup>20</sup> Then,  $\theta^A$  is chosen to be  $\frac{1}{2}$ .

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 as it can be seen in [Table 1](#). [Figure 2](#) 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 college spend more time in school prior to graduation (4.5 years vs. 3.84 years). The evidence suggests that  $\theta^C = 0$ .

The remaining parameters,  $q_\mu^i(\eta)$  are estimated using a Simulated Method of Moments. Sim-

---

<sup>20</sup>Students transfer before obtaining a degree or completing the course-work at academic 2-year colleges. Only 12.5% of students that transfer from  $A$  to  $C$  in the NLS-72 sample hold a degree.

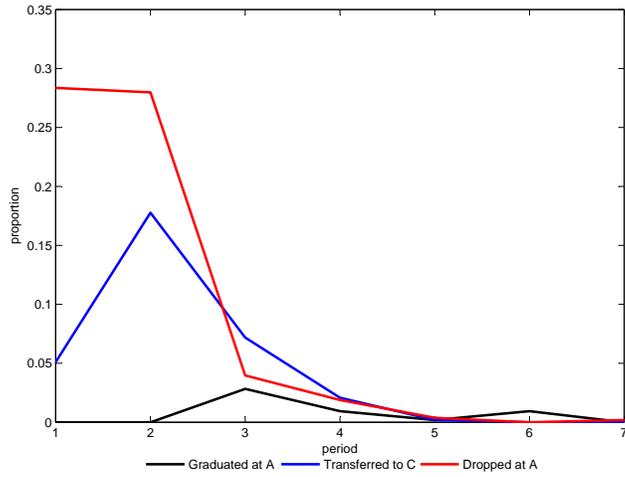


Figure 1: **Transitional dynamics for students initially enrolled in academic 2-year colleges. Source: NLS-72.**

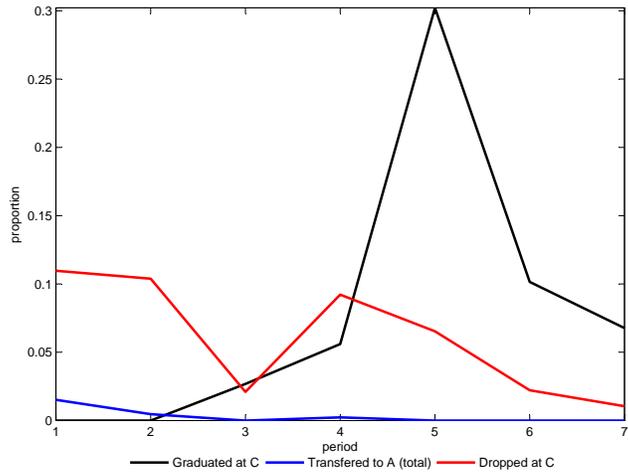


Figure 2: **Transitional dynamics for students initially enrolled in 4-year colleges. Source: NLS-72.**

ulating the model requires to solve numerically the model, produce transition probabilities and simulate priors. The prior  $p_0$  can be produced in different ways. One way comprehends specifying a functional form for  $j_\mu(\vartheta)$  and thus producing  $p_0$  by Bayes' rule. This approach implies that also

the parameters of  $j_\mu(\vartheta)$  have to be estimated. An alternative is to drive the estimation more heavily in the data. Let

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

where  $X$  is a vector that includes all the observable characteristics of high-school graduates that are correlated with the talent level of the agent and  $\beta$  is the vector of factor loadings, identified by the ordered probit for the initial choice of agents presented in [Table 21](#). The parameter  $\iota$  is a parameter that can not be identified directly from data that acts as a re-scaling parameter. The scale of the prior  $p_0$ , even though not important for the enrollment pattern (the only thing that matters here is the ordering) plays an important role for transfer and dropout behavior as the initial prior  $p_0$  is an unbiased estimator of the actual probability of an agent being high talented. This parameter will be jointly calibrated with  $q_\mu^i(\eta)$ .

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  $\beta$  can be estimated by an ordered probit on the initial choice ([Table 21](#) - Version B) and then upper and lower bounds for  $\varepsilon$  can be computed using  $X'\hat{\beta}$  and the choice of the agent. Further the cutoffs shown in the Table are monotonic transformations of the threshold for  $p$ . The exponent on  $p_0$ , the parameter  $\iota$ , can not be identified from data.

Indirect inference, through Simulated Method of Moments, is used to estimate the remaining 9 parameters (the 8 learning parameters and  $\iota$ ). The chosen moments are (see [Table 1](#)): (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 college that dropped-out, (5) proportion of students initially enrolled in academic 2-year college that transfer to 4-year colleges, (6) proportion of students initially enrolled in academic 2-year college that graduate at 2-year colleges (highest degree), (7) proportion of students initially enrolled in 4-year college that dropped-out, (8) proportion of students initially enrolled in 4-year college that transfer to 4-year colleges, (9) proportion of students initially enrolled in 4-year college that graduate at 4-year college, (10) mean wage for academic 2-year college graduates after 1<sup>st</sup> spell of education, and (11) mean wage for 4-year college graduates after 1<sup>st</sup> spell of education. The estimated parameters are presented in [Table 8](#). Note that  $\frac{q_1^A(F)}{q_0^A(F)} < \frac{q_1^C(F)}{q_0^C(F)}$  and  $\frac{q_1^A(E)}{q_0^A(E)} < \frac{q_1^C(E)}{q_0^C(E)}$  which depicts an important feature of Postsecondary education: Classes at academic 2-year colleges are easier than in 4-year colleges. A low grade in academic 2-year college induces a stronger re-evaluation of beliefs than in 4-year colleges. In the same way, a high grade induces a stronger re-evaluation of beliefs in 4-year colleges than in academic 2-year colleges.

$q_1^A(F) = 0.00001$	$q_0^A(F) = 0.23$	$q_1^A(E) = 0.13$	$q_0^A(E) = 0.075$
$q_1^C(F) = 0.01$	$q_0^C(F) = 0.22$	$q_1^C(E) = 0.18$	$q_0^C(E) = 0.001$
$\iota = 2$			

Table 8: **Parameters estimated by Simulated Method of Moments.**

A drawback of the model's simplicity can be seen in the inability to match completely the moments from the data. [Table 9](#) presents the value of the moments in data and the simulated values by the model. In particular, the model over-estimates the proportion of students that transfer from *A* to *C* compared to the amount that drop out.

The evolution of thresholds as a function of accumulated credits is presented in [Figure 3](#). As

% of High-school graduates that	Data	Model
join workforce	59.4	57.3
enroll in $A$	15.2	16.95
enroll in $C$	25.4	25.74
drop at $A$ (1 <sup>st</sup> spell)	9.57	5.09
transfer from $A$ to $C$ (1 <sup>st</sup> spell)	4.86	10.49
graduate at $A$ (1 <sup>st</sup> spell)	0.77	1.37
drop at $C$ (1 <sup>st</sup> spell)	10.41	10.37
transfer from $C$ to $A$ (1 <sup>st</sup> spell)	0.51	0.2
graduate at $C$ (1 <sup>st</sup> spell)	14.48	15.17
Mean Wage for Graduates at (1 <sup>st</sup> spell)		
academic 2-year college	0.0445	0.045
4-year college	0.212	0.215

Table 9: Moments in data (NLS-72) and Model.

discussed in Proposition 6, the inaction region in both academic 2- and 4-year colleges increase with credits.<sup>21</sup>

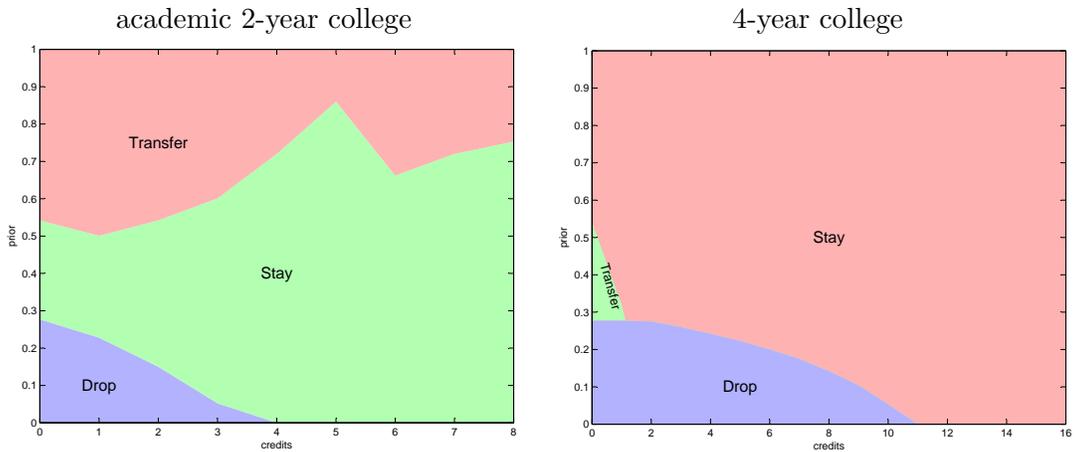


Figure 3: Dropout and Transfer thresholds as a function of accumulated credits  $s$ .

<sup>21</sup>The non-monotonicity in the transfer threshold in Ac. 2-year college is due to the discreteness of credits  $s$ .

## 5 Model's Predictions

The prior  $p_0$  is positively correlated with the measure of talent  $X'\beta$  obtained from the ordered probit regression (Table 21 and Table 7) so that the enrollment pattern generated 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 4. The initial prior  $p_0$  affects the decision making of the student and the dynamic pattern in two different ways. First, 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  $p_0 = \Pr[\mu = 1]$ .

Figure 4 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 the figure 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  $p_0$ .<sup>22</sup> 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

---

<sup>22</sup>To produce the figure the model was simulated 100000 times and then a cubic polynomial was fitted to the simulated probabilities of each educational history.

aspect observed in the figure 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).<sup>23</sup>

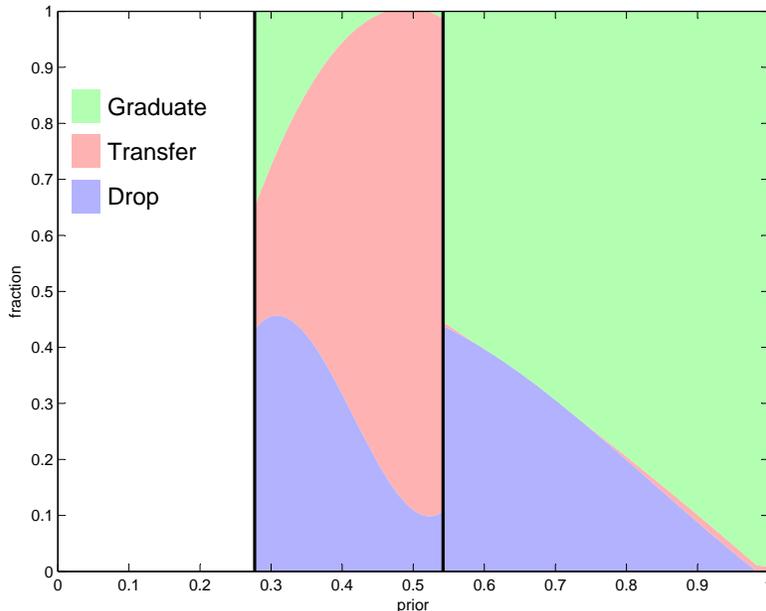


Figure 4: **Probability of Dropout (D), Transfer (T) and Graduation (G) for students initially enrolled in academic 2-year colleges and 4-year colleges.** The model was simulated 100000 times and then a cubic polynomial was fitted to the simulated distribution.

To evaluate the predictions of the model regarding the transition probabilities, [Table 10](#) tabulates students by behavior (dropout, transfer, graduation) in 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 model's predictions. 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 as very few students are included in this cells. Still, the evidence in these two cases does not conflict with the model's predictions.

<sup>23</sup>The shape of the figure is robust to different calibrations of the model. In particular, calibrations that match

	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

Table 10: **Proportion of agents initially enrolled in  $i$  with particular history (conditional on initial enrollment status).**

An alternative, and more involved, way of evaluating [Figure 4](#) with data is, conditional on the initial enrollment choice of agents, to estimate the density of each of the different educational histories non-parametrically (see [Figure 5](#)). Inspection of the figure provides similar results to what was discussed in [Table 10](#). Weighting accordingly these densities it is possible to produce the empirical counterpart of [Figure 4](#). 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 6](#) presents the results. The figure shows a very similar pattern to [Figure 4](#).

## 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 unaltered the primitives of the model, the model is solved again letting  $\gamma$  tend to zero, as this case maps to risk full insurance. Comparing the benchmark model (i.e.  $\gamma = 8$ ) with the risk-neutral case provide insights regarding the interaction

---

more closely dropout and transfers (in detriment of enrollment moments).

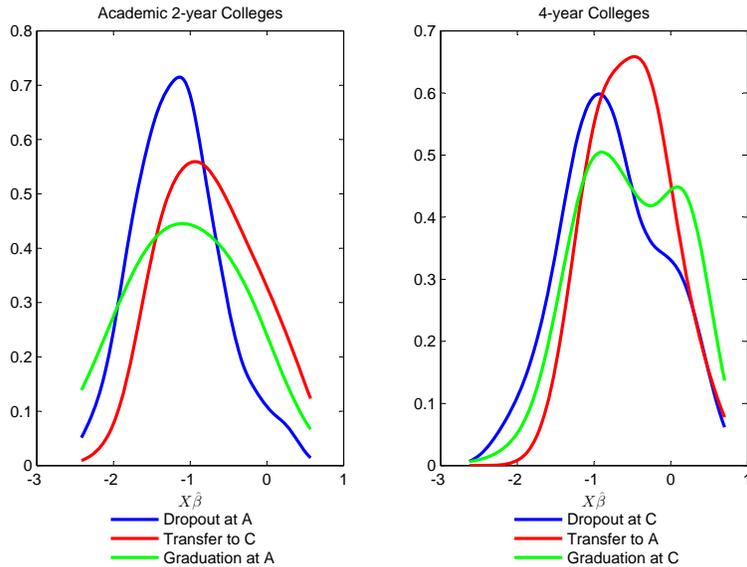


Figure 5: **Fraction of students a function of the measure of talent  $X'\hat{\beta}$ .** The estimation is performed conditional on the initial enrollment choice of agents.

of risk and insurance with the optimal policy and returns in this economy.<sup>24</sup> A similar strategy is followed to evaluate the size of the option value. Keeping the primitives unaltered, the model is solved two more times. In the first one the transfer option is discarded and in the second one both the transfer and dropout options are eliminated. Using a decomposition of returns, the value-added of each option is then evaluated.

## 6.1 Insurance

For a high-school graduate with any given prior  $p_0$ , Figure 7 presents the returns for the benchmark model ( $\gamma = 8$ ) and the model where  $\gamma \rightarrow 0$ . The comparison is important as risk aversion is tightly connected to market completeness. The more complete the markets, the lower the value for  $\gamma$ . It follows that Figure 7 compares the benchmark model with an economy where markets are complete.

<sup>24</sup>The analysis will abstract from Moral Hazard that can potentially arise from credit provision.

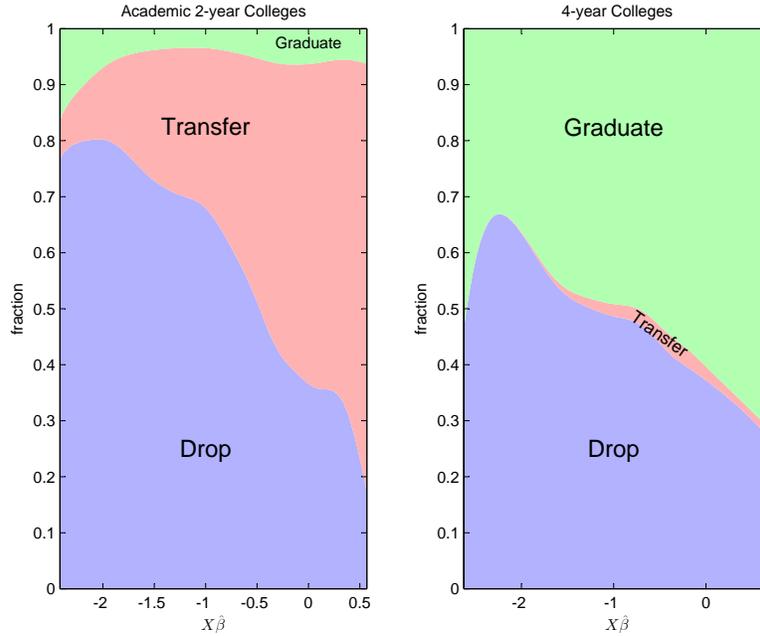


Figure 6: **Empirical Probability of Dropout (D), Transfer (T) and Graduation (G) for students initially enrolled in academic 2-year colleges and 4-year colleges as function of talent  $X'\hat{\beta}$ .** First, a conditional (on the initial enrollment choice) nonparametric estimation of each event was 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.

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  $C$  and  $A$  than for the one between  $A$  and  $W$  is not casual: enrollment at  $C$  is more risky than enrollment in  $A$  (simple comparison of the ratio of wages). It follows that a decrease in risk aversion have a stronger effect on  $C$  than in  $A$ . [Figure 7](#) also shows that risk aversion hinders the returns to education in an important way, and the effect is stronger the more uncertain the prior.

[Table 11](#) presents the mean return (and standard deviation) for the cross-section of agents of enrollment at academic 2-year colleges and 4-year colleges for both the benchmark model and the

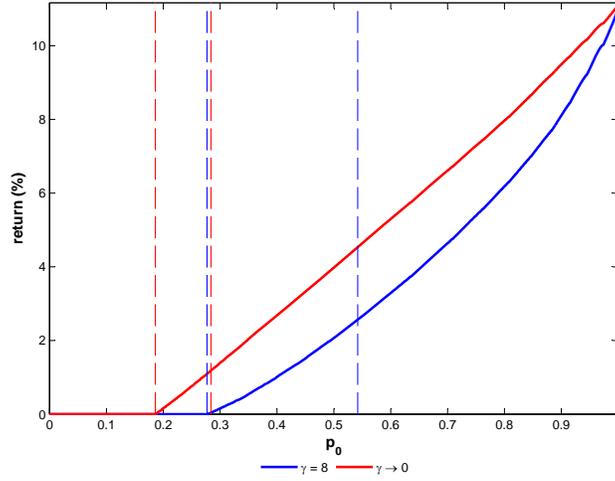


Figure 7: Return of Enrollment for a given initial prior  $p_0$ . The vertical lines define the indifference prior for enrollment between work and academic 2-year colleges and between academic 2-year colleges and 4-year colleges.

risk-neutral model for 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.

	Ac. 2-year C.	4-year C.
$\gamma = 8$	1.07 (0.43)	4.05 (0.42)
$\gamma \rightarrow 0$	0.13 (0.09)	4.7 (1.71)

Table 11: **Returns for  $\gamma = 8$  and  $\gamma \rightarrow 0$  for the estimated distribution of priors  $p_0$ .** All the numbers in the table are in percentage points.

Providing insurance not only increases returns for every prior but also affects enrollment decisions (this can be seen in [Figure 7](#) where the vertical dotted lines denote the enrollment thresholds).

[Table 12](#) computes the distribution of initial enrollment for both cases. As expected, full insurance

increases total enrollment (total enrollment increase by 10.6%) and enrollment in 4-year colleges where risk matters the most as the wedge in wages and cost of education are higher and time until graduation longer (enrollment in 4-year colleges increase 17.82%). Finally, the mass of students still enrolling in academic 2-year colleges with full insurance depicts 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. On the extensive level, providing full insurance increases total enrollment. On the intensive level, the provision of full insurance affects the composition of enrollment as risk, prices and prizes differ across types of institutions.

	workforce	Ac. 2-year C.	4-year C.
$\gamma = 8$	57.39	16.98	25.62
$\gamma \rightarrow 0$	52.54	5.17	42.29

Table 12: **Distribution of initial enrollment for  $\gamma = 8$  and  $\gamma \rightarrow 0$  for the estimated distribution of priors  $p_0$ .** All the numbers in the table are in percentage points.

## 6.2 How much Option Value?

To evaluate the size of the option value the model is solved again but reducing the amount of options. First, the transfer option is discarded and so 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, are 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, correspondingly, 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. 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 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 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 with no options available. **Figure 8** shows that 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 8** accounts for an important role for the transfer option to explain 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 is a small part of the role of academic 2-year colleges and thus explains little of the returns.

**Table 13** produce the same decomposition but for the mean return for the population distribution of priors  $p_0$ . The transfer option is very valuable in academic 2-year colleges accounting for 52% of total value. The dropout option is valuable in both types of institutions but more in 4-year colleges (71% vs. 39%).

	ac. 2-year C.		4-year C.	
	value-added	cumulative	value-added	cumulative
Enrollment	8.56	8.56	28.57	28.57
Dropout Option	39.17	47.73	71.37	99.94
Transfer Option	52.27	100	0.06	100

**Table 13: Proportion of returns explained by each option (and cumulative) for the estimated distribution of priors  $p_0$ .** All the numbers in the table are in percentage points.

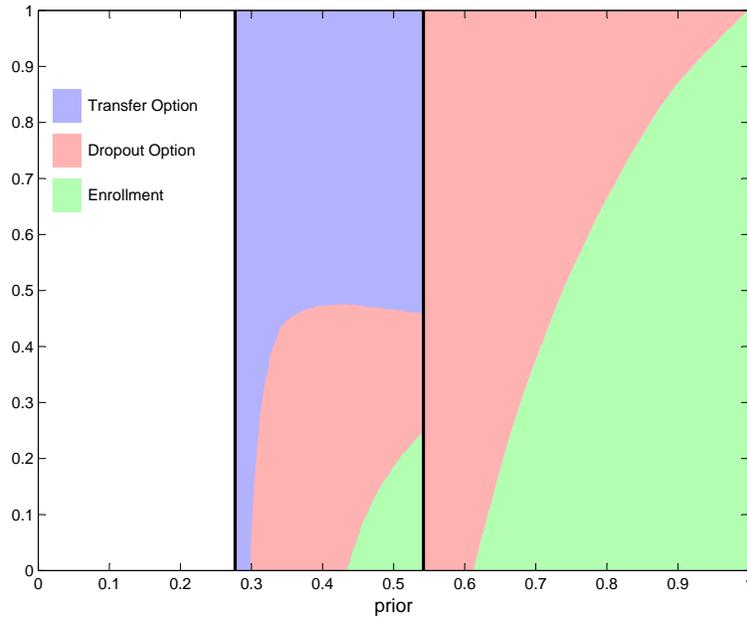


Figure 8: **Decomposition of Returns.**

## 7 Supporting Evidence from NLS-92

NLS-92 follows the cohort that graduate from high-school in 1992 up to the year 2000. As previously discussed, there are several reasons to use NLS-72 over this newer data set. First, NLS-72 presents wage information gathered 13 years after high-school graduation while NLS-92 gathers the information only 8 years after. Second, the questionnaire in NLS-72 has a specific section that allow for an easy distinction of academic 2-year colleges and vocational school, while this distinction can only be made in NLS-92 by looking at the credits and subjects taken by a student and deciding whether they are vocational or academic credits. Third, NLS-72 presents the dynamics of enrollment behavior by year allowing for an understanding of the dynamic pattern of education while in NLS-92 there is no easy way to do so. Finally, NLS-72 has detailed cost information (tuition, room and board, books, etc.) while this information is not available in NLS-92.

Still, the fact that NLS-72 is outdated raise concern about the validity of the results discussed in the paper. With this in mind, this section intends to replicate part of the evidence and implications of the model for NLS-92 and contrast it with NLS-72. [Table 14](#) shows that the distribution of initial enrollment changed from the 70's to the 90's as noted by the important increase in the share of high-school graduates enrolling in 4-year colleges in detriment of the share joining the workforce. As noted in [Heckman and LaFontaine \(2008\)](#) this can be explained by the increase in the amount of high-school students obtaining a GED. [Table 15](#) presents the aggregate dynamics for students initially enrolled in 2- and 4-year colleges for both NLS-72 and NLS-92. The table shows that the patterns of education are similar for both data sets.

%	W	V+A	C
NLS-72	51	24	25
NLS-92	27	23	50

Table 14: **Distribution of Initial Enrollment: NLS-72 vs. NLS-92.**

[Table 16](#) shows that the hypothesis of sorting discussed for NLS-72 also holds for NLS-92 (an ordered probit regression was performed using the same variables as regressors as those used in [Table 21](#)) but collapsing together academic 2-year colleges and vocational school.

	%	D at (V+A)	D at C	T to (V+A)	T to C	G at (V+A)	G at C
NLS-72	V+A	73	45	-	21	6	55
	C	89	41	5	-	11	54
NLS-92	V+A	37	27	-	45	18	73
	C	78	26	1	-	22	73

Table 15: **Aggregate Dynamics: NLS-72 vs. NLS-92.**

The evidence shows that the idea of educational ladder also holds for the data set constructed from NLS-92.

To test whether the predictions of the model also hold for the newer data set, [Table 17](#) replicates the analysis performed in [Table 10](#) that can be contrasted with [Figure 4](#). Comparison of [Table 17](#)

	W	V+A	C
NLS-72	-1.71 (0.54)	-1.38 (0.57)	-0.94 (0.61)
NLS-92	-2.67 (0.683)	-1.87 (0.767)	-1.23 (0.77)

Table 16: **Sorting: NLS-72 vs. NLS-92.**

and Table 10 show that the model’s predictions also hold for newer cohorts as those included in NLS-92.

	Ac. 2-year C.			4-year C.		
	D	T	G	D	T	G
Low $X'\hat{\beta}$	55%	30%	15%	29%	13%	58%
Med $X'\hat{\beta}$	36%	42%	22%	20%	7%	73%
High $X'\hat{\beta}$	21%	66%	13%	6%	3%	91%
# of obs.	200	248	91	248	106	990

Table 17: **Proportion of agents initially enrolled in  $i$  with particular history (conditional on initial enrollment status and measure of talent  $X'\hat{\beta}$ ) - NLS-92.**

## 8 Conclusion

This paper presents new evidence regarding why returns to 4-year college graduation are large relative to the low enrollment and graduation rates: the wedge in returns for the marginal student in 4-year college is explained by 2-year colleges. Evidence in terms of sorting and Internal Rates of Return points in this direction.

The ex-post return to graduating from an academic 2-year college is low, but there is a moderate return to dropping out and a large return to transferring to a 4-year college: academic 2-year colleges act as a *stepping stone* in which agents learn about themselves in a cheaper and less demanding environment. A model of educational of post-secondary educational choice that

incorporates academic 2-year colleges together with 4-year colleges and work is explored where high-school graduates are uncertain about their ability to accumulate human capital. Depending on the value of their beliefs, agents sort across enrollment alternatives. During tenure at school, students receive signals that produce a re-evaluation of beliefs that, by interacting with the current amount of accumulated credits, produce dropouts and transfers. The properties of academic 2-year colleges, a type of 2-year college, are such that makes these institutions an ideal practice ground for students with aspirations of graduation at 4-year colleges but with low expectations about their ability.

Is there any support for the learning mechanism over competing mechanisms such as non-pecuniary costs of education? [Stinebrickner and Stinebrickner \(2008\)](#) provides the first piece of evidence as it shows that bad grades in exams explain both dropouts and statements about disliking college (causality). A second piece of evidence arises from contrasting the predictions of a parameterized version of the model with data. These predictions are a unique feature of the learning mechanism and the interaction of beliefs with observable measures of ability.

Finally, the parameterized version of the model is used to evaluate first, how the provision of full insurance affects the allocation and measured returns and second, the value-added of the transfer and dropout option to total value. The results show that risk plays an important role on explaining why agents enroll in academic 2-year colleges and that the availability of dropout and transfer options explain much of the returns to academic 2-year college enrollment. A similar idea applies for 4-year colleges. First, the provision of full insurance increases total enrollment at 4-year colleges mostly by a shift in the composition. Second, the decomposition of returns show that the dropout option is an important source of value. All of these results points towards a single conclusion: the interaction of risk and option value is a major force in post-secondary education.

## References

- Altonji, J. G.**, “The Demand for and Return to Education when Education Outcomes are Uncertain,” *Journal of Labor Economics*, 1993, 11 (1).
- Belley, P. and L. J. Lochner**, “The Changing Role of Family Income and Ability in Determining Educational Achievement,” *Journal of Human Capital*, 2008, 1 (1).
- Belzil, C. and J. Hansen**, “Unobserved Ability and the Return to Schooling,” *Econometrica*, 2002, 70 (5).
- Cameron, S. and C. Taber**, “Borrowing Constraints and the Returns to Schooling,” *Journal of Political Economy*, 2004, 112.
- and **J. J. Heckman**, “The Dynamics of Educational Attainment for Blacks, Whites and Hispanics,” *Journal of Political Economy*, 2001, 109 (3).
- Carneiro, P., K. T. Hansen, and J. J. Heckman**, “Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measuring of the Effects of Uncertainty on College Choice,” *International Economic Review*, 2003, 44 (2).
- Comay, Y., A. Melnik, and M. A. Pollatschek**, “The Option Value of Education and the Optimal Path for Investment in Human Capital,” *International Economic Review*, 1973, 14 (2).
- Cunha, F., J. J. Heckman, and S. Navarro**, “Separating uncertainty from heterogeneity in life cycle earnings,” *Oxford Economics Papers*, 2005, 57 (2).
- Grubb, W. N.**, “The Varied Economic Returns to Postsecondary Education: New Evidence from the Class of 1972,” *Journal of Human Resources*, 1993, 28 (2).
- Heckman, J. J. and P. A. LaFontaine**, “The American High School Graduation Rate: Trends and Levels,” *IZA Discussion Papers*, 2008, (3216).
- and **S. Urzua**, “The Option Value of Educational Choices and the Rate of Return to Educational Choice,” *unpublished*, 2008.
- , —, **L. J. Lochner, and P. E. Todd**, “Earning Functions and Rates of Return,” *Journal of Human Capital*, 2008, 2.
- , —, —, and **P. Todd**, *Earnings Functions, Rates of Return, and Treatment Effects: The Mincer Equation and Beyond*, Vol. Handbook of the Economics of Education, Elsevier Science, 2006.
- , —, —, and —, “Earnings Functions and Rates of Return,” *Journal of Human Capital*, 2008, 2 (1).
- Johnson, W. R.**, “A Theory of Job Shopping,” *Quarterly Journal of Economics*, 1978, May.
- Jovanovic, B. and Y. Nyarko**, “Stepping-stone mobility,” *Carnegie-Rochester Conference Series in Public Policy*, 1997, 46.

- Judd, K. L.**, “Is Education as Good as Gold? A Portfolio Analysis of Human Capital Investment,” *unpublished*, 2000.
- Kane, T. J. and C. E. Rouse**, “Labor-Market Returns to Two- and Four-Year College,” *American Economic Review*, 1995, 85 (3).
- and –, “The Community College: Educating Students at the Margin between College and Work,” *Journal of Economic Perspectives*, 1999, 105 (3).
- Keane, M. P. and K. I. Wolpin**, “The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment,” *International Economic Review*, 2001, 42 (4).
- Lochner, L. J. and A. Monge-Naranjo**, “Human Capital Formation with Endogenous Credit Constraints: Government Student Loan Programs and Private Lenders,” *unpublished*, 2007.
- Lucas, R. E. Jr. and N. Stokey**, *Recursive Methods in Economic Dynamics*, Harvard University Press, 1989.
- Mehra, R. and E. C. Prescott**, “The Equity Premium: A puzzle.,” *Journal of Monetary Economics*, 1985, 22.
- Miao, J. and N. Wang**, “Experimentation under Uninsurable Idiosyncratic Risk: An Application to Entrepreneurial Survival,” *unpublished*, 2007.
- Miller, R.**, “Job Matching and Occupational Choice,” *Journal of Political Economy*, 1984, 92 (6).
- Stange, K.**, “An Empirical Examination of the Option Value of College Enrollment,” *unpublished*, 2007.
- Stinebrickner, T. R. and R. Stinebrickner**, “Learning About Academic Ability and the College Drop-out Decision,” *working paper*, 2008.

## Appendix

### A Proof of Proposition 1

Solving for  $c$  from the budget constraint and substituting back into [equation \(1\)](#) reduces the problem to a single variable problem. Further, it is straightforward to check that the conditions for unique solution to [equation \(1\)](#) 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 \(1\)](#) 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  $W(a; h) = -\frac{1+r}{\gamma r} e^{-\gamma(ra+h)} + \frac{1+r}{\gamma r}$

### B Proof of Proposition 2

Solving for  $c$  from the budget constraint and substituting back into [equation \(3\)](#) 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 proposition shows that there exists a unique solution to [equation \(3\)](#).

**Proposition 7**  $V_i(a, s, p)$  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  $\eta$  that arrive produce updating in the state  $p$ , which is thus stochastic.<sup>25</sup> 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 \(3\)](#) provides the maximized value function,

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

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

Conject that  $\tilde{V}_i(a', s', p') = -\frac{1+r}{\gamma r} e^{-\gamma(ra' + \tilde{v}_i(s, p))} + \frac{1+r}{\gamma r}$  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 \(4\)](#) reads,

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

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  $V_i(a, s, p)$ .

## C Proof of [Lemma 3](#)

See [equation \(7\)](#). For any given prior  $p$ , each of the exponentials inside the  $\max\{\}$  as function of  $\eta$  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  $[e^{-\gamma h^w}, e^{-\gamma h_1^C}]$ , 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

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

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

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

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

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

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

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

#### Lemma 4

$$\int_{\eta} \max \{ J_1(\gamma, b(p, \eta), s + \Omega(\eta)), J_2(\gamma, b(p, \eta), s + \Omega(\eta)), J_3(\gamma, b(p, \eta), s + \Omega(\eta)) \} H_i(d\eta, p) \quad (11)$$

*differentiable with respect to  $\gamma$ .*

**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  $\eta$  (that depends on  $s$ ,  $p$ ,  $\gamma$ , 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  $\gamma$ , these thresholds are also differentiable by construction. It follows that [equation \(11\)](#) can be rewritten as

$$\int_{-\infty}^{\eta^L(\gamma, p, s)} J_1(\gamma, p', s') H_i(d\eta, p) + \int_{\eta^L(\gamma, p, s)}^{\eta^H(\gamma, p, s)} J_2(\gamma, p', s') H_i(d\eta, p) + \int_{\eta^H(\gamma, p, s)}^{\infty} J_3(\gamma, p', s') H_i(d\eta, p) \quad (12)$$

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

Let  $Q(\gamma, p, s)$  denote the object in [equation \(11\)](#).

#### Proposition 8

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

**Proof.** Follows by applying Leibniz's rule to [equation \(12\)](#) and by noting that

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

by construction of the thresholds  $\eta^L$  and  $\eta^H$ . ■

Now, **Lemma 3** follows by applying L'Hopital's rule for the case where  $\gamma \downarrow 0$  and by the results of **Proposition 8**.

## D Proof of Proposition 3

The proof for  $p$  follows by induction. The ultimate goal in Postsecondary 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.

## E Proof of Proposition 6

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 lemma shows that, in this case,  $V_A(a, s_1, p_t^A(s_0)) > V_C(a, \theta^A(s_1), p_t^A(s_0))$ .

**Lemma 5**  $V_A(a, s_1, p_t^A(s_0)) - V_C(a, \theta^A(s_1), p_t^A(s_0)) > 0$ .

**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,

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

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

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

Finally, as  $V_A(a, s, p)$  increasing and convex in  $s$ ,  $\theta^A(s)$  increasing and concave and  $\frac{\partial \theta^A(s_0)}{\partial s} < \frac{\partial \theta^A(0)}{\partial s} < 1$ ,

$$V_A(\Upsilon_1^0) - V_C(\Psi_1^0) > 0$$

which completes the proof. ■

**Lemma 6** *For any  $p > \min \{p_d^A(s), p_d^C(s)\}$ , 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 (see [Lemma 6](#))  $V_A(a, s_1, p_t^A(s_0)) - V_C(a, \theta^A(s_1), p_t^A(s_0)) < 0$  which violates [Lemma 5](#). 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.

## F 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  $i$ ,  $\tau_i$  the flow cost of attendance,  $\varpi_i^D$  the increase in wages due to dropping out at institution  $i$ ,  $\varpi_i^G$  the graduation premium,  $\alpha_i$  the increase in wages due to experience,  $G_i$  a dummy that accounts for graduation at institution  $i$  and  $D_i$  a dummy that accounts for dropping out at institution  $i$ . Further, assume that students transfer only once.<sup>26</sup>

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

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

<sup>26</sup>Very few students transfer more than once in NLS-72. Also, the assumption makes the presentation of the methodology much easier.

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

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

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 3** 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  $\varpi_i^D$  and  $\varpi_i^G$  can be obtained from that table (see Version A).

**Table 4** 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  $\alpha_i$  (see Version A).

## G Construction of Data set from NLS-72

Individuals with the following characteristics were eliminated: (1) students that join Postsecondary institution after October of 1972, (2) individuals with two or more spells of Postsecondary education, (3) individuals with missing values for observable characteristics (discussed later), (5) individuals with Professional, Master or PhD. degrees.

The set of observable characteristics used here are: race, gender, Socioeconomic Status of Family, Maximum Educational Level of Father, rank in high-school class and location of high-school.

The cost of education presents several missing values. To input values here a regression using of the cost of education on observable characteristics was used to produces estimates.

The final data set contains 3480 individuals.

**TO BE IMPROVED.**

## H Other Tables and Plots

Initially enrolled in V	time at V	time at A	time at C
graduate at V	3.26	-	-
dropout at V	1.45	-	-
transfer to A	1.22	-	-
transfer to C	1.11	-	-
graduate at A	-	3.5	-
dropout at A	-	1.28	-
graduate at C	-	-	4.75
dropout at C	-	-	2.8

Table 18: Mean times for different educational histories for students initially enrolled in vocational school.

Initially enrolled in A	time at V	time at A	time at C
graduate at A	-	3.84	-
dropout at A	-	1.72	-
transfer to V	-	1.43	-
transfer to C	-	2.21	-
graduate at V	1.66	-	-
dropout at V	1.05	-	-
graduate at C	-	-	3.36
dropout at C	-	-	2.42

Table 19: Mean times for different educational histories for students initially enrolled in academic 2-year colleges.

Initially enrolled in C	time at V	time at A	time at C
graduate at C	-	-	5.23
dropout at C	-	-	3.02
transfer to V	-	-	1.61
transfer to A	-	-	1.53
graduate at V	1.66	-	-
dropout at V	1.44	-	-
graduate at A	-	3.5	-
dropout at A	-	2.05	-

Table 20: Mean times for different educational histories for students initially enrolled in 4-year colleges.

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

Table 21: **Evidence on Sorting: Ordered Probit Regression (NLS-72)**. Ordered probit estimation of the initial enrollment choice on observable measures of ability. Version B merges vocational school with work.