College Risk and Return*

Gonzalo Castex†
University of Rochester
Central Bank of Chile

August 2010

Abstract

Attending college is thought of as a very profitable investment decision, as its estimated annualized return ranges from 8% to 13%. However, a large fraction of high school graduates do not enroll in college. Using a simple risk premium approach, I reconcile the observed high average returns to schooling with relatively low attendance rates. A high dropout risk has two important effects on the estimated average returns to college: selection bias and risk premium. Taking into account dropout risk, a simple calculation of risk premium accounts for 51% of the excess return to college education. In order to explicitly consider the selection bias, I further explore the dropout risk in a life-cycle model with heterogeneous ability. The risk-premium of college participation accounts for 29% of the excess returns to college education for high-ability students and 27% of the excess return for low-ability students. Risk averse agents are willing to reduce their return to college in order to avoid the dropout risk. The effect is not uniform across ability levels.

JEL Classification: I21, J17, J24
Keywords: College attendance, Risk premium, Return to education, College dropout

1 Introduction

Attending college is considered a most profitable investment, as its estimated annualized return ranges from 8% to 13% (Card 1999). However, a large fraction of high school graduates do not enroll in college. According to the National Longitudinal Sample of Youth 1979 (hereafter NLSY79), 59% of high school graduates do

*Special thanks are due to Yongsung Chang for his advice, encouragement and support. I also thank Geni Dechter and Andrew Davis for useful comments and participants at EALE and SOLE 2010 Conference and Bank of Israel Seminar. All errors are my own.
†Correspondence: gcastex@bcentral.cl

not continue formal education. Moreover, around 30% of high school graduates from the highest quartile of the cognitive ability distribution, measured by AFQT scores, do not enroll in college.[1] This is despite the availability of student loan funds for financing college and the fact its returns easily compensate the earnings forgone while in college (Ionescu 2009).

Selection bias is a common explanation used by the existing literature to explain the high return to college. Different schooling levels may be attributed to differences in individual aptitudes and tastes for schooling relative to work (Card 2001). Using a simple risk premium approach, I reconcile the observed high average returns to schooling with relatively low attendance rates.

Previous literature often omits an important factor about the college investment decision: the *risk*. According to Restuccia and Urrutia (2004) and Mayer (2008), the average college dropout rate is around 45%-55%.[2] This high dropout risk affects the estimated average returns to college through two importat channels: selection bias (as in the traditional literature by Card 2001, Willis and Rosen 1979) and risk premium (as in the equity premium literature such as Mehra and Prescott 1985). Taking into account dropout risk, a simple calculation of risk premium accounts for 51% of the excess return to college education.[3] In order to explicitly consider the selection bias and the ability level effect on the risk premium, I further explore the dropout risk in a life-cycle model with heterogeneous ability. The risk premium of college participation accounts for 27% of the excess return to college education for students from the lowest quartile of the ability distribution and about 29% of the return to college education for students from the highest quartile of the ability distribution.

According to the NLSY79, the college dropout rate is about 55%. The average tuition paid by students who dropped out was, at the beginning of the 1980s, $3,300. 5% of this group paid more than $10,000.[4] On average, students who drop out owe around $9,350 to financial and educational institutions; 15% of this group owes more than $24,000. The decision to enroll in college seems not always to be ex-post optimal, in particular for students who dropout and have larger accumulated debts for college financing.

I propose two approaches to quantify how much of the excess of return to college education is due to the college risk. I first evaluate in a simple framework the risk premium of this investment decision. I follow Mehra and Prescott (1985) to quantify the excess return to college education. The human capital investment

[1] Armed Forces Qualification Test scores are widely used in the literature as a measure of cognitive achievement, aptitude and intelligence.

[2] Also consistent with Census and NLSY79.

[3] The return to college education is compared to the return of a risk free bond to quantify the exess return.

[4] Values expressed in 2007 dollars adjusted by CPI.

decision can be analyzed with a two-period model to evaluate its return. In the first period agents acquire human capital by attending college. In the second period the payoff of the investment decision is realized. Agents may successfully graduate from college or may fail. I calibrate the model for different points in time for the U.S. economy, from 1960 to 2007, using Census data and the American Community Survey (ACS). Under this approach I evaluate the role played by two sources of uninsurable risk on college enrollment: the dropout risk and the larger volatility in permanent income shocks of college graduates compared to high school graduates (as documented by Chen 2002). Then, I explore the role played by heterogeneity in ability levels and build a life-cycle model with endogenous enrollment. Agents who enroll in college face dropout risk. The model is calibrated to match college participation rates and dropout rates observed in NLSY79. Under this model specification, I evaluate the role played by the dropout risk for each level of ability and quantify how much of the return to college education is explained by its risk.

The rest of the paper is organized as follows: Section 2 analyzes the return to college in an risk premium environment, describing the data, calibration strategy and results. Section 3 describes a more structured model to account for heterogeneous ability and their role with risk premium. Section 4 concludes the paper.

2 Risk Premium Approach

This section applies the theory of consumption behavior and asset pricing in a static representative agent model. I use the consumption based capital asset pricing model to quantify the risk premium of the risky college investment decision, as in the traditional equity premium literature by Mehra and Prescott (1985). Given the nature of schooling choice as a human capital investment, a two-period model is suitable to quantify how much of the return to college is explained by its risk. I consider two sources of uninsurable risk: a permanent shock on wage for college graduates and the risk of failing college. The permanent shock on wages arises in the form of a stochastic return to college education after college-graduation (Chen 2002). The risk of failing college arises in the form of dropout risk.

2.1 The environment

Consider an economy with a representative agent that lives for two periods. The life-time utility for a college attendant is described by:

$$V^c = \mathbb{E}\{u(c) + \beta[\psi u(c \times (1 + \widetilde{r})) + (1 - \psi)u(c \times (1 + r_d))]\}$$

where c corresponds to consumption level of a single good in both periods; $\widetilde{r}$ is the stochastic return to college education; r_d is the return to college dropout; ψ corresponds to the probability of graduating from college; β is the discount factor, $u(.)$ is a current period utility function, and $\mathbb{E}\{.\}$ is an expectation operator over the stochastic return to college in period 2. There is no saving or borrowing technology in this environment and there is no insurance mechanism against the stochastic return to college.

Education can only be acquired during the first stage of the life-cycle ($t = 1$). The outcome of this investment opportunity is uncertain and realized at the second stage of the life-cycle ($t = 2$). For agents who do not acquire college education, consumption levels in the first and second period of their life-cycle correspond to high-school labor income, W_1 and W_2 respectively. If agents decide to attend college in the first period they forgo earnings and pay tuition. This cost is summarized by τ. In the second period these agents face uncertainty about college completion. With probability ψ an agent successfully graduates from college, in which case they obtain the return to college investment, denoted by $\widetilde{r} \sim N(r_c, \sigma^2_{r_c})$. This stochastic return to college education corresponds to the second source of uncertainty faced when agents make their enrollment decision. If the individual fails to graduate, they receive a fraction of the return to college denoted by r_d. Table 1 summarizes the payoffs of the college-enrollment decision problem.

	stage 1	stage 2
No College	W_1	W_2
College	$W_1(1-\tau)$	$\psi \int_{\widetilde{r}} W_2(1+\widetilde{r})dF(\widetilde{r}) + (1-\psi)W_2(1+r_d)$

Table 1: Payoffs of the model

Note: The table displays the payoff for the two period representative agent model for two educational alternatives. Attending college implies a cost τ in the first period and an uncertain payoff at the second stage.

The cost of the human capital investment decision is paid in the first stage of the life-cycle, when college students forego earnings and pay tuition, while the stochastic benefits are received in the second stage. The risk premium of the college decision is obtained by pricing tomorrow's consumption stream, considering an environment with and without uncertainty.

There are two sources of uninsurable risk under consideration in this environment. First, a permanent shock on wages after college graduation, when college-graduates draw their return to education $\widetilde{r}$. Second, the dropout risk of attending college, $1 - \psi$.

Taking attending college as the risky asset, following Lucas (1978), the price of attending college, p, is obtained from the optimality condition.

$$-pu'(c_1) + \beta\mathbb{E}[\psi(1+\widetilde{r})u'(c_1 \times (1+\widetilde{r})) + (1-\psi)(1+r_d)u'(c_1 \times (1+r_d))] = 0 \quad (1)$$

The mathematical expectation is over the return to education drawn conditional on completing college, as shown in Table 1. The Previous specification corresponds to the price of attending college in an environment with two sources of risk.

Computing the return to college education in environments with and without risk allows me to quantify how much of the excess of return to college is explained by the dropout risk and permanent shock on wages (stochastic return).

2.2 The stochastic return to college education

In this subsection, I estimate the stochastic return to college education to quantify how much of the excess of return to college is due to the permanent shock on wages and college dropout risk.

The internal rate of return plays a key role in economics of human capital: an additional level of schooling is considered profitable if the internal rate of return exceeds the opportunity cost. Carnoy and Marenbach (1975) estimate the rate of return to college education by using the standard discount formula:[5]

$$0 = \int_0^T (Y_t - C_t)e^{-rt}dt$$

where Y_t corresponds to the difference in average wage income in period t between those workers with college education and those without college education (high school graduates). C_t corresponds to the cost of schooling in period t, which includes tuition cost and foregone earnings while in school, $C_t = 0$ after individuals join the labor force. r represents the marginal internal rate of return of college education. T corresponds to the total number of periods under consideration, from the beginning of the college education to the end of working life.

[5] See also Heckman et al. (2008).

Incorporating the dropout risk and the permanent income shock in wages, the stochastic internal rate of return is obtained from:[6]

$$0 = \int_0^{t_2=2} (Y_t - C_t) e^{-\widetilde{r}t} dt + \int_{t_2=2}^{T} (\psi(\widetilde{Y}_t - C_t) + (1-\psi)(Y_t^D - C_t)) e^{-\widetilde{r}t} dt \qquad (2)$$

Where $\widetilde{Y}_t$ corresponds to the stochastic wage differential between college graduates and high school graduates, which depends on the wage draw after college graduation (permanent income shock as in Chen 2002). Y_t^D corresponds to the earning differential between college-dropouts and high school graduates. Agents successfully graduate from college with probability ψ.

The college premium plays a key role in determining the stochastic internal rate of return to college. Agents draw a stochastic wage premium after college graduation that is permanent in their lives, a permanent income shock in wages. Given this stream of income is possible to obtain the internal rate of return, $\widetilde{r}$, from equation 2.

The difference between these two measures of the stochastic return to education, internal rate of return and college-wage premium, is that the former one already incorporates the cost of the investment opportunity, in terms of foregone earnings and tuition cost. The second measure just captures the benefit of acquiring college education and not the investment cost.

Next subsection describes the data and calibration strategy to estimate first the internal rate of return (monetary return to college) and then the risk premium using the Lucas (1978) asset pricing model as described in the previous subsection.

2.3 Data and Calibration

In order to analyze the risk and return to college education across time for the US economy, I use data from the Census from 1960 to 2000 and from the American Community Survey (ACS) from 2001 to 2007, both provided by the IPUMS-CPS project,[7] where I obtain individual earnings at different educational levels across age. In particular I analyze total wage and salary income,[8] considering white males aged 18 to 65

[6]For simplicity is assumed that dropout occurs at the end of the second year of college education.
[7]King et al. 2008, see technical details on the CPS and ACS data at http://cps.ipums.org/cps/samples.shtml.
[8]This variable reports the respondents total pre-tax wage and salary income from previous calendar year.

years old that are part of the labor force.[9]

For each cohort, I estimate average wage for each age and educational level to construct the wage profile along the agent's life-cycle. Standard deviations are also computed to quantify the difference between those workers with college education and those with high school education (permanent income shock in wages). With the wage profile streams is possible to estimate the internal rate of return to college education, incorporating the tuition cost and foregone earnings while in college (see equation 2).[10]

The return to college education is quantified according the descriptions of the previous subsection. r_c corresponds to the average monetary return to college graduation, r_d corresponds to the monetary return to college dropout, σ_r corresponds to the standard deviation of the monetary return. s corresponds to the college graduate wage premium, s_d corresponds to the college dropout wage premium and σ_s the standard deviation of the wage premium.[11]

Year	1960	1970	1980	1990	2000	2005
r_c	6.72%	9.18%	9.13%	13.54%	16.82%	18.55%
r_d	5.6%	6.68%	6.86%	10.06%	12.42%	13.28%
σ_r	2.71%	2.78%	3.35%	2.75%	2.91%	2.89%
s	31.06%	38.33%	34.12%	56.85%	64.79%	73.80%
s_d	17.74%	21.42%	23.30%	21.80%	25.83%	21.53%
σ_s	11.80%	12.30%	11.05%	18.49%	20.82%	22.81%

Table 2: Monetary Return to College and College Premium

Note: The table displays return to college education, r_c, return to college dropouts, r_d, standard deviation of the return to college, σ_r, college wage premium, s, dropout wage premium, s_d and standard deviation of the college wage premium, σ_s. Source Census, computed as described in equation 2.

As documented by Goldin and Katz (2007), the college premium has been increasing during the last decades and technological advance has outpaced the number of students enrolling in college. Both measures of the return to college, monetary terms and wage differential, are 98.6% correlated across time. The difference is explained by the investment cost that is not considered in the college wage premium approach.

The probability of success in college is estimated from the data. It is the fraction of people who graduated from college conditional on college enrollment for each particular cohort. The values are consistent with reports in previous literature, such as Restuccia and Urrutia (2004) and Mayer (2008) and are similar to those calculated using the NLSY79 sample. The values across time are reported in Table 3.[12]

[9] See summary statistics in Appendix A.

[10] Tuition cost from 1960 to 2007 for U.S. is reported in Appendix A, source: College Board.

[11] Values from 2001 to 2007 are available upon request to the author.

[12] Dropout rates from 2001 to 2007 are available upon request to the author.

Year	1960	1970	1980	1990	2000	2005
ψ	0.4786	0.4565	0.4458	0.5216	0.5161	0.4780

Table 3: Probability of college success (white male, source: Census, ACS)
Note: Computed as the fraction of students who do not graduate conditional in college enrollment. Source Census.

To quantify the risk premium associated with the college enrollment decision, I parameterize the model such that the first period occurs when agents are 18 to 22 years old, representing the period in which individuals decide whether to acquire human capital by attending college. The second period occurs when agents are 23 to 65 years old and actively participate in the labor market.[13]

At time zero, a risk adverse agent decides whether to attend college during the first stage of the life-cycle or to join the labor force as an unskilled worker. This decision is not only based on the return to college, but also on its risk.

To quantify how much of the return to college is explained by its risk, I first solve the model considering a risk-free environment. Then I add the dropout risk and the volatility of the college return. The difference in excess returns to college obtained in each setup allows me to evaluate the risk premium involved in the college enrollment decision. The price of the college decision is computed for each of the risky frameworks as described in Table 4.

Environment	Optimality condition
Model 1, No risk	$-p_{M_1}u'(c_1)+\beta[(1+r_c)*u'(c_1(1+r_c))]=0$
Model 2, Permanent shock	$-p_{M_2}u'(c_1)+\beta\mathbb{E}[(1+\widetilde{r})*u'(c_1(1+\widetilde{r}))]=0$
Model 3, Dropout risk	$-p_{M_3}u'(c_1)+\beta[\psi(1+r_c)u'(c_1(1+r_c))+(1-\psi)(1+r_d)u'(c_1(1+r_d))]=0$
Model 4, Dropout and Permanent shock	$-p_{M_4}u'(c_1)+\beta\mathbb{E}[\psi(1+\widetilde{r})u'(c_1(1+\widetilde{r}))+(1-\psi)(1+r_d)u'(c_0(1+r_d))]=0$

Table 4: Optimality conditions for college attendance
Note: The expressions correspond to the pricing formula for one unit of college education as a consumption good, under 4 scenarios.

The utility function is restricted to the constant relative risk aversion class, $u(c)=\frac{c^{1-\gamma}}{1-\gamma}$. The parameter γ measures the curvature of the utility function. Its value is assumed to equal 2, as standard in the literature. The discount factor, β, is assumed to equal 0.96. Consumption values are expressed in per-year units.

[13] Parameters and consumption levels are converted in annual terms.

2.4 Results

To evaluate to what extent the excess return to college is explained by dropout risk and the permanent income shock on college-wages, I use the proposed two-period model to quantify the excess return to college education under 4 scenarios: no risk, considering only the permanent shock on college-wages, considering only the college dropout risk, and considering both sources of uncertainty. Each framework is evaluated for the US economy from 1960 to 2005. I quantify the excess return to college education under linear utility first to also evaluate the effect of the curvature of the utility function on the risk premium of college education.

In a risk neutral environment, it is possible to quantify the dropout effect on the return to college education directly from Tables 2 and 3. The return to college education in 1980 is 9.13% when considering successful students only. This implies a 5.13% excess return to college education, compared to a 4% risk free asset. Adding into the evaluation the dropout probability reported in Table 3, the return to college decreased to 7.87% (9.13% $\times$ 44.58% + 6.86% $\times$ 55.42%), implying an excess return to college education of 3.87%. Considering the risk of college education, the excess or return to college decreased by 25% in 1980s (21% in average from 1960 to 2007). Results are shown in second column of Table 5.

Returning to CRRA utility, I start by estimating Model 1. This specification does not consider any source of uncertainty. When there is no dropout risk, individuals who choose to acquire college education join the labor market as college educated workers. Additionally, in this setup there is no uncertainty about the college premium they will draw, i.e., no permanent shock on wages. Excess returns to college education in this environment are reported in the first column of Table 5.

The next step is to estimate Model 2. This model specification departs from the benchmark environment by incorporating one source of uncertainty: a permanent shock on college wages. The wage shock is summarized by σ_r^2. However, as there is no college dropout risk in this setup, all agents successfully graduate from college and draw a monetary return $\widetilde{r} \sim N(r_c, \sigma_r^2)$. The third column of Table 5 reports the excess return to college education under this risky environment.

I proceed by estimating Model 3. In this setup the only source of risk is the possibility of dropping out from college. Successful graduation for participating individuals occurs with probability ψ, but with probability $1-\psi$ agents fail and join the labor force as college dropouts. Agents who successfully graduate do not face uncertainty about the monetary return they draw. Column 4 in Table 5 provides the estimated excess return to college education across time under this scenario. Differences between values reported in column 4 and column 2 are explained by risk aversion.

Finally I estimate the complete model, Model 4. In this framework I allow for two sources of risk: permanent shock on college premium and college dropout risk. The results are shown in the last column of Table 5.

For each of the 5 model specifications, the excess return is computed with respect to a risk free asset, an asset who pays one unit of the consumption good in the second period. I take 4% as the return of the risk free asset.

The following table shows the quantified excess of return to college education. Values are shown in monetary return units, as in table 2.[14]

Year	Model 1 No risk	Linear utility considering dropout	Model 2 Permanent shock	Model 3 Dropout risk	Model 4 Perm. shock + Dropout risk
1960	2.72%	2.14%	2.65%	2.13%	-0.17%
1970	5.18%	3.82%	5.11%	3.81%	1.36%
1980	5.13%	3.87%	5.03%	3.86%	1.36%
1990	9.54%	7.88%	9.47%	7.85%	5.59%
2000	12.82%	10.69%	12.75%	10.65%	8.30%
2005	14.55%	11.80%	14.48%	11.74%	9.19%

Table 5: Excess of return to college education

Note: The table displays the excess return to college education compared to a 4% risk free asset. Return to college education is computed as described in Table 4. Data source: Census.

Results summarized in Table 5 show the excess returns to college education under each model specification. Values are reported per year of college education.

The model without uncertainty, Model 1, shows the return to college education across time in an environment without dropout risk and permanent shock on wages. Agents who decide to accumulate human capital successfully graduate from college and obtain the average monetary return as a payoff. Agents earn in the second stage of their life-cycle a wage that fully incorporates college premium with no uncertainty and consume out of all income. Values under this model specification match the values reported in Table 2, the monetary return with no uncertainty.

The second column in Table 5, shows the monetary return to college education under a linear utility specification, considering the dropout risk from Table 3. The excess of return to college education is reduced

[14] Considering a 4% risk free asset. Values in college wage premium units are shown in Appendix A. The mapping to convert the units corresponds to the one described in Table 4.

by 21% (25% in 1980). Agents are willing to reduce the college return to avoid college dropout.

The difference in returns to college education estimated by Model 1 and Model 2 is explained by the inclusion of the permanent income shock. This source of uncertainty explains about 1% of the excess of return to college education (2% in 1980). (Chen (2002), who analyzed data from NLSY79, reports that 23% of the college return is explained by this average risk differential.[15] Risk averse agents who face uninsurable risk in college returns, specifically the uncertain college premium, require a larger return to compensate for the risk.) The effect of the permanent income shock on wages is much larger in an environment with dropout risk, see below in Model 4.

Adding the second source of uncertainty, by only considering the dropout risk, yields the specification of Model 3. Risk averse agents who decide to accumulate human capital, face the probability of failing to graduate from college. The probability of college success is reported in Table 3. The differences between the excess of returns to college education under Model 1 and Model 3 specifications, are explained by the dropout risk, that accounts for 22% of the estimated excess of return to college education (25% in 1980). Dropping out implies a lower return to college investment, that corresponds to a fraction of the total college premium and corresponds to 55% of the college premium in 1960, decreasing up to 27% of college premium in 2006. Comparing outcomes of model 3 and linear utility specification, it is possible to see the risk aversion effect on the excess of return generated by dropout risk, that corresponds to 0.52% during the time period analyzed.

Model 4 combines both sources of uncertainty, agents who decide to attend college may fail to graduate and obtain only a fraction of the college premium (as in Model 3). Those who successfully graduate face a second source of risk: a permanent shock on wages (as in Model 2). The full model including both sources of uninsurable risk accounts for 51% of the excess of return to college, 73% in 1980 (this values are obtained by comparing estimation outcomes from Model 1 and Model 4).

This sections applied a simple risk premium approach to estimate the excess of return explained by the risk in college education, that have been estimated around 51%. Next section explores further this issue analyzing the role played by heterogeneous ability and selection in a life-cycle setup.

[15]Chen (2002), performs a certainty equivalent approach to estimate how much return to college education is explained by the wage volatility differential between high school and college graduate wages.

3 A life-cycle model with heterogeneous ability

In this section I analyze the role played by heterogeneity in ability levels and the interaction between individual ability and risky college investment. The model is built and calibrated taking college dropout as the sole source of uninsurable risk of college enrollment. The model reproduces wage distributions by educational and ability levels, college participation and dropout rates by ability levels as observed in the data. A second simulation of the model is performed without considering the dropout risk, all students who enroll successfully graduate from college. The non-risky college environment generates larger enrollment rates across all ability levels, since life time utility for college participants is higher for students from all ability levels. I then adjust the college premium to match the participation rates observed in the data for each ability level, obtaining a measure of return to college in environments with and without dropout risk. I interpret the differential return between these two measures as an estimate of the college risk premium.

3.1 The model

In this section I describe the model used to explain to what extent the return to college is explained by its risk and evaluate the role played by heterogeneous ability. I develop a three stage life-cycle model with a discrete choice of college enrollment and exogenous college dropout. I assume that the economy is populated by a unitary mass of heterogeneous agents that derive instantaneous utility solely from consumption. Schooling decisions are made in the first stage and are based on the lifetime utility maximization problem that each agent faces. Individuals may obtain three levels of education: no college, some college and college education. Education and employment are mutually exclusive in each period.

The life-cycle of an agent has three different phases. In the first phase, agents draw their type, a pair $\{x, y\} \equiv \Omega$, that corresponds to ability and family income levels from the joint distribution $H(x, y), x \in [\underline{x}, \overline{x}] \equiv \Phi$, $y \in [\underline{y}, \overline{y}] \equiv \Gamma$ and H: $\Phi \times \Gamma \rightarrow [0, 1]^2$. The second dimension in individual heterogeneity, family income, is included primarily for calibration purposes. Results are integrated out over this variable. In this stage agents decide whether to enroll in college or join the labor force. This decision is a function of their type and current wage offer. Agents receive this wage offer from the non-college wage distribution, $w^N \sim F^N(w)$. They simultaneously observe the probability of dropping out and the wage distributions of college graduates and college dropouts. The dropout probability is a function of individual types, $\psi = \psi(\Omega)$. I assume that individuals have perfect foresight about the skill price distributions and that wages depend on ability and are constant over the life-cycle. With the high-school wage offer in hand and observable college-success probability, college-graduate and college-dropout wage distributions, agents make their decisions about college attendance and labor market participation.

The decision for agent i about whether or not to participate in college at the first stage is based on the following optimization problem: $V(\Omega_i; w_i^N) = max\{V^C(\Omega_i), V^N(\Omega_i; w_i^N)\}$, where $V^C(\Omega_i)$ is the life-time utility of attending college and $V^N(\Omega_i; w_i^N)$ is the life-time utility of not attending college.[16]

Agents who attend college during the first stage consume c and are allowed to borrow at a subsidized interest rate, ρ; they also have to pay a college tuition cost, which equals τ. In order to finance their education, students receive grants and scholarships which are a function of ability and family income, represented by $g(\Omega)$. These agents also receive instant utility from college attendance, $\phi(\Omega_j)$. This utility is treated as a residual in the model since it is not measured in the data and is a function of an effort cost of college and the consumption value of schooling in utils. $\phi(.)$ has an important interpretation as the consumption value of schooling, summarizing non-pecuniary benefits of acquiring college education. The natural borrowing limit is also imposed to rule out Ponzi schemes.

The discounted life-time utility of college attendees of type Ω_j is given by the following expression:

$$V^C(\Omega_j) = \max_{c,a'} \left\{ u(c) + \phi(\Omega_j) + \beta \int_w \{\psi(\Omega_j) V^{CS}(\Omega_j, a') + (1 - \psi(\Omega_j)) V^{CD}(\Omega_j, a'; w)\} dF^{CD}(w) \right\} \quad (3)$$

$$s.t.$$

$$c + \rho a' I_{(a'<0)} + a' I_{(a'>0)} + \tau(\Omega_j) = g(\Omega_j)$$

$$a' \geqslant -\underline{a}$$

College attendees derive utility from consumption and discount future utility at rate β. They maximize life-time utility subject to budget and borrowing constraints. Agents successfully graduate or dropout from college in the second stage of their life-cycle, with probabilities $\psi(\Omega_j)$ and $1 - \psi(\Omega_j)$, in which case their discounted life-time utilities are given by $V^{CS}(\Omega_j, a')$ and $V^{CD}(\Omega_j, a'; w)$, respectively.

Equation (3) also describes the trade off faced by high school graduates when making the college enrollment decision. Graduating from college is associated with a higher expected wage, but the cost of financing

[16] See Appendix B for a detailed description of the solution method.

education may lead to negative asset holdings. The additional potential for dropout, makes college enrollment a risky investment decision. Individuals who attend college in the first period but drop out in the second period are likely to receive a lower wage offer and may have accumulated debt when they join the labor force.

Agents who choose not to enroll in college in the first stage of their life-cycle and join the labor force after high school graduation face a consumption-saving decision problem described by:

$$\begin{aligned} V^N(\Omega_j; w_j^N) = \quad & \max_{c,a'} \left\{ u(c) + \beta W\left(a'; w_j^N\right) \right\} \\ & s.t. \\ & c + a' = w_j^N \\ & a' \geqslant -\underline{a} \end{aligned} \tag{4}$$

Agents who decide to join the labor force consume c, borrow or save a' and receive ability-dependent wage compensation. They face the natural borrowing limit and discount future utility at rate β. $W(a; w)$ corresponds to a life-time utility of agents at the working stage. The value of $W(a; w)$ is obtained through a standard consumption-saving utility maximization problem with no uncertainty, where the state variables are the asset/debt level and wage. A complete description of $W(a; w)$ is provided in equation (7).

The life-time utility of individuals without college education increases with wage. On the other hand, receiving a higher wage offer in the first stage of the life-cycle reduces the probability of college enrollment.

In the second phase of their life-cycle, agents who are enrolled in college face the dropout risk, with probability $\psi(\Omega_j)$ agents successfully graduate college and obtain utility defined by $V^{CS}(\Omega_j, a')$. With probability $1 - \psi(\Omega_j)$ agents dropout, in which case agents draw a wage from the college-dropout wage distribution and obtain a life-time utility described by $V^{CD}(\Omega_j, a'; w)$.

If agents continue college, they will continue paying tuition, will receive grants and will be allowed borrow at the subsidized interest rate in the second stage of their life-cycle. Agents continue deriving utility from college as in the previous stage. The maximization problem of a college student in the second stage of the life-cycle takes the following form:

$$
\begin{aligned}
V^{CS}(\Omega_j, a) = \quad & \max_{c,a'} \left\{ u(c) + \phi_2(\Omega_j) + \beta \int_w W(A; w) dF^C(w) \right\} \\
& s.t. \\
& c + \rho a' I_{(a'<0)} + a' I_{(a'>0)} + \tau(\Omega_j) = g(\Omega_j) + a(1+r) I_{(a>0)} \\
& a' \geqslant -\underline{a} \\
& a I_{(a<0)} + a' = A \\
& A \geqslant -\underline{a}
\end{aligned} \tag{5}
$$

Note that in the maximization problem, the borrowing constraint faced by the agent considers the total accumulated debt, A.

Agents who drop out after spending the first stage of their life-cycle in college have to pay their accumulated debt. In each remaining period of their life-cycle they maximize a consumption-saving problem, $W(a; w)$. Their discounted life-time utility in the current second stage is described by:

$$
\begin{aligned}
V^{CD}(\Omega_j, a; w) = \quad & \max_{c,a'} \{ u(c) + \beta W(a'; w) \} \\
& s.t. \\
& c + a' = a(1+r) + w \\
& a' \geqslant -\underline{a}
\end{aligned} \tag{6}
$$

The third stage of the life-cycle is the working stage for those who graduate from college; the working stage arrives earlier for those who decided not to enroll or for those who receive the dropout shock. For simplicity it is assumed that the dropout shock arrives in the middle of the college education process.

Agents who graduate from college draw a wage from the college wage distribution, $w^C \sim F^C(w)$. In the following periods they face a consumption-saving problem, given by $W(a, w)$. The life time utility at the third stage of the life-cycle is a function of acquired education in the earlier periods. It is specified as follows:

$$W(a,w) = \max_{c,a'} \{u(c) + \gamma\beta W(a',w)\} \quad (7)$$
$$st.$$
$$c + a' = a(1+r) + w$$
$$a' \geqslant -\underline{a}$$

Equation (7) describes the consumption-saving problem agents face when they join the labor force. I include the survival probability γ to match the expected duration of labor force participation (and the retirement age). The maximization problem is solved given the budget constraint and natural borrowing limit. Utility in the working stage is increasing in assets and wages. Since wages are ability dependent, the life-time utility is also increasing in ability level.

3.2 Data and Calibration

3.2.1 Data

The analysis uses data from the National Longitudinal Survey of Youth 1979 cohort. It provides a nationally representative sample of young men and women aged 14-22 at the beginning of 1979. For the individuals in the sample, college attendance decisions took place in the early 1980s. excluding youths who are part of the minority and poor white oversamples (I use only the full random samples in the analysis). The data contains detailed information on individuals, including their ability level, family income and other family and personal characteristics.[17]

The data source contains a measure of ability, Armed Forces Qualification Test - AFQT scores, widely used in the literature as a measure of cognitive achievement, aptitude and intelligence. I use the AFQT89 variable as a proxi for cognitive ability.[18]

Another key variable in this analysis is family income. NLSY79 reports family income measured in early survey years. I use average family income when respondents are ages 16-17.[19] I denominate the family

[17] See Appendix C for summary statistics.

[18] AFQT89 is not adjusted by age. I thank Lance Lochner for pointing this out. I follow the age-correction procedure suggested by Carneiro et al. (2005).

[19] When income is available only for age 16 or age 17 and not both, I use the available measure.

income measure in 2007 dollars using the consumer price index for all urban consumers.

Whithin the sample, college attendance and dropout decisions took place in the early 1980s. Following Belley and Lochner (2007), an individual is considered to have attended college if their highest grade attended is equal or greater than 13. Similarly, it is considered a college dropout if the agent attends college but does not graduate.

The raw NLSY79 contains information on 6,111 individuals.[20] Individuals with no information about AFQT scores, family income or schooling were dropped from the sample. The final sample contains 2,477 individuals. Descriptive statistics for the variables used in the analysis are provided in Appendix C.

3.2.2 Calibration

The proposed model is calibrated to match college participation, college dropout and wage distributions observed in the data.

This section discusses the choice of parameters used in the model. The model has a set of 18 parameters. I divide the parameter space into three subsets. The first corresponds to parameters that I impose in the model from pre-existing estimates in previous literature. The second set corresponds to parameters estimated directly from the data. I calibrate the remaining parameters to match certain moments in the sample.

External parameters: The first subset of parameters and their values are reported in table 6.

I use a CRRA utility function with coefficient of risk aversion σ. Agents enter the first stage as an 18-year-old high school graduate. The duration of the first and second stages of the life-cycle is two years each; during the first period agents make their college enrollment decision. The third stage has an infinite horizon; I add a survival probability to the specification to match the retirement age.

Parameter		Value	Target/Source
Coeff. of risk aversion	σ	2	standard
Discount factor	β	0.96	standard
Prob. to survive	γ	0.957	to match 65 yrs.
Interest rate	r	4%	standard
Subsidized int. rate	ρ	0.9246	to set $r_{t=1} = 0$

Table 6: Imposed parameters in the model

The coefficient of risk aversion, the discount factor and the interest rate are chosen following standard

[20] Considered only the cross-sectional representative sample.

practice in the literature. I use a survival probability in the third stage (working stage) maximization problem to match a retirement age of 65 years. Finally, the subsidized interest rate was chosen to create a zero cost for those agents who borrow to finance their education.

Estimated parameters: The second subset of parameters are estimated from the data set. They include the tuition cost, grants awarded and the ability/family income distribution.

I follow the previous literature and estimate an average tuition (Akyol and Athreya 2005, Caucutt and Kumar 1999, Gallipoli et al. 2007, Garriga and Keightley 2007). Tuition is reported only in 1979. The annual average tuition cost is estimated to be \$4,350 (2007 dollars).

To estimate grants and scholarships awarded, I follow the mehodology of Gallipoli et al. (2007), who suggest a linear specification in ability and family income. Using data from NLSY79 I estimate the following equation for log-grants:

$$g(\Omega_i) = \alpha_0 + \alpha_1 x_i + \alpha_2 y_i + \theta X_i + \alpha_\lambda \widehat{\lambda}_i + \varepsilon_i, \tag{8}$$

where X_i is a set of controls for individual and family characteristics.

To estimate the above equation requires correction for selection bias: grants are not observed for those high school graduates who do not enroll in college. To correct for this selection, I implement the conventional two-step selectivity adjustment procedure suggested by Heckman (1979). In the first stage I formulate an econometric model to estimate the probability of attending college, which is used to predict the probability of college enrollment for each individual, $\widehat{\lambda}_i$.[21] In the second stage, I correct for the selection problem by including the predicted individual probability as an additional explanatory variable in the grants specification. See Appendix D for details, parameter values and estimated effects of individual and family characteristics on college attendance. The parameters of interest are reported in the following table and estimation results show that grants increase in ability and decrease in family income.

Grants	*constant*	*ability*	*family income*
NLSY79	11.40	0.23	-0.35
	(1.99)	(0.11)	(0.18)

Table 7: Parameters estimated for grant equation

To solve the model I impose the joint probability distribution of ability and family income observed in

[21] Parental education is used as instrumental variable in the first stage estimation.

NLSY79. Ability is normalized to have zero mean and standard deviation equal to one. Family income is expressed in natural logs. Figure 1 shows the marginal densities:

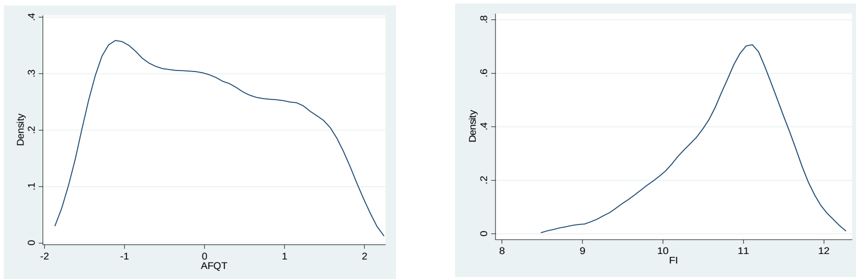

Figure 1: Marginal densities NLSY79: (a)Ability, (b)Family Income distributions

Calibrated parameters: The third subset of parameters is calibrated using simulated method of moments. I calibrate parameters of the wage distributions, probability of college success and the college utility function to match key moments observed in the data.

Wage distributions are calibrated to match means and standard deviations of life-cycle wages for each ability and educational level. Probability of college success is calibrated to match average dropout rates by ability and family income levels. College utility, which I assume to have a linear specification in ability and family income, is calibrated to match the profile of college enrollment rates.

I use the NLSY79 data to obtain wage distributions for high school graduates, college graduates and college dropouts. The annual mean wage is \$22,549 for high school graduates, \$25,399 for those who drop out from college and \$29,158 for college graduates.[22] The underlying college premium is consistent with the one reported by Goldin and Katz (2007) who use the 1980 Current Population Survey (CPS).

To generate the model wage profile, I estimate the return to ability and experience for different educational levels.[23] These estimates are used to project wage profiles along the life-cycle for different ability and education groups. I use the estimated wage paths to calculate a mean wage and its standard deviation for each ability and education level. Parameter estimates are reported in Appendix D.

At each educational level agents receive ability dependent wage offers. I define the wage offer function as follows: $w^{educ} = \varpi_0 + \varpi_1 x_i + \varepsilon_w$. I assume that these wage offers are constant throughout each agent's life and are generated such that the accepted wages for each ability and educational level match the mean and

[22] See Appendix D for details. Wages are nominated in 2007 dollars.
[23] I use the standard Mincer approach.

standard deviation of the life-cycle wages estimated in the data.

Probability of college success is a function of the agent's type. The probability parameter was calibrated to match dropout rates observed in the data per each ability and family income group. The college utility parameters are not measured in the data and are considered as a residual in the model. These parameters are calibrated such that the model reproduces some relevant features observed in the data, particularly the college attendance rates for each ability level. Appendix D summarizes the parameters estimated from the model.

Calibration results: Here I document how the model performs by comparing model generated statistics with those observed in the data.

Grants and scholarships from the model and data are reported in the following table.[24]

Grants	Data	Model
NLSY79	2,128	2,394
	(556)	(819)

Table 8: Annual grants: Model vs. Data

As can be seen from Table 8, the model slightly overestimates average and standard deviations for annual grants and scholarships observed in the data (values are nominated in 2007 dollars).

The following tables show mean log wages by ability quartile and educational levels.

	1980	
Ability quartile	Model	Data
lowest - Q1	10.030	10.049
	(0.62)	(0.63)
Q2	10.164	10.196
	(0.62)	(0.63)
Q3	10.324	10.360
	(0.62)	(0.63)
highest - Q4	10.513	10.517
	(0.62)	(0.63)

Table 9: Average log-wage for college graduate

Note: Average wage over the life-cycle for different ability levels, measured as AFQT score. Standard deviation in parentheses. Source is NLSY79.

Table 9 shows that the model performs well in replicating the wage premium for each ability level. High-ability college graduates (Q4) from the model get a 48% higher wage relative to those of low-ability level

[24] Note: standard deviations reported in parentheses.

(Q1). The wage structure for high school graduates is reported in the following table:

	1980	
Ability quartile	Model	Data
lowest - Q1	9.855	9.828
	(0.64)	(0.68)
Q2	9.964	9.952
	(0.69)	(0.68)
Q3	10.090	10.091
	(0.68)	(0.68)
highest - Q4	10.237	10.224
	(0.70)	(0.68)

Table 10: Average log-wage for high school graduates
Note: Average wage over the life-cycle for different ability levels, measured as AFQT score. Standard deviation in parentheses. Source is NLSY79.

As can be seen in table 10, in the high school wage distribution the ability premium is lower compared to the one for college graduates, and it is approximately 36%.

3.3 Results

Using the calibrated parameters I simulate the model and evaluate its ability to reproduce schooling attendance rates, dropout rates and labor income patterns observed in the data. To examine the performance of the model I match the detailed college participation and dropout profile observed in NLSY79. I name this college participation profile the risky allocation. The college particiation and dropout rates outcome's from the model and comparison with the data is shown in figure 2.

The outcome of a model with no college utility parameter ($\phi(\Omega_j)$) is of practical interest and is presented in Appendix E. The results show that this simpler model reproduces 45% of the overall college participation rate observed in NLSY79. Figure E.1 in Appendix E demonstrates that college participation rates are overestimated for low-ability students and underestimated for high-ability students. These results suggest that low ability individuals have some disutility from attending college, while high-ability students gain utility from college participation. Alternatively, it is possible that students who attend college must exert some effort which is decreasing with ability level.

The matched participation rates for 1980 are shown in Figure 2, termed here as the risky allocation, which serve as a starting point for the risk premium measurement. To quantify to what extent the dropout risk can explain the return to college observed in the data, I update the simulation of the model without

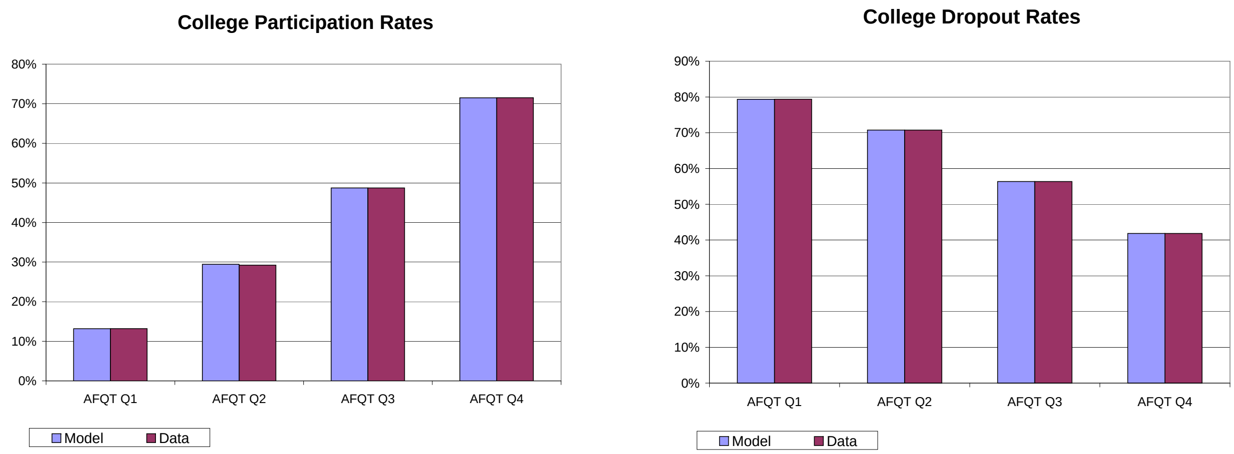

Figure 2: College Participation and College Dropouts rates

considering the dropout risk: if a high school graduate decides to attend college, he/she will successfully graduate. In this environment, $\psi(\Omega_j) = 1 \ \forall \Omega_j$. This alternative specification increases the life-time utility of college participation since all agents receive the college premium with certainty. This implies that enrollment rates increase for the entire ability profile. I name this college participation profile the non-risky allocation. College participation rates in an environment with and without dropout risk are presented in figure 3.

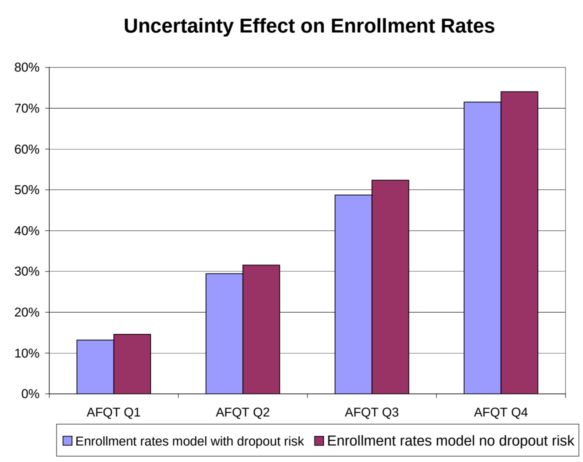

Figure 3: Dropout Risk Effect on College Enrollment Rates

To measure how much of the return to college observed in the data is due to risk premium, I adjust the college premium downwards in the model specification with no dropout risk, as depicted in Figure 3, such that college participation decreases from the non-risky allocation to the risky allocation profile. The amount of adjustment in the college wage premium corresponds to a quantification of the risk premium. The following table shows the college premium of the model with risk and without risk required to generate the participation rates from the risky allocation. The college premium is measured as the difference in average log wage between a college educated worker and a high school graduated worker.

	AFQT Q1	AFQT Q2	AFQT Q3	AFQT Q4
College Premium	22.70%	23.32%	24.61%	26.52%
College Premium (no Risk)	18.36%	19.29%	19.69%	20.86%

Table 11: College Wage Premium

Note: The college wage premium reported corresponds to the average log wage differential that equates college participation rates between college graduate and high school graduates, in risky and non-risky environments. See Figure 3.

For students with an ability level in the lowest quartile, AFQT Q1, the dropout risk explains about 19% of the observed college wage premium. For students from the top quartile of the ability distribution, AFQT Q4, the respective figure is 21%.

In terms of the monetary returns to education, to make these results comparable with the ones reported in the previous section (risk premium approach), the following table shows the excess monetary return in an environment with and without dropout risk.

	AFQT Q1	AFQT Q2	AFQT Q3	AFQT Q4
Monetary Return	3.36%	3.50%	3.78%	4.18%
Monetary Return (no Risk)	2.44%	2.64%	2.72%	2.98%

Table 12: Excess of Monetary Return to College Education

Note: Monetary return for a college educated worker under a risky and non-risky college scenario that generates the non risky college participation rates.

The dropout risk in an environment in which I allow for selection and heterogeneity in ability level explains about 27% of the excess return to college education for low ability students and about 29% for those students with a high ability level. Dropout rates in the NLSY79 sample reach about 52% in average across ability levels. Dropout risk explains about one third of the return to college education. The effect is increasing in ability level.

4 Conclusion

This paper analyzes the excess return to college education in an environment where human capital accumulation is risky. The risk arises from two sources: a permanent income shock on wages after college graduation (as in Chen 2002) and the potential for college dropout (as measured about 52% of college participants). Risk adverse individuals prefer a lower return to college education if the risk associated with college completion is reduced

Attending college has been considered one of the most profitable investments, annualized returns ranging from 8% to 13% (Card 1999). I utilize a simple approach, as in Mehra and Prescott (1985), to quantify how much of the excess return to college is explained by its risk. I also explore the role played by heterogeneous ability by developing a life-cycle model with endogenous enrollment and analyzing the dropout risk effect on college returns.

Under the risk premium approach, the permanent income shock on wages explains 1% of the excess return to college education. Dropout risk explains 22% of the estimated excess return. Both sources of risk combined explain 51% of the excess return to college education.

To explore the role of individual heterogeneity, a life-cycle model setup is used to estimate how much of the college return is explained by dropout risk. The model is calibrated to match key moments observed in the data and it is simulated in an environment with and without risk. The college dropout risk explains about 27% of the college return for students with a low ability level and about 29% of the return for students with a high ability level.

Dropout risk reconciles two empirical facts: a high return to college education with low enrollment rates. Previous literature usually relies on selection bias as an explanation for these facts.

References

Akyol, A. and K. Athreya (2005). Risky higher education and subsidies. *Journal of Economic and Dynamic & Control 29*, 979–1023.

Belley, P. and L. Lochner (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human Capital 1(1)*, 37–89.

Board, C. (2007). Trends in higher education series. College Board.

Card, D. (1999). The caussal effect of education on earnings. *Handbook of Labor Economics 5*, 1801–63.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69, No 5*, 1127–1160.

Carneiro, P., J. Heckman, and D. Masterov (2005). Labor market discrimination and racial differences in premarket factors. NBER Working Papers 10068, National Bureau of Economic Research, Inc.

Carnoy, M. and D. Marenbach (1975). The return to schooling in the united states, 1939-69. *The Journal of Human Resources 10 No3*, 312–331.

Caucutt, E. and K. Kumar (1999). Higher education subsidies and heterogeneity a dynamic analysis. mimeo.

Chen, S. (2002). Is investing in college education risky? University of Rochester and University of New York at Albany. Mimeo.

Gallipoli, G., C. Meghir, and G. Violante (2007). Equilibrium effects of education policies: a quantitative evaluation. mimeo.

Garriga, C. and M. Keightley (2007). A general equilibrium theory of college with education subsidies, in-school labor supply, and borrowing constraints. mimeo.

Goldin, C. and L. Katz (2007). The race between education and technology: The evolution of u.s. educational wage differentials, 1890 to 2005. NBER working paper 12984.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica 47*, 153–61.

Heckman, J., L. Lochner, and P. Todd (2008). Earnings functions and rates of return. *Journal of Human Capital 2(1)*, 1–31.

Ionescu, F. (2009). Federal student loan program: Quantitative implications for collge enrollment and default rates. *Review of Economics Dynamics 12 (1)*, 205–231.

Mayer, A. (2008). The old college try: Estimating the role of uncertainty, costs and rewards in the experiment called college. Working paper, Texas A & M University.

Mehra, R. and E. C. Prescott (1985). The equity premium: A puzzle. *Journal of Monetary Economics 15*(2), 145–161.

Restuccia, D. and C. Urrutia (2004). Intergenerational persistence of earnings: The role of early and college education. *American Econ Review 94(4)*, 1354–1378.

Willis, R. and S. Rosen (1979). Education and self-selection. *J Political Econ 87(5)*, 7–36.

A Data: U.S. From 1960 to 2007

In this appendix I report summary statistics from the CPS data used to quantify the return to education from 1960 to 2007.[25]

	High school			College dropouts			College graduates		
year	number obs.	average wage	sd wage	number obs.	average wage	sd wage	number obs.	average wage	sd wage
1960	93,372	5,154	2,930	31,370	5,911	3,690	34,182	7,228	4,487
1970	129,755	8,053	4,317	43,316	9,323	5,561	51,564	12,216	7,244
1980	160,370	15,556	8,726	67,550	17,603	10,527	83,987	22,572	13,255
1990	159,181	22,600	15,248	113,636	27,596	19,953	104,233	40,708	31,356
2000	161,915	29,068	22,844	131,272	36,201	30,142	123,097	58,502	55,268
2001	68,196	30,174	22,339	58,180	37,843	29,434	62,272	64,793	60,706
2002	59,569	30,529	22,535	51,280	38,397	30,127	56,003	62,612	56,892
2003	65,775	30,910	22,848	56,797	38,343	29,080	62,602	63,868	56,156
2004	64,068	31,495	23,287	57,114	40,101	30,747	63,875	65,724	50,363
2005	154,685	32,399	24,342	135,202	41,261	31,647	147,662	69,588	54,214
2006	160,055	32,582	25,339	136,440	41,301	32,026	150,614	71,691	65,889
2007	159,335	34,035	27,385	136,517	43,273	34,622	154,188	75,977	70,935

Table A.1: Average wage and salary income by educational level across time

Tuition cost for the period analyzed is reported in Table A.2. For details see Board (2007)

Year	1960	1970	1980	1990	2000	2001	2002	2003	2004	2005	2006	2006
Tuition	3,200	3,500	3,800	8,000	8,000	8,400	8,700	9,100	9,500	9,600	9,900	10,300

Table A.2: College tuition costs. Source: College Board, values in 2007 dollars adjusted by CPI.

Table A.3 reports the excess return to college education under the four model scenarios, as described in Section 2 Table 12. The measure of the excess return corresponds to a log wage difference.

	Model 1	Model 2	Model 3	Model 4
Year	No risk	Permanent shock	Dropout risk	Perm. shock + Dropout risk
1960	14.64%	14.56%	13.34%	10.80%
1970	20.17%	20.09%	17.08%	14.34%
1980	20.06%	19.94%	17.20%	14.39%
1990	30.28%	30.20%	26.34%	23.72%
2000	38.16%	38.07%	32.97%	30.17%
2005	42.40%	42.31%	35.61%	32.54%

Table A.3: Excess return to college education. Measured in low wage differential

B Solution Method

In this Appendix I propose an analytical solution to the model. Given the nature of the life-cycle environment, the model is solved backwards from the third stage of the agent life-cycle.

[25]Raw data available upon request to the author.

The individual's maximization problem at the working stage, $W(a,w)$, has a simple analytical solution since there is no uncertainty in the final stage of the life-cycle. Given the asset level at the beginning of this stage and the wage drawn, life-time utility is given by the following specification:

$$\begin{aligned} W(a_o,w) &= \max_{\{c_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t) \\ & s.t. \\ c_o + \frac{c_1}{1+r} + ... &= a_0 + w + \tfrac{w}{1+r} + ... \end{aligned} \tag{B.1}$$

Life-time utility at the working stage is increasing in assets and wages. Since wages are ability dependent, the life-time utility is also increasing in ability level.

Equation(B.1) also describes the trade off faced by high school graduates when facing the college enrollment decision. Graduating from college is associated with a higher expected wage, but financing education costs may lead to negative asset holdings. Taking into account the college dropout shock at the second stage of the life-cycle, college enrollment is also a risky investment decision. Individuals who attend college in the first period but drop out in the second period are likely to receive a lower wage offer but may have a larger accumulated debt when they join the labor force.

The saving decision of agents who choose not to enroll in college is given by equation (B.2).

$$a^*_{NC} = \frac{\left[\beta(1+r)\Psi_1\Psi_2\right]^{\frac{1}{\sigma}} w - \frac{w(1+r)}{r}\Psi_1}{(1+r)\Psi_1 + \left[\beta(1+r)\Psi_1\Psi_2\right]^{\frac{1}{\sigma}}} \tag{B.2}$$

where $\Psi_1 \equiv \left(\frac{1+r-[\beta(1+r)]^{\frac{1}{\sigma}}}{1+r}\right)$ and $\Psi_2 \equiv \left(\frac{1}{1-\beta[\beta(1+r)]^{\frac{1-\sigma}{\sigma}}}\right)$.

The saving decision is explained by consumption smoothing. Agents who do not enroll in college in the first period of their life-cycle do not face any uncertainty in future stages and therefore have no precautionary motives affecting their saving decisions. The discounted life-time utility of individuals who join the labor force as high school graduates is given by:

$$V^N(\Omega_j, w_j^N) = u(c^*) + \beta W\left(a^*, w_j^N\right) \tag{B.3}$$
$$= \frac{\left(w_j^N - a^*\right)^{1-\sigma}}{1-\sigma} + W(a^*, w_j^N)$$

The life-time utility of individuals without college education increases with their wage. On the other hand, receiving a higher wage offer in the first stage of the life-cycle reduces the probability of college enrollment.

To evaluate the saving decision of individuals who enroll in college in the initial stage of the life-cycle I numerically solve the following equation.

$$\left(g(\Omega_j) - \tau - a^*\left[\rho I_{(a^*<0)} + I_{(a^*>0)}\right]\right)^{-\sigma} = \tag{B.4}$$
$$\beta \int_w \left[\left((a+a^*)(1+r) + w\frac{1+r}{r}\right)\Psi_1\right]^{-\sigma}(1+r)\Psi_1\Psi_2 dF^C(w) \tag{B.5}$$

This equation is obtained from the first order conditions of the college participation problem. The savings of individuals who enroll in college is a decreasing function of both their value of college participation, determined by transfers, and expected wage following graduation or dropout. Their saving/borrowing decisions are also affected by consumption smoothing motives and precautionary motives.

Given solutions for optimal saving decisions and wage offers, I estimate the life-time utilities at working stages for each ability and family income level. With the set of estimated life-time utilities in hand, I proceed to evaluate college enrollment rates and labor force participation patterns for each ability-family-income group.

C Data: Summary Statistics NLSY 1979

NLSY Descriptive statistics for the sample from NLSY79.

Table C.2 shows the summary statistics for different quartiles of ability and family income. Source is NLSY79, AFQT as measure of cognitive ability.

Sample Descriptive Statistics	NLSY79
Male	49.80%
Black	12.53%
Hispanic	8.25%
Completed high school	75.63%
Attended college	40.73%
Completed at least one year of college	28.16%
Urban residence at age 12	76.05%
Number of siblings	2.9
Mother HS graduate	67.36%
Father HS graduate	66.55%
Family income ($10,000, 2007 dollars)	5.789
Sample size	2,477

Table C.1: Summary statistics for NLSY79 sample

Note: AFQT was normalized to a mean zero and a unit standard deviation.

D Parameter Estimates

Estimated effects on college participation (selection equation used to implement the Heckman (1979) two-step procedure) are shown in Table D.1.

The parameters estimated for the grant equation, $g(\Omega_i) = \alpha_0 + \alpha_1 x_i + \alpha_2 y_i + \theta X_i + \alpha_\lambda \widehat{\lambda_i} + \varepsilon_i$ are reported in Table 7

Wage estimation from the data proceed as follows: I first estimate a wage profile along the life-cycle. I impose rational expectations over the wage structure, i.e., agents can observe the wage profile along the life-cycle for each educational level and ability level. I control for experience and ability and perform estimations for each educational level. The wage equation is estimated for the first cohort using NLSY79 data available for the 1979 - 2006 period. The wage specification is: $w_{it}^{educ} = \lambda_0 + \lambda_1 exp_{it} + \lambda_2 exp_{it}^2 + \lambda_3 x_i + \lambda_4 X_i + \varepsilon_{it}$. This equation represents the log-wage structure for an individual i in period t who has an educational level *educ*. The variable *exp* corresponds to experience and variable x to ability level. X_i is a set of controls for individual and family characteristics. In these estimations I use information on white males only, whose annual wages are between $3,000 and $280,000. In the sample I include only individuals who participate in the labor force and are not attending school at that time. Estimates are presented in the Table D.2.

Standard errors are in parenthesis. I project a hump shaped wage profile and estimate the average wage

	NLSY1979
AFQT*	0
	(1)
F. income	57,889
	(33,157)
AFQT Q1	-1.228
	(0.202)
AFQT Q2	-0.490
	(0.235)
AFQT Q3	0.353
	(0.259)
AFQT Q4	1.362
	(0.330)
F. income Q1	20,673
	(7,310)
F. income Q2	43,805
	(5,834)
F. income Q3	63,821
	(6,278)
F. income Q4	103,254
	(24,363)

Table C.2: Family income and AFQT summary statistics

	NLSY79
Sex	4.4048
	(1.34)
Hispanic	2.1059
	(2.67)
Race	0.4380
	(2.16)
Urban	0.2722
	(1.69)
Highest grade mother	0.0671
	(0.01)
Highest grade father	0.0431
	(0.01)
Siblings	0.4771
	(0.44)

Table D.1: Estimated effects on college participation

along the life-cycle. Average wages are reported in table D.3.

E Model Predictions

First, I solve the model without considering college utility. The model is employed to generate college attendance rates for each ability and family income group using parameters I estimatd from NLSY79. These simulated outcomes are compared to those observed in the data. This exercise allows me to evaluate the performance of the model while only using parameters estimated from the data. Figure E.1 presents the simulated and actual college attendance profiles of the NLSY79 cohort.

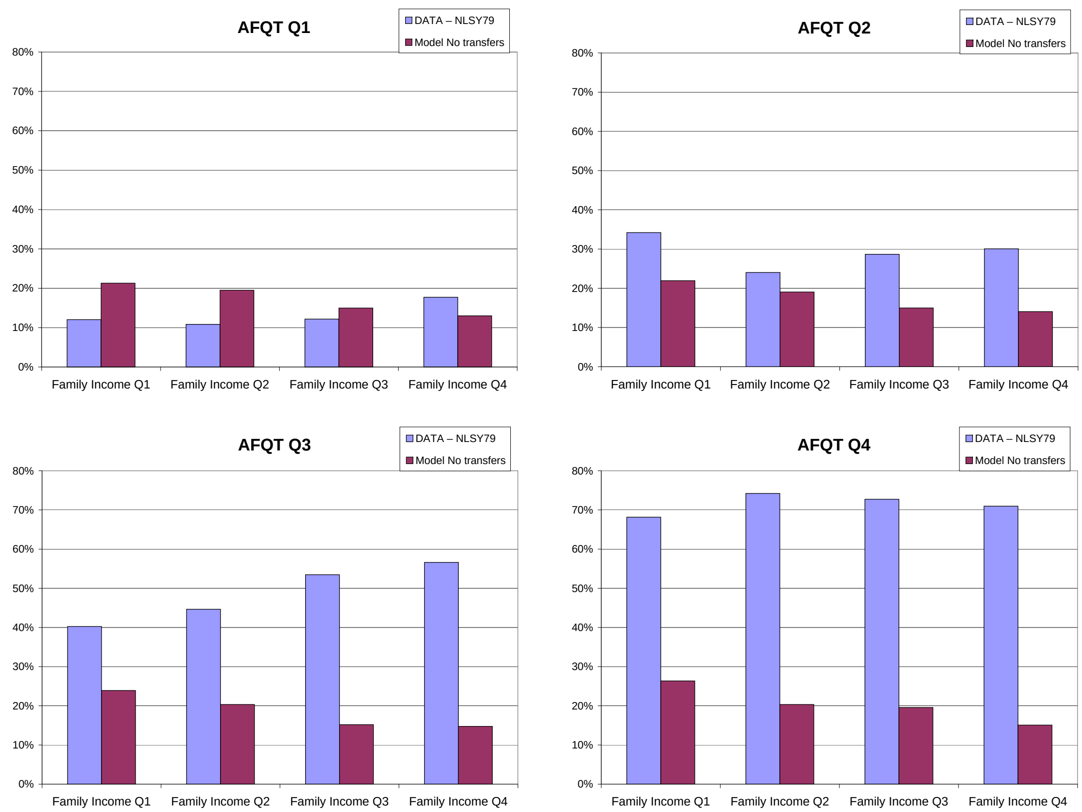

Figure E.1: College participation profile NLSY79-Data vs Model (no effor/cost utility parameter)

Adding the college utility parameters, $\phi(\Omega_j) = \delta_0 + \delta_1 x_i + \delta_2 y_i$, to the calibration of the model generates the participation rates provided in Figure 2 - left panel.

	High school	College dropout	College graduate
λ_0	9.046	9.071	9.344
	(0.02)	(0.04)	(0.09)
λ_1	0.123	0.122	0.118
	(0.004)	(0.007)	(0.012)
λ_2	-0.0027	-0.0026	-0.0027
	(0.0001)	(0.0002)	(0.0004)
λ_3	0.157	0.037	0.186
	(0.006)	(0.012)	(0.014)

Table D.2: Estimated parameters for wage equations

	college wage	dropout wage	non-college wage
NLSY79	29,158	25,339	22,549

Table D.3: Average wage per educational level, NLSY79 and NLSY97

	Data				Model			
	Ability							
Family Income	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Q1	12.09%	34.19%	40.25%	68.15%	12.09%	33.03%	50.37%	68.14%
Q2	10.83%	24.03%	44.65%	74.21%	13.12%	35.58%	51.00%	67.54%
Q3	12.17%	28.66%	53.46%	72.73%	16.26%	35.01%	53.57%	68.81%
Q4	17.72%	30.07%	56.60%	70.97%	17.70%	36.25%	56.77%	70.96%

Table E.1: College participation profile, Model vs Data, NLSY79