

# Insuring College Failure Risk\*

Satyajit Chatterjee<sup>†</sup>

Felicia Ionescu<sup>‡</sup>

*Federal Reserve Bank of Philadelphia*

*Colgate University*

September 2009

## Abstract

Participants in student loan programs must repay loans in full regardless of whether they complete college. However, the dropout rate among students from low-income background, the ones most likely to access the student loan program, is between 33 to 50 percent. We examine whether insurance against college-failure risk can be offered, taking into account moral hazard and adverse selection. To do this, we develop a model that accounts for US college enrollment, drop out, and completion rates. We find that optimal insurance raises enrollment and completion rates and raises average welfare by 3 percent.

JEL Codes: D82; D86; I22;

Keywords: College Risk; Government Student Loans; Optimal Insurance

---

\*The authors would like to thank participants at the NBER-EFACR group, the Society of Economics Dynamics, Econometrics Society and Midwest Macroeconomic Meetings and seminar series at Cornell University and University of Connecticut, especially to Orazio Attanasio, Lutz Hendricks, Narayana Kocherlakota, Luigi Pistaferri, José Víctor Ríos Rull, and Viktor Tsyrennikov.

<sup>†</sup> Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia PA, 19106; (215)574-3861, satyajit.chatterjee@phil.frb.org

<sup>‡</sup> Department of Economics, 13 Oak Drive, Hamilton NY, 13346; (315)228-7955, Fax: (315)228-7033, fionescu@colgate.edu.

# 1 Introduction

Recent research in the education literature provides support for the fact that financial constraints during college-going years are not crucial for college enrollment (Carneiro and Heckman (2002), Cameron and Taber (2001)). Rather, it is student characteristics, such as learning ability, that determine the decision to enroll. Given the generosity of the student loan program, funds are readily available and eligible high school graduates invest in college if they perceive the returns to a college education to be high enough (Ionescu (2009)).

However, there is considerable financial risk in taking out a student loan to pay for college. Using the 1990 PSID, Restuccia and Urrutia (2004) document that 50 percent of people who enroll in college drop out. Similarly, using the NCES data and surveys, we find that 36.8 percent and 35.2 percent of students enrolled in 1989-90 and 1995-96, respectively, do not possess a degree and are not enrolled five years after their initial enrollment. For these people, the foregone earnings and out-of-pocket expenses incurred while enrolled in college yield little or no financial return. This financial risk may discourage some people from taking out a loan and enrolling in college. Thus, even though prospective students do not appear to be credit constrained, a mechanism to share the risk of failing to complete college – *college failure risk* – might improve the welfare of enrolled students and encourage more people to enroll and complete college.<sup>1</sup>

The aim of this paper is to study whether the student loan program can offer insurance against college failure risk. The current operation of the program suggests that it is administratively feasible to offer some insurance. Under the current system, a borrower can choose from a menu of fairly sophisticated repayment options (standard, graduated, income-contingent and extended repayment). Nevertheless, under each of these payment options,

---

<sup>1</sup>Given a dramatic shift against unskilled labor in the U.S. labor market between 1980-1990 (decline in wages, labor activity, longer unemployment spells for low skilled), Heckman (1999) has argued that increasing the supply of college labor can reverse this trend. Specifically, for 1990 workforce of 120 million, about 5.4 million would have to become college equivalents to reverse the 1980-1990 erosion of real wage, and about 1 million additional skilled persons would need to be added to the workforce each year on top of the once and for all change of 5.4 million.

the borrower is required to repay the entire loan and associated interest expenses regardless of whether he or she completed college. We will examine whether it is feasible to forgive a portion of the loan for students who fail out of college.<sup>2</sup>

We conduct our investigation under two important constraints on the provision of failure insurance. First, we require that the insurance scheme not distribute resources from people with high probability of completion to people with low probability of completion (and vice versa). Formally, this requires that the insurance program be self-financing with respect to each person who chooses to participate. The current programs enforce this self-financing constraint regardless of whether the program participant actually graduates from college. We will permit failures to pay less than graduates, but each participant will pay the full cost of college in expectation. Second, we require that the insurance program guard against moral hazard (the possibility that insurance will increase the risk of failure due to shirking) and adverse selection (the possibility that poor risks will attempt to pool with the good risks). In this regard, it is worth noting that the moral hazard problem is not so severe as one might (initially) think because there is a large college premium in earnings. Shirking to take advantage of insurance means giving up on earning more later in life.

In the theoretical section of the paper, we develop a simple model of a student's enrollment, dropout and college effort decisions. The model postulates the necessary heterogeneity in student characteristics in order to be consistent with the diversity of enrollment, drop out and effort decisions we see in reality. The heterogeneity is in a student's utility cost of putting effort in college and his or her outside option, neither of which are directly observable to loan administrators. The unobserved heterogeneity complicates the task of providing insurance. These complications are analyzed in the theoretical section and the constrained optimization problem that delivers the optimal insurance program is developed.

In the quantitative section, we calibrate the model to US data on college enrollment, early leaving, and completion rates as well as the average indebtedness of program participants,

---

<sup>2</sup>The borrower is permitted to discharge her loan only if a repayment effort over 25 years does not fully cover all obligations.

distinguishing between students of different ability levels. We focus on individuals from low-income families – the ones for whom this sort of insurance is most relevant. These are individuals for whom the risk of college failure appears to be the highest currently.<sup>3</sup> Furthermore, these are also the individuals for whom, conditional on dropping out of college, the burden of a student loans can be substantial. We solve the optimal insurance problem for the calibrated model. We quantify the effects of insurance on enrollment and completion rates as well as welfare. The optimal insurance offered in case of non-completion ranges between 10 to 40 percent of total college cost. The insurance scheme induces an increase in enrollment rate of 5.4 percentage points and an increase in college graduates of 5.4 percentage points. Although insurance draws in students with a high risk of failure, completion rates rise because fewer students drop out voluntarily from college. Insurance increases welfare by 3 percent on average.

The rest of the paper is organized as follows. Section 2 presents the choices available to a student. Section 3 lays out the key predictions of this model when no insurance is offered and compares these predictions to patterns in the data. Section 4 develops the constrained optimization problem that delivers the optimal insurance scheme. Parameter selection and calibration of the model is presented in Section 5. Section 6 presents the results of offering insurance in the calibrated model and Section 7 concludes.

## 2 Environment

Time is discrete and indexed by  $t = \{0, 1, 2, \dots\}$ . In period 0, a prospective student makes a one-time decision to enroll in college or not. If she does not enroll, she can work in a low-paid job with disutility of effort  $\theta \geq 0$  and, starting in period 1, earn  $y \geq 0$ . The earnings  $y$  is drawn from a distribution  $H(y)$ . At the time of the enrollment decision, the student knows  $\theta$  but not the realization of  $y$ .

If the individual chooses to enroll in college, she learns the cost of making effort in college.

---

<sup>3</sup>See Manski (1992) and Stinebrickner and Stinebrickner (2007)

Effort,  $e$ , is a binary variable that can take values 0 (no effort) or 1 (effort). The cost of making effort is denoted  $\gamma$  and the student draws  $\gamma \geq 0$  from a distribution  $G(\gamma)$ . After she learns  $\gamma$  the student decides whether to continue on in college or not. If she chooses to leave, she incurs the cost of effort  $\theta$  in the low-paid job and draws her (life-time) earnings  $y$  in period 1. She also incurs some partial college expenses  $x/4$ , where  $0 < \phi < 1$ .<sup>4</sup> At the time of choosing whether to continue in college or not, the student knows  $\gamma$  and  $\theta$  but not her earnings in period 1 and beyond.

If the student continues in college she incurs the annual college cost of  $x$ . A continuing student must choose between putting in effort or not. If she chooses to shirk ( $e = 0$ ), she will fail with probability 1 but she will not incur effort costs of any kind in period 0 and will start life in period 1 with an earnings draw  $y$  from the distribution  $H(y)$  and a debt of  $x$ . If she chooses to put in effort ( $e = 1$ ), she will complete her first year with probability  $\pi \in (0, 1)$ . If she completes successfully she begins period 1 as a college student with one more year to go and debt of  $x$  (no interest accumulates on the debt as long the student continues in college). If she fails to complete, she starts period 1 with an earnings draw  $y$  from  $H(y)$  and a debt of  $x$ .

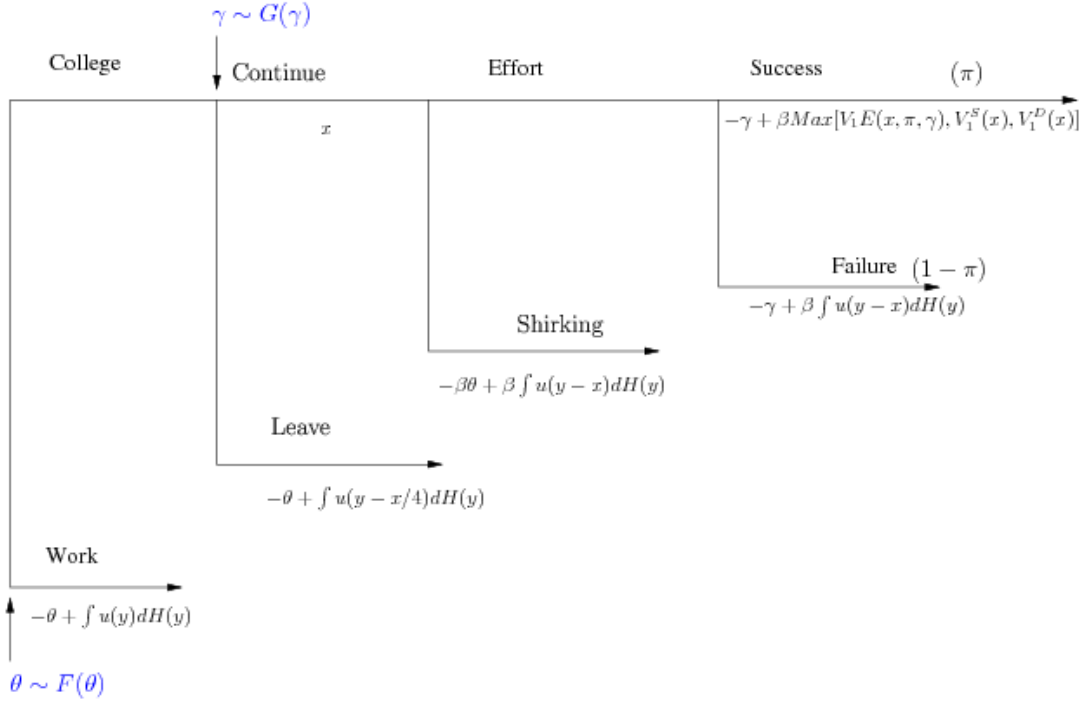
Figure 1 illustrates this timing of period 0 decisions. In the case where a student succeeds in completing the first year of college, she faces a similar decision tree in period 1 (which we will describe below).

In period 1, a student with one more year to go has to choose again whether to continue in college or not. If she does not continue, she gets an earnings draw  $y$  from the distribution  $H(y)$  starts her life with debt  $5x/4$ . If she continues, she incurs another year of college expense  $x$ . And, as in period 0, she must choose between putting in effort or shirking. If she shirks, she fails with probability 1 but does not incur any effort cost in period 1 and starts life in period 2 with an earnings draw  $y$  from the distribution  $H(y)$  and a debt of  $2x$ . If she puts in effort, she completes college with probability  $\pi$ . If she succeeds in completing,

---

<sup>4</sup>We assume that if a student voluntarily withdraws from college, he or she pays a cost that is some (relatively small) proportion of a year's college costs. We fix this proportion to be  $1/4$ .

Figure 1: Timing of Decisions



she draws her life-time earnings  $y$  from the distribution  $M(y)$  has debt of  $2x$ . If she fails to complete college, she starts period 2 with an earnings draw  $y$  from  $H(y)$  and a debt of  $2x$ .

In order to describe individuals decision problems in period 0 and 1 (these are the only periods in which there are decisions to be made), we will start with describing the utility (payoffs) to students at the start of period 1 (students that have one more year of college to go).

1. A student who drops out gets

$$V_1^D(x) = \int U(y - 5x/4) dH(y).$$

2. A student who continues but shirks gets

$$V_1^S(x) = \beta \int U(y - 2x) dH(y).$$

3. A student who continues and puts in effort gets

$$V_1^E(\pi, x, \gamma) = -\gamma + \beta \left[ \pi \int U(y - 2x) dM(y) + (1 - \pi) \int U(y - 2x) dH(y) \right].$$

Turning to period 0, the payoffs are as follows

1. An individual who does not enroll gets

$$W(\theta) = -\theta + \beta \int U(y) dH(y).$$

2. An Individual who enrolls, but drops out gets

$$V_0^D(x, \theta) = -\theta + \int U(y - x/4) dH(y).$$

3. An individual who enrolls, continues and shirks gets

$$V_0^S(x, \theta) = -\beta\theta + \beta \int U(y - x) dH(y).$$

4. An individual who enrolls, continues and puts in the effort gets

$$V_0^E(\pi, x, \gamma) = -\gamma + \beta \left[ \pi \max[V_1^E(\pi, x, \gamma), V_1^S(x), V_1^D(x)] + (1 - \pi) \int U(y_N - x) dH(y_N) \right].$$

The structure of payoffs is generally self-explanatory. One aspect worth remarking on is that leaving or shirking in period 0 forces the individual to work in the low-paid job for 1 period. In contrast, if the student fails in period 0 despite putting in effort, she does not have to work in the low-paid job. This assumption is a convenient way to capture the fact that exerting effort in college has benefits even if it does not lead to college credits. Also, since anyone who is in college in period 1 must have successfully completed one year of college (and therefore exerted effort in period 0), she can drop out or shirk and not have to work in the low-paid

job. Thus,  $\theta$  does not appear in either  $V_1^D(x)$  or  $V_1^S(x)$ .

We make the following set of assumptions on the primitives.

**Assumption 1:**  $U(c) : R \rightarrow R^{++}$  with  $U'(\cdot) > 0$  and  $U''(\cdot) < 0$ .

**Assumption 2:**  $\beta^2 \int U(y - 2x)dM(y) > \int U(y)dH(y)$  (college degree is profitable financial investment).

**Assumption 3:**  $\int z(y)dM(y) > \int z(y)dH(y)$  for any  $z(y)$  strictly increasing in  $y$  (the distribution  $M$  first-order stochastically dominates the distribution  $H$ ).

### 3 College Enrollment, Dropout and Failure Under the Current System

We begin by studying the choice problem in period 1. There are three options open to the student. She could drop out, or continue on in college but not put in any effort, or she could continue on in college and exert effort.

**Proposition 3.1.** *In period 1, there is a cut-off  $\gamma_1(x, \pi) \geq 0$  such that for  $\gamma > \gamma_1(x, \pi)$ , students drop out and for  $\gamma \leq \gamma_1(x, \pi)$  they continue on with effort. Furthermore,  $\gamma_1(x, \pi)$  is increasing in  $\pi$ .*

*Proof.* Since  $5x/4 < 2x$  and  $\beta < 1$ ,  $V_1^D(x) > V_1^S(x)$ . Hence, dropping out is strictly better than shirking in period 1. Therefore, the student chooses between continuing on with effort or dropping out. Denote the difference in payoffs between these two choices by  $V_1(x, \pi, \gamma) = V_1^E(x, \pi, \gamma) - V_1^D(x)$ . Observe that  $V_1(x, \pi, \gamma)$  is continuous and strictly decreasing in  $\gamma \in [0, \infty)$ . If  $V_1(x, \pi, 0) \leq 0$ , then  $\gamma_1(x, \pi) = 0$ . If  $V_1(x, \pi, 0) > 0$ , by continuity and strict monotonicity with respect to  $\gamma$ , there exists a unique  $\hat{\gamma} > 0$  such that  $V_1(x, \pi, \hat{\gamma}) = 0$ . Hence  $\gamma_1(x, \pi) > 0$ .

To prove  $\gamma(x, \pi)$  is increasing in  $\pi$  note that

$$\begin{aligned}
V_1(x, \pi, \gamma) &= -\gamma + \beta\pi \left[ \int U(y - 2x)dM(y) - \int U(y - 2x)dH(y) \right] \\
&\quad + \beta \int U(y - 2x)dH(y) - \int U(y - x)dH(y).
\end{aligned}$$

By Assumption 2,  $V_1(x, \pi, \gamma)$  is strictly increasing in  $\pi$ . Now consider  $\hat{\pi} < \tilde{\pi}$ . If  $V_1(x, \hat{\pi}, 0) < V_1(x, \tilde{\pi}, 0) \leq 0$ , then  $\gamma_1(x, \hat{\pi}) = \gamma_1(x, \tilde{\pi}) = 0$ . If  $V_1(x, \hat{\pi}, 0) \leq 0 < V_1(x, \tilde{\pi}, 0)$ , then  $0 = \gamma(x, \hat{\pi}) < \gamma(x, \tilde{\pi})$ . Finally, if  $0 < V_1(x, \hat{\pi}, 0) < V_1(x, \tilde{\pi}, 0)$ , then  $0 < \gamma(x, \hat{\pi}) < \gamma(x, \tilde{\pi})$ . This establishes that  $\gamma(x, \pi)$  is increasing in  $\pi$ .  $\square$

It is perhaps worth noting that the threshold  $\gamma$  will be zero for sufficiently low probability of success  $\pi$ . Observe that  $V_1(x, 0, 0) < 0$  and, by Assumption 2,  $V_1(x, 1, 0) > 0$ . Thus, when no effort in school is required ( $\gamma = 0$ ) there exists  $\bar{\pi}_1 > 0$  such that  $V_1(x, \bar{\pi}_1, 0) = 0$ . For all  $\pi < \bar{\pi}_1$ ,  $V_1^E(x, \pi, \gamma) - V_1^D(x) < 0$  for all  $\gamma \geq 0$ . Therefore, the threshold  $\gamma_1(x, \pi)$  is 0 for all  $\pi \leq \bar{\pi}_1$ .

We now study the choices in period 0. The choice problem can be broken down into two parts. First, conditional on not putting in effort in college, is it better to drop out or shirk? And, second, given the answer to the first question, is it better to put in effort in college?

**Proposition 3.2.** *In period 0, there exists a cut-off  $\theta_0(x) > 0$  such that conditional on not putting in effort in college students drop out for  $\theta < \theta_0(x)$  and shirk for  $\theta \geq \theta_0(x)$ .*

*Proof.* Consider the function  $V_0^D(x, \theta) - V_0^S(x, \theta) = -\theta(1 - \beta) + \int U(y - x/4)dH(y) - \beta \int U(y - x)dH(y)$ , which is continuous and strictly decreasing in  $\theta \in [0, \infty)$ . We have  $V_0^D(x, 0) - V_0^S(x, 0) = \int U(y - x/4)dH(y) - \beta \int U(y - x)dH(y) > 0$ . By continuity and strict monotonicity with respect to  $\theta$ , there exists  $\theta_0(x) > 0$  such that  $V_0^D(x, \theta_0(x)) - V_0^S(x, \theta_0(x)) = 0$ . For any  $\theta$  below this cut-off, dropping out is strictly preferred to shirking and at or above this cut-off, shirking is weakly or strictly preferred to dropping out.  $\square$

Proposition 3.2 shows that conditional on not putting in effort in college, some students would rather spend time in college shirking than dropping out so as to delay paying the cost  $\theta$ . Students who choose to do this are using the student loan program to borrow and consume leisure.

The next proposition deals with the decision to put in effort in college in period 0.

**Proposition 3.3.** *In period 0, there exists a cut-off  $\gamma_0(x, \pi, \theta) \geq 0$  such that for  $\gamma < \gamma_0(x, \pi, \theta)$  (if applicable), students put in effort in period 0 and for  $\gamma \geq \gamma_0(x, \pi, \theta)$  they either drop out or shirk. Furthermore,  $\gamma_0(x, \pi, \theta)$  is increasing in  $\pi$  and  $\theta$ .*

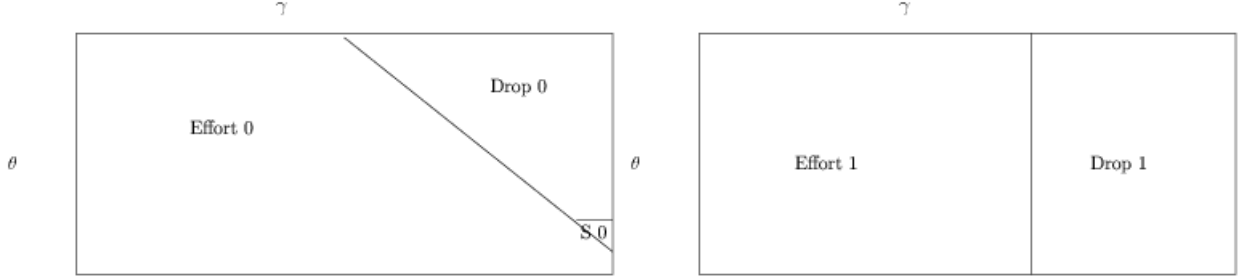
*Proof.* Consider the function  $V_0(x, \pi, \gamma, \theta) = V_0^E(x, \pi, \gamma) - \max[V_0^D(x, \theta), V_0^S(x, \theta)]$  which is continuous for all  $(\pi, \gamma, \theta) \in [0, 1] \times [0, \infty) \times [0, \infty)$  and strictly increasing in  $\pi$  (by Assumption 2), strictly decreasing in  $\gamma$  and strictly increasing in  $\theta$ . If  $V_0(x, \pi, 0, \theta) \leq 0$ , then  $\gamma_0(x, \pi, \theta) = 0$ . If  $V_0(x, \pi, 0, \theta) > 0$ , by continuity and strict monotonicity with respect to  $\gamma$  there exists a unique  $\hat{\gamma} > 0$  such that  $V_0(x, \pi, \hat{\gamma}, \theta) = 0$ . Thus,  $\gamma_0(x, \pi, \theta) > 0$ .

The fact that  $\gamma_0(x, \pi, \theta)$  is increasing in  $\pi$  can be established exactly along the lines of the proof given in Proposition 3.1.

To prove  $\gamma_0(x, \pi, \theta)$  is increasing in  $\theta$ , consider  $\tilde{\theta} < \hat{\theta}$ . If  $V_0(x, \pi, \gamma, \tilde{\theta}) < V_0(x, \pi, \gamma, \hat{\theta}) \leq 0$ , then  $\gamma_0(x, \pi, \tilde{\theta}) = \gamma_0(x, \pi, \hat{\theta}) = 0$ . If  $V_0(x, \pi, \gamma, \tilde{\theta}) \leq 0 < V_0(x, \pi, \gamma, \hat{\theta})$ , then  $0 = \gamma_0(x, \pi, \tilde{\theta}) < \gamma_0(x, \pi, \hat{\theta})$ . Finally, if  $0 < V_0(x, \pi, \gamma, \tilde{\theta}) < V_0(x, \pi, \gamma, \hat{\theta})$  then  $0 < \gamma_0(x, \pi, \tilde{\theta}) < \gamma_0(x, \pi, \hat{\theta})$ . This establishes that  $\gamma_0(x, \pi, \theta)$  is increasing in  $\theta$ .  $\square$

These Propositions can be conveniently seen in Figures 2. The left (right) figure presents the choices that the student makes in period 0 (period 1) in terms of the effort levels required on the job,  $\theta$ , and the effort level required in college,  $\gamma$ .

Figure 2: Choices in periods 0 and 1



The propositions 3.1 and 3.3 give us two thresholds for  $\gamma$ . It is important to understand the relationship between them because it will play an important role in the discussion of optimal insurance. We have the following proposition.

**Proposition 3.4.** *Assume that  $\pi > \bar{\pi}_1$ . For sufficiently low value of  $\theta$ ,  $\gamma_0(x, \pi, \theta) < \gamma_1(x, \pi)$  and for sufficiently high value of  $\theta$ ,  $\gamma_0(x, \pi, \theta) > \gamma_1(x, \pi)$ .*

*Proof.* We will evaluate  $V_0(x, \pi, \gamma, \theta)$  at the value of  $\gamma$  for which the student is indifferent between putting in effort or dropping out in period 1.

For  $\pi > \bar{\pi}_1$  and  $\theta < \theta_0(x)$ ,  $V_0(x, \pi, \gamma_1(x, \pi), \theta) = -\gamma_1(\pi, x) + \theta - \beta[\int U(y - x/4)dH(y) - \int U(y - x)dH(y)] - \beta\pi[\int U(y - x)dH(y) - \int U(y - 5x/4)dH(y)]$ . This implies that for  $\theta$  sufficiently close to 0,  $V_0(x, \pi, \gamma_1(x, \pi), \theta) < 0$ . Hence, for  $\theta$  sufficiently small,  $\gamma_0(x, \pi, \theta) < \gamma_1(x, \pi)$ .

For  $\pi > \bar{\pi}_1$  and  $\theta > \theta_0(x)$ ,  $V_0(x, \pi, \gamma_1(x, \pi), \theta) = -\gamma_1(\pi, x) + \beta\theta - \beta\pi[\int U(y - x)dH(y) - \int U(y - 5x/4)dH(y)]$ . This implies that for  $\theta$  is sufficiently large  $V_0(x, \pi, \gamma_1(x, \pi), \theta) > 0$ . Hence for  $\theta$  sufficiently large,  $\gamma_0(x, \pi, \theta) > \gamma_1(x, \pi)$ .  $\square$

The significance of these results is that for a student with  $\gamma < \gamma_0(x, \pi, \theta) < \gamma_1(x, \pi)$  it is optimal to put in effort in period 0, and if she is successfully completes college in period 0, to also put in effort in period 1. In contrast, for a student with  $\gamma_1(x, \pi) < \gamma < \gamma_0(x, \pi, \theta)$ , it is optimal to put in effort in the first year of college but then drop out even if he or she is

successful. This is a student for whom the cost of effort is high enough that exerting effort throughout both years of college is not optimal but it is low enough (and disutility from the low-paid job high enough) that it is optimal to exert effort in the first year of college and thereby avoid  $\theta$ .

Next we will determine who enrolls in college. Observe that since enrolling in college and then leaving gives people about the same utility as working, there is a small cost to a student to enroll in college and learn her  $\gamma$ . However, if the student's probability of success is sufficiently low, she may choose not to enroll because regardless of the value of  $\gamma$  she will find it optimal to leave rather than continue with college. Similarly, for a student of a given probability of success, if the effort in the low paid job is sufficiently high, she may choose to enroll.

The following proposition gives the cut-off value of effort required on the job that makes the student indifferent between working and enrolling in college. For every effort less than that, the student strictly prefers not to enroll.

**Proposition 3.5.** *In period 0, there exists a cut-off  $\theta_C(x, \pi) \geq 0$  such that for  $\theta > \theta_C(x, \pi)$  enrolling gives at least as much utility as working and  $\theta \leq \theta_C(x, \pi)$  working gives at least as much utility as enrolling. Furthermore,  $\theta_C(x, \pi)$  is decreasing in  $\pi$ .*

*Proof.* Consider the function  $V_C(x, \pi, \theta) = \int \max\{V_0^E(x, \pi, \gamma), V_0^D(x, \theta), V_0^S(x, \theta)\}dG(\gamma) - W(\theta)$ . We will show that this function is increasing in  $\theta$ . Observe that

$$V_C(x, \pi, \theta) = \int_0^{\gamma_0(x, \pi, \theta)} V_0^E(x, \pi, \gamma)dG(\gamma) + \int_{\gamma_0(x, \pi, \theta)} \max[V_0^D(x, \theta), V_0^S(x, \theta)]dG(\gamma) - W(\theta).$$

Let  $\theta$  increase by  $\Delta > 0$ . Consider the effect of this change on  $V_C(x, \pi, \theta)$  in 2 parts:

$$V_C(x, \pi, \theta + \Delta) - V_C(x, \pi, \theta) = [V_C(x, \pi, \theta + \Delta) - \bar{V}_C(x, \pi, \theta + \Delta)] + [\bar{V}_C(x, \pi, \theta + \Delta) - V_C(x, \pi, \theta)].$$

where

$$\bar{V}_C(x, \pi, \theta + \Delta) = \int_0^{\gamma_0(x, \pi, \theta)} V_0^E(x, \pi, \gamma) dG(\gamma) + \int_{\gamma_0(x, \pi, \theta)} \max[V_0^D(x, \theta + \Delta), V_0^S(x, \theta + \Delta)] dG(\gamma) - W(\theta + \Delta).$$

Then  $[\bar{V}_C(x, \pi, \theta + \Delta) - [V_C(x, \pi, \theta)]]$  is given by

$$\begin{aligned} & \int_{\gamma_0(x, \pi, \theta)} \max\{-(\theta + \Delta) + \int u(y - x/4)H(dy), -(\theta + \Delta)\beta + \beta \int u(y - x)H(dy)\} dG(\gamma) - \\ & \int_{\gamma_0(x, \pi, \theta)} \max\{-\theta + \int u(y - x/4)H(dy), -\theta\beta + \beta \int u(y - x)H(dy)\} dG(\gamma) + \Delta \end{aligned}$$

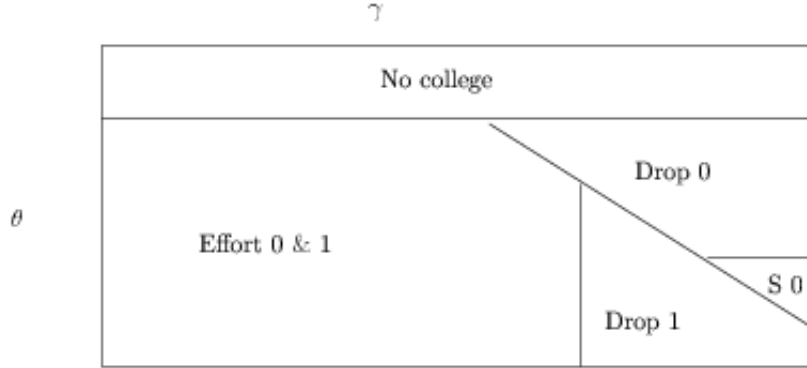
Observe that the above change is non-negative because the positive  $\Delta$  term contributes  $\Delta$  while the negative  $\Delta$  term contributes either  $-\Delta G(\gamma_0(x, \pi, \theta))$  (in the case where  $\theta + \Delta < \theta_0$ ) or  $-\beta\Delta G(\gamma_0(x, \pi, \theta))$  (in the case where  $\theta + \Delta \geq \theta_0$ ). Furthermore, the term  $[V_C(x, \pi, \theta + \Delta) - \bar{V}_C(x, \pi, \theta + \Delta)]$  is non-negative by optimality. Hence,  $V_C(x, \pi, \theta + \Delta) - V_C(x, \pi, \theta) \geq 0$ . Thus  $V_C(x, \pi, \theta)$  is increasing in  $\theta$ .

Since  $V_C(x, \pi, \theta)$  is increasing in  $\theta$ , if enrolling is optimal for some  $\theta$ , enrolling must be also be optimal for any  $\hat{\theta}$  greater than  $\theta$ . Therefore, there must be a cut-off value  $\theta_C(x, \pi) \geq 0$  such that for all  $\theta > \theta_C(x, \pi)$  the student will find it optimal to enroll and for  $\theta \leq \theta_C(x, \pi)$  the student will find it optimal to not enroll.

To establish that the threshold is decreasing in  $\pi$  observe that  $V_0^E(x, \pi, \gamma)$  is strictly increasing in  $\pi$  and, therefore,  $V_C(x, \pi, \theta)$  is strictly increasing in  $\pi$ . It follows that the cut-off  $\theta_C(x, \pi)$  cannot be strictly increasing in  $\pi$ .  $\square$

Our model of college enrollment and college completion is consistent with a diversity of student behavior. First, it predicts that not every student will enroll in college. Second, among those who enroll some will leave college voluntarily or shirk in period 0. These are the students who discover that their disutility from putting in effort in college is higher than  $\gamma_0(x, \pi, \theta)$ . Third, there will be students who continue on in college (and put in effort) in

Figure 3: Choices in college



period 0, but fail to graduate with probability  $1 - \pi$ . Fourth, among students who successfully complete period 0 in college, some will leave college voluntarily in period 1. These are the students whose disutility from putting in effort in college happens to be between  $\gamma_0(x, \theta, \pi)$  and  $\gamma_1(x, \pi)$ . Fifth, there will be students who continue on in college (and put in effort) in period 1, but fail to graduate, with probability  $1 - \pi$ . Finally there are students who enroll in college and complete their degrees. Figure 3 sums up this diversity of behavior as determined by the two types of effort costs,  $\theta$  and  $\gamma$ .

Next, we turn briefly to the observable implications of the theory. Among other things, the theory implies specific patterns regarding enrollment, non-completion and earnings with respect to the probability of success  $\pi$ . If prospective students can be classified by some observable index of their probability of success in college conditional on putting in effort – by their scholastic ability – the theory makes predictions about the variation in student performance across scholastic ability groups. In what follows, we will assume that there is an observable index  $a$  that varies positively with probability of success  $\pi$ . That is,

**Assumption 4:**  $\pi(a)$  is increasing in  $a$ .

We study how the cut-offs illustrated above change with  $a$  holding all other primitives constant. The purpose is to show that the model is consistent with the basic qualitative patterns in the data regarding enrollment, non-completion and earnings across observed

ability groups. As we will document in the next section, if  $a$  is proxied by SAT scores we find that enrollment rates are increasing in  $a$ , non-completion rates are decreasing in  $a$  and earnings are increasing in  $a$ .

Proposition 3.5 delivers that  $\theta_C(x, \pi)$  is decreasing in the probability of success  $\pi$ . Since  $\pi(a)$  is increasing in  $a$  this implies that the enrollment cut off is declining in  $a$ . Hence, enrollment rates – defined as the fraction of students of a particular ability group who enroll for college – is increasing in  $a$ .

For each ability level  $a$  define the *non-completion rate*,  $n(a)$ , as the sum of the fraction of students who enroll in college but drop out, shirk or fail in period 0, or, drop out or fail in period 1. That is,

$$\begin{aligned} n(a) &= [1 - G(\gamma_0(x, \pi(a), \theta))] + [1 - \pi(a)]G(\gamma_0(x, \pi(a), \theta)) + \pi(a)G(\gamma_0(x, \pi(a), \theta)) \\ &\times \{[1 - \tilde{G}(\gamma_1(x, \pi(a), \theta))] + [1 - \pi(a)]\tilde{G}(\gamma_1(x, \pi(a)))\}, \end{aligned}$$

where  $\tilde{G}(\gamma) = \min\{1, \frac{G(\gamma)}{G(\gamma_0(x, \theta, \pi(a)))}\}$  is the distribution of  $\gamma$  conditional on  $\gamma < \gamma_0(x, \pi(a), \theta)$ .

**Proposition 3.6.** *The non-completion rate  $n(a)$  is decreasing in  $a$*

*Proof.* The expression for  $n(a)$  simplifies to  $1 - \pi(a)^2 G(\gamma_0(x, \theta, \pi(a))) \tilde{G}(\gamma_1(x, \pi(a)))$ . Substituting in the expression of  $\tilde{G}(\gamma)$  we get

$$\begin{aligned} n(a) &= 1 - \pi(a)^2 G(\gamma_0(x, \theta, \pi(a))) \min \left\{ 1, \frac{G(\gamma_1(x, \pi(a)))}{G(\gamma_0(x, \theta, \pi(a)))} \right\} \\ &= 1 - \pi(a)^2 \min \{ G(\gamma_0(x, \theta, \pi(a))), G(\gamma_1(x, \pi(a))) \} \end{aligned}$$

The result follows from Propositions 3.3 and 3.1 which established that  $\gamma_0(x, \pi, \theta)$  and  $\gamma_1(x, \pi)$  are increasing in  $\pi$  and the assumption that  $\pi(a)$  is increasing in  $a$ .  $\square$

Next we show that average earnings is increasing in scholastic ability. By average earnings

of a scholastic group  $a$  we mean

$$e(a) = F(\theta_C(x, \pi(a))) \int y dH(y) + [1 - F(\theta_C(x, \pi(a)))] [n(a) \int y dH(y) + (1 - n(a)) \int y dM(y)]$$

**Proposition 3.7.** *Average earnings  $e(a)$  is increasing in  $a$*

*Proof.* Follows from Proposition 3.5 which established that  $\theta_C(x, \pi)$  is decreasing in  $\pi$  and therefore  $\theta_C(x, \pi(a))$  is decreasing in  $a$ , the Proposition 3.6 which delivered that  $n(a)$  is decreasing in  $a$  and Assumption 3 which implies  $\int y dM(y) > \int y dH(y)$ .  $\square$

These propositions relied on the assumption that  $a$  affected  $\pi$  only. It is possible that  $a$  also affects other primitives, for instance, the distribution from which the effort cost  $\gamma$  is drawn, the distribution from which earnings  $y$  is drawn and the college cost  $2x$ . Indeed, in the quantitative section, we will permit  $a$  to affect these distributions and the college cost as well.

## 4 Insuring College Failure Risk

Can the student loan program gainfully offer insurance against college failure risk? As noted in the introduction, we wish to answer this question recognizing that the student loan program cannot redistribute resources from students with high probability of success (high ability) to students with low probability of success (low ability) and recognizing that insurance against college failure may encourage shirking (and therefore failure).

It is best to break up the answer in two parts. Consider first the nature of optimal insurance in period 1 when loan administrators can observe effort so that moral hazard is not an issue. Conditional on the student having put in effort, the student loan program gives a transfer  $f_1$  to a student if she fails college and collects a premium  $s_1$  if she completes college. Since

the insurance is required to be self-financing (no cross-subsidies) we must have

$$-\pi \cdot s_1 + (1 - \pi) \cdot f_1 = 0. \quad (1)$$

Ignoring the  $-\gamma$  term, expected utility given these transfer is then

$$\pi \cdot \int U(y - 2x - [(1 - \pi)/\pi]f_1)dM(y) + (1 - \pi) \cdot \int U(y - 2x + f_1)dH(y), \quad (2)$$

where we have used equation (1) to express  $s_1$  in terms of  $f_1$ . Maximizing the above expression with respect to  $f_1$  yields the following first-order condition:

$$\int U'(y - 2x - [(1 - \pi)/\pi]f_1)dM(y) = \int U'(y - 2x + f_1)dH(y).$$

Hence the value of  $f_1$  that attains the maximum is one that equalizes the expected marginal utility of consumption following failure and success. Denote this value of  $f_1$  by  $f_1^*$ . Because there is a college premium in earnings (meaning that the distribution  $M(y)$  first-order stochastic dominates the distribution  $H(y)$ ) the value of  $f_1^*$  will typically far exceed the cost of the college  $2x$ . Henceforth, we will proceed under the assumption that this is so.

**Assumption 5:**  $f_1^* > 2x$  (first best insurance exceeds college costs)

Since our goal is to study the possibility of offering insurance against the risk of paying for college but failing to graduate, we limit the maximum insurance that can be offered against failure to  $2x$ . The following is then true.

**Lemma 4.1.** *Given Assumption 5,  $V_1^E(x, \pi, \gamma, f_1) = -\gamma + \beta[\pi U(y - 2x - f_1\pi/(1 - \pi)) + (1 - \pi)U(y - 2x + f_1)]$  is strictly increasing in  $f_1 \in [0, 2x]$*

*Proof.* The result follows from noting that  $\partial V_1^E(x, \pi, \gamma, f_1)/\partial f_1 > 0$  for all  $f_1 \in [0, 2x]$ .  $\square$

When effort is not observable, however, actuarially fair insurance up to the full cost of college cannot generally be offered. Under full-cost insurance, a student who shirks receives

$\beta \int U(y)dH(y)$ . In contrast, the student gets  $\int U(y - 5x/4)dH(y)$  from dropping out. For  $\beta$  close to 1, shirking will dominate dropping out. In fact, we will proceed under the assumption that it does.

**Assumption 6:**  $\beta \int U(y)dH(y) > \int U(y - 5x/4)dH(y)$  (full cost insurance induces shirking)

Thus, with full-cost insurance, students who chose to drop out prior to the introduction of insurance (and by Proposition 3.4 such students do exist) now *may* be motivated to shirk instead. If at least some students shirk, the failure rate will exceed  $\pi$  and the premia collected will fail to cover loss claims.<sup>5</sup>

We first consider optimal insurance schemes that do not induce shirking. This is a restrictive but simpler problem to analyze. It is simpler because with a “no-shirking” insurance arrangement, the probability of failure is simply  $\pi$ . In contrast, less restrictive insurance schemes may induce shirking and raise the probability of failure above  $\pi$  since shirkers fail with probability 1. The endogeneity of the failure probability makes the general insurance problem difficult (this is the moral hazard problem). The solution to the restrictive “no-shirking” insurance problem provides some guidance on how to set up the general optimal insurance problem.

We will denote the indemnity in period  $t$  (i.e., the payment received in the event of failure in period  $t$ ) as  $f_t$  and the payment in case of success as  $s_t$ . We will assume that students who succeed pay their premia when they leave college. Assuming that program administrators cannot tell the difference between genuine failures and those who fake failure by shirking, the payoffs in period 1 are as follows:

1. A student who drops out gets

$$V_1^D(x, s_0) = \int U(y - 5x/4 - s_0)dH(y).$$

---

<sup>5</sup>It is not certain that these students will find it optimal to shirk. The reason is that insurance also increases the value of putting in effort in college.

2. A student who continues but shirks gets

$$V_1^S(x, f_1, s_0) = \beta \int U(y_N - 2x - s_0 + f_1) dH(y).$$

3. A student who continues and puts in effort gets

$$V_1^E(x, \pi, \gamma, f_1, s_0, s_1) = -\gamma + \beta[\pi \int U(y - 2x - s_0 - s_1) dM(y) + (1 - \pi) \int U(y - 2x - s_0 + f_1) dH(y)].$$

And, the payoffs in period 0 are as follows:

1. Individuals who do not enroll get

$$W(\theta) = -\theta + \beta \int U(y) dH(y).$$

2. Students who enroll but leave get

$$V_0^D(x, \theta) = -\theta + \int U(y - x/4) dH(y).$$

3. Students who enroll, do not leave and shirk get

$$V_0^S(x, \theta, f_0) = -\beta\theta + \beta \int U(y - x + f_0) dH(y).$$

4. Students who enroll, do not leave and put in the effort get

$$V_0^E(\pi, x, \gamma, f_0, f_1, s_0, s_1) = -\gamma + \beta[\pi \max[V_1^E(\pi, x, \gamma, s_0, s_1, f_1), V_1^S(x, s_0, f_1), V_1^D(x, s_0)] + (1 - \pi) \int U(y - x + f_0) dH(y)].$$

Define the welfare of a student with utility costs  $(\theta, \gamma)$  as

$$W(\pi, x, \theta, \gamma, f_0, f_1, s_0, s_1) = \max\{V_0^E(\pi, x, \gamma, f_0, s_0, f_1, s_1), V_0^S(x, \theta, f_0), V^D(x, \theta)\}$$

The optimal insurance problem with the no-shirking constraint is:

$$\sup_{\{f_0, f_1, s_0, s_1\}} \int_{\theta} \left[ \int_{\gamma} \max\{W(x, \pi, \gamma, \theta, s_0, f_0, s_1, f_1), W(\theta)\} dG(\gamma) \right] dF(\theta) \quad (3)$$

subject to:

$$V_0^D(x, \theta) - V_0^S(x, f_0) > 0 \text{ for all } \theta$$

$$V_1^D(x, s_0) - V_1^S(x, s_0, f_1) > 0$$

$$s_0\pi - f_0(1 - \pi) = 0$$

$$s_1\pi - f_1(1 - \pi) = 0$$

The no-shirking constraints put upper bounds on the level of insurance that can be offered in periods 0 and 1.

**Proposition 4.2.** *In an optimal no-shirking insurance arrangement  $f_0$  must be 0 and  $f_1$  must be strictly less than some level  $\bar{f}_1 > 0$ .*

*Proof.* Consider the incentive constraint in period 0. This constraint requires that

$$-\theta(1 - \beta)/\beta + \left[ \int U(y - x/4)dH(y) - \beta \int U(y - x + f_0)dH(y) \right] > 0$$

Since  $\int U(y - x/4)dH(y) - \beta \int U(y - x)dH(y) > 0$ , for any  $f_0 > 0$ , there exists a  $\theta(f_0)$  such that constraint holds exactly. Since the distribution  $F(\theta)$  has unbounded support the constraint is violated for all  $\theta \geq \theta(f_0)$ . Thus the optimal “no-shirking”  $f_0$  must be 0. By the feasibility constraint, the optimal “no-shirking”  $s_0$  must also be 0.

Since  $\int U(y - 5x/4)dH(y) - \beta \int (y - 2x)dH(y) > 0$ , there exists  $\bar{f}_1 > 0$  such that  $\int U(y - 5x/4)dH(y) - \beta \int (y - 2x + \bar{f}_1)dH(y) = 0$ . For  $f_1 \geq \bar{f}_1$ , the period 1 no-shirking constraint is violated. Thus, the optimal “no-shirking”  $f_1$  must be less than  $\bar{f}_1$ .  $\square$

**Proposition 4.3.** *The supremum of the no-shirking insurance program exists and feasible  $f_1$  exist that come arbitrarily close to attaining the supremum.*

*Proof.* Since payoffs are bounded above by the quantity  $\int U(y)dM(y)$  (the expected utility of a person with a college degree and no debt), ex-ante utility, namely,

$$\int_{\theta} \left[ \int_{\gamma} \max\{W(x, \pi, \gamma, \theta, 0, 0, \pi/(1 - \pi)f_1, f_1), W(\theta)\} dG(\gamma) \right] dF(\theta)$$

is bounded above by the same quantity for every feasible choice of  $f_1$ . Thus the set of attainable ex-ante utility must have a least upper bound.

From Assumption 6 we have that  $\bar{f}_1 < 2x$ . By Lemma 4.1 we have  $V_1^E(x, \pi, \gamma, 0, 0, \pi/(1 - \pi)f_1, f_1)$  is strictly increasing in  $f_1 \in [0, \bar{f}_1)$ . Thus, ex-ante utility is strictly increasing in  $f_1 \in [0, \bar{f}_1)$ . It follows that the supremum is not attained by any feasible  $f_1$  but  $f_1$  exist that come arbitrarily close to attaining it.  $\square$

We now turn to the general insurance problem wherein we allow for insurance levels that induce shirking. The failure rate will now exceed  $1 - \pi$  because shirkers fail with probability 1. Students who succeed must pay a higher premium to cover the losses imposed by shirkers. This raises two issues. First, the increase in the cost of insurance might induce more students to shirk and a positive feedback between higher insurance costs and the measure of shirkers might make it impossible to offer such insurance. Second, even if such insurance levels are feasible they may be too costly in terms of the “tax” on the successful students and worse than “no-shirking” insurance.

We will now permit  $f = (f_0, f_1)$  to be any element of the the set  $[0, x] \times [0, 2x]$ . It is helpful to think of the premia  $s = (s_0, s_1)$  as being made up of two parts. One part is the “base”

premiums that cover losses when there is no shirking and is given by  $b(f) = (b_0(f_0), b_1(f_1)) = (\pi/(1-\pi)f_0, \pi/(1-\pi)f_1)$ . The other part is the additional premium that needs to be collected to cover the losses imposed by shirkers. Denote these as  $\tau(f) = (\tau_0(f), \tau_1(f))$ .

Define  $\gamma_0(x, \pi, \theta, f, b(f) + \tau(f)) \geq 0$  as the cut-off value of  $\gamma$  above which an enrolled student will not put in effort in college in period 0 (i.e., she will either dropout or shirk). This cut-off solves

$$V_0^E(x, \pi, \gamma, f, b(f) + \tau(f)) = \max\{V_0^D(x, \theta), V_0^S(x, \theta, f_0)\} \quad (4)$$

The existence of this cut-off follows from the same logic as in Proposition 3.3.

Define  $\theta(x, f_0)$  as the cut-off value of  $\theta$  above which, conditional on not putting in effort in college, a student would prefer to shirk and below which she would prefer to drop out. This cut-off solves

$$V_0^S(x, \theta, f_0) = V_0^D(x, \theta) \quad (5)$$

Existence follows from the same logic as in Proposition 3.2.

Finally, define  $\gamma_1(x, \pi, f_1, b(f) + \tau(f))$  as the cut-off value of  $\gamma$  above which the student does not put effort in college in period 1. This cut-off solves

$$V_1^E(x, \pi, \gamma, f_1, b(f) + \tau(f)) = V_1^S(x, f_1, b_0(f_0) + \tau_0(f_0))\chi_{\{f_1 \geq \bar{f}_1(f_0)\}} + V_1^D(x, b_0(f_0) + \tau_0(f_0))[1 - \chi_{\{f_1 \geq \bar{f}_1(f_0)\}}] \quad (6)$$

where  $\chi_{\{f_1 \geq \bar{f}_1(f_0)\}}$  is an indicator function that takes on the value 1 if the expression in  $\{\cdot\}$  is true and  $\bar{f}_1(f_0)$  is such that  $\int U(y - 5x/4 - b_0(f_0) - \tau_0(f_0))dH(y) - \beta \int (y - 2x - b_0(f_0) - \tau_0(f_0) + \bar{f}_1) dH(y) = 0$ . We have incorporated the fact that if  $f_1$  is at least as large as  $\bar{f}_1(f_0)$ , the student finds it optimal to shirk. Given an outside option (dropping out or shirking), existence follows from the same logic as in Proposition 3.1.

We can state the requirement for feasibility of  $f$ .

**Definition 4.4.** *Insurance levels  $f \in [0, x] \times [0, 2x]$  is feasible if there exist  $\tau^* = (\tau_0^*(f), \tau_1^*(f))$  such that*

$$\begin{aligned} & \pi \cdot G(\gamma_0(x, \pi, \theta, f, b(f) + \tau^*(f))) \cdot \tau_0^*(f) \\ &= [1 - G(\gamma_0(x, \pi, \theta, f, b(f) + \tau^*(f)))] \cdot [1 - F(\theta(x, f_0))] \cdot f_0 \end{aligned} \quad (7)$$

and

$$\begin{aligned} & \pi^2 \cdot G(\gamma_0(x, \pi, \theta, f, b(f) + \tau^*(f))) \cdot \tilde{G}(\gamma_1(x, \pi, f_1, b(f) + \tau(f))) \cdot \tau_1^*(f) \\ &= [1 - \tilde{G}(\gamma_1(x, \pi, f_1, b(f) + \tau(f)))] \cdot \chi_{\{f_1 \geq \bar{f}_1\}} \cdot f_1 \end{aligned} \quad (8)$$

where  $\tilde{G}(\gamma) = \min\{1, G(\gamma)/G(\gamma_0(x, \pi, \theta, f, b(f) + \tau^*(f)))\}$ .

The term multiplying  $\tau_0^*(f)$  on the lhs of (10) is the measure of enrolled students who put in effort in period 0 and succeed. Each of them pay the additional premia  $\tau_0^*(f)$ . The term on the rhs of (10) is the measure of enrolled student who do not put in effort in college *and* shirk. Each of them collect  $f_0$  from the insurance scheme. For feasibility, the two sides must balance. Similarly, the term multiplying  $\tau_1^*(f)$  on the lhs of (11) is the measure of students who put in effort in period 1 and succeed (as before  $\tilde{G}$  is the distribution of  $\gamma$  conditional on the set of  $\gamma$  for which students put in effort in period 0). Each of them pay the additional premium  $\tau_1^*(f)$ . The term on the rhs of (11) is the measure of students who do not put in effort in period 1. If the insurance scheme offers  $f_1 \geq \bar{f}_1$  then all these students shirk; otherwise they drop out. For feasibility the two sides must balance.

Let  $\Phi \subset [0, x] \times [0, 2x]$  be the set of  $f$  which are feasible.  $\Phi$  is non-empty because any insurance scheme in which  $f_0 = 0$  and  $f_1 < \bar{f}_1$ ,  $\tau = (0, 0)$  satisfies both equations (these are the set of no-shirking insurance levels). The general optimal insurance problem can be

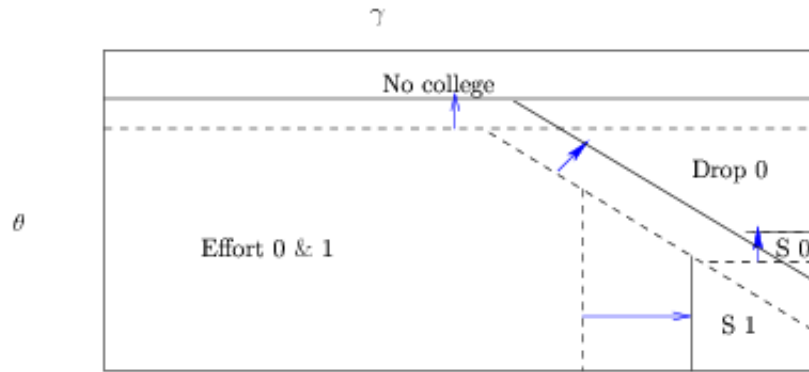
stated compactly as follows:

$$\sup_{f \in \Phi} \int_{\theta} \left[ \int_{\gamma} \max\{W(x, \pi, \gamma, \theta, f, b(f)) + \tau(f), W(\theta)\} dG(\gamma) \right] dF(\theta).$$

The fact that  $\Phi$  is non-empty and that all payoffs are bounded above by  $\int U(y) dM(y)$  implies that the supremum must exist. If no  $f$  attains the supremum, insurance levels exist that come arbitrarily close to attaining it.

Figure 4 indicates the effects optimal insurance.

**Figure 4: Choices when insurance is provided**



Optimal insurance increases the value of going to college and, thus, shifts up the the  $\gamma_0$  and  $\gamma_1$  loci. Optimal insurance also, potentially, shifts down the  $\theta_0$  locus (if optimal  $f_0 > 0$ ). Both effects work to lower dropout rates in period 0. Dropout rates also decline in period 1 because  $\gamma_1$  shifts up and all those who do not put in effort either continue to drop out or shirk (the latter happens if optimal  $f_1 \geq \bar{f}_1$ ). The effect of optimal insurance on non-completion rate is ambiguous because it encourages some students who were dropping out put in effort (this is the positive effect) others who were dropping out to shirk (the negative effect). Of course, optimal insurance raises the enrollement rate.

It is an open question whether optimal insurance should tolerate some amount of shirking.

Providing insurance beyond the “no-shirking” level will encourage more enrollment and more effort in college but it will also cause some students to shirk and thereby increase the cost of providing the insurance.

## 5 Calibration

First we present the values of the parameters in the model and then we present the estimation procedure for the probability of completion,  $\pi(a)$  and the distributions of the effort levels,  $F(\theta)$  and  $G(\gamma|a)$ .

### 5.1 Parameter Values

The utility function is CRRA with coefficient  $\sigma$ . We calibrate the benchmark model to the U.S. economy. We calibrate the distribution of ability,  $A(a)$  to the SAT scores of all students in a senior year in 1999 in the U.S. The mean is 1016 and standard deviation 226 (College Board 2007). In our model, this translates to 0.51 for the mean and 0.18 for the standard deviation. Parameters of the model are given in Table 1.

**Table 1: Parameter values**

Parameter	Name	Value	Target/Source
$\sigma$	Coef of risk aversion	2	standard
$\beta$	Discount factor	0.97	avg rate of 3%
$H(y_N)$	Noncollege earnings distribution	1.07(0.5)	earnings HS graduates
$M(y_C)$	College earnings distribution	1.69(0.8)	earnings college graduates
$A(a)$	Ability distribution	0.51(0.18)	SAT scores - seniors

We assume that the distributions of earnings for college and non-college are normal with known mean and variance. Statistics for lifetime earnings are based on earnings data from the CPS for 1969-2002 with synthetic cohorts.<sup>6</sup> We divide education based on years of

<sup>6</sup>For each year in the CPS, we use earnings of heads of households age 25 in 1969, age 26 in 1970, and so on until age 58 in 2002. We consider a five-year bin to allow for more observations, i.e., by age 25 at 1969, we mean high school graduates in the sample that are 23 to 27 years old. Real values are calculated using the CPI 1982-1984. There are an average of 5000 observations in each year’s sample.

education completed with 12 years and more, but less than 16 years of completed schooling for the no college group, and those with 16 years of completed schooling for college graduates. The present value of life-cycle earnings for high-school graduates who do not have a college degree is 1.07 million dollars (1999 constant dollars) and for college graduates is 1.69 million dollars. The low outcome is normalized to 1.07. Mean of earnings for high school graduates and college graduates are given by  $\mu^N(y) = 1.07$  and  $\mu^C(y) = 1.69$ . In our model, the ratio between the two education groups represents the earnings premium for an individual who completes college, relative to the same individual in the counterfactual situation of not attending college. This cannot be directly mapped into the rates of return for college education obtained from standard Mincer regressions, since these rates do not typically correct for the ability bias implied by the selection of individuals between college and non-college choices. Willis and Rosen (1979), adjusting for selection in ability, find a lifetime rate of return of attending college of 9.9% per year. Later research estimates the rate of return on college to be between 8% and 13% per year (see Willis (1986) and Card (2001)). This high return to college education also implies very large differences in lifetime income. We use a lifetime college premium of 1.58 (as in Murphy and Welch (1992)), which is consistent with the estimates of 13% rate of return per year.<sup>7</sup> We set the standard deviations to  $\sigma^N(y) = 0.5$  and  $\sigma^C(y) = 0.8$ . To obtain these numbers we compute the coefficient of variation in each year's sample, average them out and report this average to the present value of life cycle earnings to pin down standard deviations for the two education groups.

In the theoretical part of the paper, we showed that high ability students have higher earnings on average, given the higher chances of completing college. In addition to the effect of the probability of success from college on earnings, we consider that the distribution of life-cycle earnings of each education group also depends on the ability level, i.e.  $M(y|a)$  and  $H(y|a)$ . This exercise is motivated by the fact that conditional on completing college, the ability levels of students may have significant effects on life-cycle earnings (see

---

<sup>7</sup>Restuccia and Urrutia (2004) use the 10% rate of return that corresponds to a lifetime college premium of 1.5.

Cunha and Heckman (2009), Hendricks and Schoellman (2009)).

We use the High-School and Beyond data set for high-school graduates who acquired and did not acquire a college degree and are full time workers employed 5 years later after they acquired their highest degree. We use their annual earnings 5 years after acquiring their highest degree across SAT scores. The set of ability,  $a$  (preparation for college) has 5 elements,  $a_i$  with  $i = 1, 2, \dots, 5$  corresponding to the following 5 groups of SAT scores:  $< 700$ ,  $700 - 900$ ,  $901 - 1100$ ,  $1101 - 1250$ ,  $> 1250$ . We use these differences in earnings across the five ability groups together with the moments in the life-cycle distribution from the CPS data to estimate the means of life-cycle earnings across ability groups for the two education levels. As expected, earnings increase in ability. We obtain  $\mu_i^C(y) = 1.54, 1.66, 1.74, 1.84, \text{ and } 1.92$  and  $\mu_i^N(y) = 0.99, 1.05, 1.11, 1.17, \text{ and } 1.21$ .<sup>8</sup>

The loan amount that the government collects is computed as the fixed payment for a console with the discounted present value that equals the weighted average college cost for private and public institutions. The cost for college is \$20706/year for private universities and \$8275/year for public universities in 1999. Among the students who borrowed for their education, 67% went to public and 33% to private universities. The enrollment-weighted average cost is \$49508 in constant 1999 dollars (College Board (2001)). We consider a heterogeneous cost of college. Using this enrollment-weighted procedure, we estimate college costs across ability groups using data from the Princeton Review on college rankings in terms of the average SAT score of accepted students. We estimate the following college costs for the 5 groups of ability levels in the model: \$32545, \$36812, \$41016, \$59165, and \$77269 (in 1999 dollars). We obtain an increasing cost of college, which reflects the fact that high ability students enroll to better colleges. Thus, we set the loan amounts in the model to  $2x(a_i) = 0.0256, 0.0352, 0.0370, 0.0564, \text{ and } 0.0734$ , which represent 1.66% to 3.8% of life-cycle earnings of college graduates.

---

<sup>8</sup>We also look at earnings 1 year, 3 years and 10 years after college graduation occurs using the B&B data for college graduates in 1992/1993 and at the earnings in 2001 from the BPS data for students who enroll in college in 1996 and complete college by 2001 and also for those who do not complete college.

This monotonicity of the college cost coupled with the fact that the distribution of earnings depends on the ability level  $a_i$  captures the idea that high ability students go to better colleges and also experience higher returns to their investment. In addition, there are two more factors that affect both future and current returns from college in our model that will induce higher earnings for high ability students, apart from the dependent draw of earnings: 1) High ability students have better chances to the high draw of earnings (thus higher future benefits), as captured by  $\pi(a_i)$  conditional on putting in effort in school,  $e = 1$  and 2) Our estimates of the conditional distribution of effort in college,  $G_i(\gamma|a)$  imply higher current benefits for high ability students. This stems from avoiding the effort on the job today (which is independently drawn) at the expense of the effort in college; our estimates, which are presented in the following section show that high ability students are required a relatively lower level of effort in college, on average, than low ability students.

## 5.2 Completion Probability and Distribution of Effort Levels

In order to estimate the function  $\pi(a_i)$  and the distributions of effort levels,  $F(\theta)$  and  $G_i(\gamma)$ , we use the National Education Longitudinal Study (NELS:88) that follows eighth-graders in 1988 and the Beginning Postsecondary Students Longitudinal Study (BPS) that follows students who enroll in a post-secondary institution for the first time in 1996. The first data set (NELS:88) collects information about post-secondary education for senior students in 1992. We use this data set for enrollment rates by ability groups as measured by SAT scores. We restrict the sample to high school graduates in 1992 who have enrolled in college without delay after high-school graduation and who are fully enrolled in 2-year and 4-year colleges in October 1992. The second data set (BPS) is designed specifically to collect data related to persistence in college and completion of post-secondary education programs for students who enroll in college in 1996. These same students are surveyed 2- and 5-years later through BPS. We use this data set for early leaving rates and completion rates by ability groups as measured by SAT scores for students beginning at 2-year and 4-year institutions. As before, we restrict the sample to students who have enrolled in college without a delay

after high-school graduation and are fully enrolled. For early leaving rates we consider all students who were last enrolled in college in 1998 and have not returned to (any) college to earn a college degree. For completion rates, we consider all students who obtained a college degree by 2001. Data are presented in Table 2.

**Table 2: Enrollment, leaving and completion rates: data**

SAT scores	$\leq 700$	701 – 900	901 – 1100	1101 – 1250	$\geq 1251$
Enrollment rates	71.5	81.2	88.8	95	95.6
Early leaving rates	16.7	9.6	4.7	1.6	1.4
Completion rates	37.2	62.7	69.6	78.2	86.5

We first identify the probability of success from college  $\pi(a_i)$ . For each ability group we know the fraction that completes college and the fraction of early leavers. Thus we use the last two rows of Table 2 and compute the fraction of students who complete college out of all students who decide to continue in college. These rates directly pin down  $\pi(a_i)$ . The probabilities of success from college across ability groups are given by 44.66, 69.36, 73.03, 79.47 and 87.73. Here we are assuming that if the student has not graduated after 5 years, the student will not graduate at all.

Second, we identify the distributions  $F(\theta)$  and  $G(\gamma|a_i)$ . We calibrate these distributions to match the rates across ability levels described in Table 2. We proceed as follows: 1) We compute the enrollment rates, early leaving rates, and completion rates in the model for an initial grid of effort levels. 2) We choose the distributions of the effort levels to best replicate the properties of U.S. data documented in Table 2. Using a parametric approach, we search over the vector of parameters that characterize the two distributions to minimize the distance between the model and the data moments: enrollment, early leaving, and completion rates. 3) For any trial of the vector describing the initial distribution, we calculate these three rates by ability groups. If the histogram that best matches the data puts strictly positive weight on the right hand side of the grids of effort levels, then the upper bounds are increased and steps 1-3 are repeated.

In order to carry out the procedure in step 1, we compute the enrollment rates, early leaving rates and completion rates in the model as follows: According to our theory, for each ability group,  $a_i$ , there is a cutoff of effort on the job,  $\theta_i^c$  that will make people within that group indifferent between enrolling and not enrolling (and ignoring application costs). Then, for each ability group we have that the student enrolls in college when  $\theta \geq \theta_i^c$ , i.e. when

$$\begin{aligned}
V_0^E(a_i, E(\gamma)) - W(\theta) = & -E(\gamma)(1 + \beta\pi(a_i)) + \theta + \pi^2(a_i)\beta^2 \left[ \int u(y - 2x(a_i))dM - \right. \\
& - \int u(y - 2x(a_i))dH \left. \right] + \pi(a_i)\beta \left[ \beta \int u(y - 2x(a_i))dH - \int u(y - x(a_i))dH \right] \\
& + \beta \int u(y - x(a_i)) - u(y)dH \geq 0.
\end{aligned} \tag{9}$$

Note that the enrollment choice takes into account the expected value of effort level in college,  $E(\gamma)$ . We use Equation 9 to compute enrollment rates by ability groups in the model. Furthermore, according to our theory, a student leaves college if  $V_0^D(\pi, x, \theta) > \max\{V_0^E(\pi, x, \gamma), V_0^S(\pi, x, \theta)\}$  or  $V_1^D(\pi, x) > \max\{V_1^E(\pi, x, \gamma), V_1^S(\pi, x)\}$ . We use this condition to compute the early leaving rates by ability groups. Finally, we compute the completion rate in the model, recognizing that some people may choose to shirk in college, and thus the probability of completing college for them is 0, whereas for people who decide to continue and put in effort, the probability of completion is  $\pi(a_i)$ .

Once we compute all three rates by ability levels for an initial guess of the distributions of effort levels in the model, we find the best distributions of effort levels (step 2 in the procedure). We restrict the initial distribution of effort during work to be normally distributed and the distribution of effort during college to be of the exponential type. We consider a general specification with different distributions of effort during college by ability levels,  $G_i(\gamma)$ . This formulation is motivated by the fact that students of different ability levels may require a different level of effort during college. The first class of distributions is characterized by 2 parameters and the second by one parameter. Thus, the problem reduces to finding the vector of parameters  $\alpha = (\mu_\theta, \sigma_\theta, \mu_{\gamma_{i=1,2,\dots,5}})$  characterizing these initial distributions to solve

the minimization problem

$$\min_{\gamma} \left( \sum_{j=1}^5 w_j ((e_j - e_j(\alpha))^2 + w_2 (l_j - l_j(\alpha))^2 + w_3 (c_j - c_j(\alpha))^2) \right),$$

where  $e_j, l_j$  and  $c_j$  represent the enrollment, early leaving and completion rates in the NELS and BPS data sets,  $e_j(\alpha)$ ,  $l_j(\alpha)$ , and  $c_j(\alpha)$  are the corresponding model rates and  $w_i$  are weights assigned to the three rates. Overall we match 15 moments.

We find the distributions  $F(\theta) \sim (0.361, 0.172)$  and  $G_1(\gamma) \sim (0.0136)$ ,  $G_2(\gamma) \sim (0.07)$ ,  $G_3(\gamma) \sim (0.0675)$ ,  $G_4(\gamma) \sim (0.0474)$ , and  $G_5(\gamma) \sim (0.0424)$ . Note that the distribution of  $\gamma$  is monotonic if we do not account for the lowest ability group. These properties of the distribution  $G_i(\gamma)$  are consistent with our interpretation of  $\gamma$  as the cost associated with school work and of  $a$  as simply the aptitude for college. In general, high ability students require less effort in college than low ability students. At the same time, for the lowest ability group, we think this is a quality of college issue – the very low ability colleges must have lower standards so that the effort required to get a degree is lower too.

To this end, our estimation procedure captures the idea that high ability people have higher chances of success provided that they put effort in college. Going against this is that the effort required at better colleges may be higher, and thus it might be harder to succeed; however, research shows that college completion is positively associated with the institution’s selectivity, tuition revenue and general expenditures per full-time equivalent student (see Hoxby (2003), Melguizo (2008), and Titus (2006)). In addition, our estimation of the distribution of efforts,  $G_i(\gamma)$  reveals that the means of effort decline in the ability levels of the student (except for the lowest ability group).

## 6 Quantitative Results

### 6.1 Benchmark

The enrollment, leaving and completion rates simulated in our benchmark economy are given in Table 3. The enrollment rate in the economy is 88.6%. Out of those who enroll, 4.8% decide to leave college and 70.1% complete college. There are 62% college graduates, on average in the benchmark economy. Out of the students who leave college, most of them decide to do so in the second part of college in the case of people with low ability. In contrast, college leavers with high ability are more likely to leave college in period 0 than in period 1. For instance, 100% and 99.5% of leavers in the first and in the second ability group leave college in period 1 and only 12% of college leavers in the fifth ability group leave college in period 1. For ability groups 3 and 4 these percentages are 91.5% and 64.9% respectively. Students with low ability levels, may find it beneficial to stay in college in period 0 to earn some credits and possibly get a 2 year degree, which could help them perform better in the job market. Students with high ability levels, however, may not need this signal in the job market. If they realize that college is not good enough for them, they walk away and eventually obtain a fairly good job with or without some college education.

**Table 3: Enrollment, leaving and completion rates: model**

SAT scores	$\leq 700$	701 – 900	901 – 1100	1101 – 1250	$\geq 1251$
Enrollment rates	71.3	80.6	88.9	95.5	98.5
Leaving rates	15.3	8.6	4.8	1.7	1.3
Completion rates	37.8	63.4	69.5	78.2	86.6

### 6.2 Insurance

We follow the analysis in Section 4. First, we consider the case where effort is observable. The model delivers that the optimal level of insurance (that equates marginal utilities across states),  $f_i^*$  is higher than the cost of college,  $2x(a_i)$  for all ability levels  $i$ . However, the loan

administrators cannot offer more insurance than the college cost. The most that they can do is to forgive the full loan amount in the case of failure. Thus, when we restrict the insurance up to the cost of college, it is optimal to offer the college cost by ability groups,  $2x(a_i)$ .

We turn now to the case where effort is not observable. Recall that our theory implies that no insurance can be offered in period 0 without inducing shirking, given unbounded support of the effort level on the job,  $\theta$ , but insurance in period 1 can be offered without inducing shirking. Thus, we first restrict attention to offering insurance in period 1 when the no-shirking contract is offered. This insurance scheme implies that no insurance is provided in period 0 and insurance in period 1 is provided up to the point where people are indifferent between shirking and dropping out. Second, we solve for the optimal insurance with shirking, with indemnities offered in both periods in college (up to the cost of college).

### 6.2.1 No-shirking Insurance: Period 1

Recall from Proposition 4.2 that it is optimal to offer insurance in period 1 in college up to  $\bar{f}_1$ , which satisfies  $\int U(y_N - 5x/4)dH(y_N) = \beta \int (y_N - 2x + \bar{f}_1)dH(y_N)$ . Given the previous discussion about  $f^*$ , it follows that the  $\min\{f^*, \bar{f}_1\} = \bar{f}_1$ . The exception is the insurance offered to the lowest ability group who draw, on average, a very low effort in school. The value of putting in effort dominates the value of shirking and thus, there is no incentive to shirk for any level of insurance up to 68.75 % of the college cost offered to this ability group. As a result, even though the indemnity that makes shirking just as good as dropping out is only 22.65 % of the college cost, it is optimal to offer 68.75% of the college cost to this ability group. The premium that is paid in case of success is  $\bar{s}_1 = (1 - \pi)\bar{f}_1/\pi$ . Table 4 presents the indemnity offered,  $\bar{f}_1(a_i)$  by ability groups, as well as the premium paid in case of success,  $\bar{s}_1(a_i)$ . The table also presents the insurance amounts as percentages of  $2x$ , the cost of college. Note that the indemnity offered increases in ability (except for the lowest ability group), with the top ability group receiving more than two times more indemnity than the second ability group. However, given that the college cost increases in the ability level, each ability group is forgiven a significant amount of their college cost in the case where

**Table 4: No-Shirking Insurance in period 1**

SAT scores	$\leq 700$	701 – 900	901 – 1100	1101 – 1250	$\geq 1251$
Indemnity $f_1$	0.0176	0.0082	0.0085	0.0136	0.0182
Perc of 2x	68.75	23.3	22.95	24.1	24.8
Premium $\bar{s}_1$	0.0218	0.0036	0.0031	0.0035	0.0026
Perc of 2x	85.16	10.25	8.5	6.25	3.48

failure occurs. The second lowest (highest) ability group is forgiven 23.3% (24.8%) of their college cost. The insurance, however is more expensive for the low ability groups relative to the high ability groups: the premium  $s_1$  is 10.25% of the college cost for the second ability group and only 3.41% of the college cost for the top ability group.

The no-shirking insurance increases enrollment and completion rates and decreases leaving rates. With insurance students find it beneficial to stay and put in effort rather than drop out voluntarily. An important observation is that when this insurance is provided, there is a positive mass of students on average (1.73%) who are indifferent between shirking and dropping out. This shirking rate comes from ability groups  $i = 2, \dots, 5$ . We next consider the case where insurance induces shirking.

### 6.2.2 Optimal Shirking Insurance

We consider the general insurance case,  $f \in [0, x] \times [0, 2x]$ , which allows people shirk. We first need to determine the feasible set of insurance schemes for each ability group. Recall that we need to collect enough in addition to the premium paid in case of success such that the mass of shirkers is covered, which means that we need to find  $\tau^* = (\tau_0^*(f), \tau_1^*(f))$  such

that the following conditions are satisfied:

$$\begin{aligned} & \pi \cdot G(\gamma_0(x, \pi, \theta, f, b(f) + \tau^*(f)) \cdot \tau_0^*(f)) \\ &= [1 - G(\gamma_0(x, \pi, \theta, f, b(f) + \tau^*(f)))] \cdot [1 - F(\theta(x, f_0))] \cdot f_0 \end{aligned} \quad (10)$$

and

$$\begin{aligned} & \pi^2 \cdot G(\gamma_0(x, \pi, \theta, f, b(f) + \tau^*(f))) \cdot \tilde{G}(\gamma_1(x, \pi, f_1, b(f) + \tau(f))) \cdot \tau_1^*(f) \\ &= [1 - \tilde{G}(\gamma_1(x, \pi, f_1, b(f) + \tau(f)))] \cdot \chi_{\{f_1 \geq \bar{f}_1\}} \cdot f_1 \end{aligned} \quad (11)$$

where  $\tilde{G}(\gamma) = \min\{1, G(\gamma)/G(\gamma_0(x, \pi, \theta, f, b(f) + \tau^*(f)))\}$ . We solve for  $(\tau_0^*(f), \tau_1^*(f))$  that must be collected for each  $f = (f_0, f_1) \in [0, x] \times [0, 2x]$ , that means to iterate simultaneously over Equations 10 and 11 until convergence is reached. As before, the most that the planner can do is to forgive some of the college cost. Once we determine the set of feasible pairs of indemnity and tax levels for each ability group, we then search over the feasible set to find the optimal contract,  $(f_0^* f_1^*, \tau_0^*, \tau_1^*)_i$  that maximizes welfare for ability group  $i$ .

The feasible set differs significantly across ability groups. For instance, the feasible indemnity levels  $f_1 \in [\bar{f}_1, 2x]$  when  $f_0 = 0$  are given by [22.65, 70], {23.3}, [22.95, 28.5], [24.1, 41.85], [24.8, 55.45] (numbers are given in % of the college cost,  $2x$ ). Note that insurance beyond  $\bar{f}_1$ , which makes shirking just as good as dropping out is not feasible for the second ability group,  $i=2$  when  $f_0 = 0$ . The jump in the mass of shirkers is too high once  $f_1$  and  $\tau_1$  are increased and Equation 11 fails to converge. The mass of shirkers for this particular group is much higher than for any of the other groups starting with  $\bar{f}_1$ , in the case where we assume that people who are indifferent between shirking and dropping out decide to shirk (4.61%). For the other ability groups insurance levels beyond  $\bar{f}_1$  are feasible. More schemes are feasible for higher ability levels ( $i = 4, 5$ ), as expected. But there are quite a few feasible levels for  $i = 3$ . For  $i = 1$ , more schemes are feasible, given that shirking occurs only for indemnity levels that exceed 68.75% of the college cost.

**Table 5: Optimal shirking insurance: period 1**

SAT scores	$\leq 700$	701 – 900	901 – 1100	1101 – 1250	$\geq 1251$
Indemnity $f_1^*$ ( $f_0 = 0$ )	0.0176	0.0082	0.0089	0.0167	0.0327
$f_1^*$ as perc of 2x	68.75	23.3	24.05	29.55	44.51
$s_1^*$ as perc of 2x	85.16	10.25	8.9	7.65	6.23
$\tau_1^*$	0	5.99e-4	3.06e-4	1.56e-4	6.12e-5
Shirking rates ( $f_0 = 0$ )	0	4.61	2.38	0.73	0.16

These results make perfect sense if we think about the ordering of the  $\gamma$  distribution. Recall that the mean of the effort level in college is declining in ability, except for the lowest ability group. The group that has the highest incentive to shirk is the second group. Thus, for this group any insurance beyond  $\bar{f}_1$  attracts a significant amount of shirkers, and so we need to collect more additional tax to cover the mass of shirkers. For the other ability groups, however, these increases are incrementally small, especially for the highest ability groups,  $i = 4, 5$ . So we get a fairly large range of feasible schemes. However, for high levels of the indemnity the insurance is too costly. An exception is the lowest ability group, who, on average, draw a very low effort in school, and thus, there is no incentive to shirk for a large range of indemnities offered since the value of putting in effort dominates the value of shirking.

For comparison purposes, we first show the results when only insurance in period 1 is considered and compare the optimal levels to those in the no-shirking contract presented before. Our results clearly illustrate the adverse selection problem. The ability group  $i = 2$  is most likely to get a high draw of effort, followed by  $i = 3$ . As a result, no insurance can be offered beyond the threshold  $\bar{f}_1$  for  $i = 2$  and slightly more for  $i = 3$ . The incentive to shirk is lower for the other ability groups given that they are less likely to draw a high effort. Thus, the optimal scheme is relatively more generous compared to the no-shirking contract. Table 5 presents the indemnity and the premium paid for the optimal contract with shirking when insurance is offered only in period 1. We also present the induced shirking rates.

To isolate the effects of offering insurance in each period in college, we also solve for the

feasible sets and the optimal levels of indemnities offered in period 0,  $f_0$  when  $f_1 = 0$ . Results deliver that shirking occurs from much lower levels of  $f_0$ . A small amount of  $f_0$  is needed for shirking to occur, given that the no-shirking constraint also depends on the effort level on the job,  $\theta$ . There are some levels of  $f_0 > 0$  such that shirking does not occur. This is because computationally we can find  $\theta_{max}$ , such that  $F(\theta_{max}) = 1$  and for some ability groups the level of  $\theta$  that makes shirking just as good as dropping out is greater than this  $\theta_{max}$ . Thus some small levels of  $f_0$  can be found without inducing shirking. The jump in the mass of shirkers is relatively small once shirking occurs and higher indemnity levels are feasible relative to offering insurance in period 1. Also, some high indemnity levels may be feasible because they will trigger less people to enroll. A high additional tax collected for a high indemnity level declines the value of acquiring a college degree to the point that enrolling in college is suboptimal. As a result, some students will decide not to enroll, which implies that overall the mass of shirkers out of all students who decide to enroll is sustainable. These high indemnity levels, however, induce a lower welfare. Thus even though more generous indemnities are feasible for  $f_0$ , the optimal level of insurance is lower relative to the optimal level of insurance offered in period 1 across all ability groups. Table 6 presents the indemnity and the premium paid for the optimal contract with shirking offered only in period 0 together with the induced shirking rates. Note that, unlike in the previous insurance scheme, for the second ability group a higher indemnity level can be offered relative to the third ability group. Most of the people who leave college for ability group 2 decide to do so at the beginning of period 1, and thus the mass of shirkers in period 0 for ability group 2 is relatively small. The opposite is true for the top ability group, and thus, the indemnity offered in period 0 to this ability group is small.

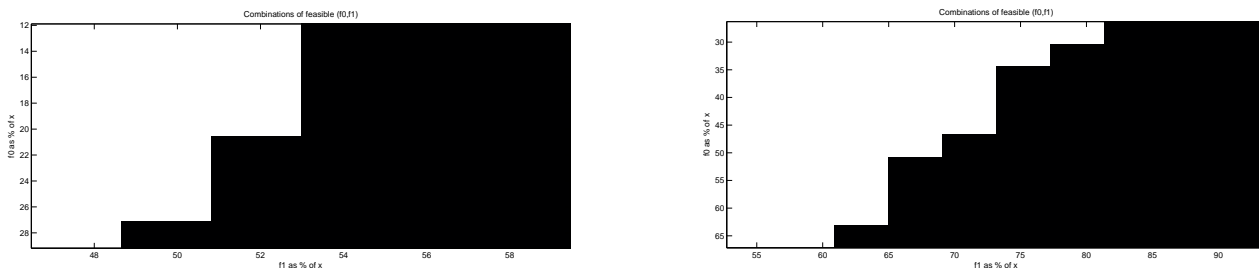
We turn now to the results regarding the feasible sets and optimal insurance levels when insurance is offered in both periods in college. The indemnity levels that are feasible in each of the two periods are less generous than in the case where insurance is offered only in one of the two periods in college. For instance, for ability level 1, recall that no shirking occurs for any indemnity offered up to the cost of college, in the case where insurance is offered

**Table 6: Optimal shirking insurance: period 0**

SAT scores	$\leq 700$	701 – 900	901 – 1100	1101 – 1250	$\geq 1251$
Indemnity $f_0^*$ ( $f_1 = 0$ )	0.0128	0.0064	0.0037	0.0065	0.0011
$f_0^*$ as perc of 2x	50	17.45	9.65	11.55	1.51
$s_0^*$ as perc of 2x	61.9	15.4	7.3	5.9	0.2
$\tau_0^*$	0	1.417e-4	7.427e-5	5.959e-5	1.554e-4
Shirking rates ( $f_1 = 0$ )	0	2.17	1.99	0.91	0.12

only in one of the two periods in college. So any level of insurance is feasible. However, if insurance is offered in both periods in college, shirking occurs for some combinations of  $(f_0, f_1)$  and insurance in period 1 is feasible only up to 45% of the college cost. Similarly, for the ability level 2, which represents the other extreme case, the only level of insurance feasible in period 1 is 23.3% of the college cost when insurance is offered only in period 1. This level of indemnity  $f_1$  is not feasible anymore in the case where insurance is offered in both periods 0 and 1. In this case, the feasible levels are up to 15% of the college cost for  $f_0$  and up to 18.5% of the college cost for  $f_1$ . The intuition behind these results is that offering insurance in period 0 induces more shirking in period 0 and thus fewer people stay in college in period 1. Consequently, the mass of shirkers out of all students who stay in college in period 1 is affected. In addition, insurance in period 0 changes the values of shirking and dropping out in period 1, given that the student in good standing in period 1 in college owes the premium  $s_0$ . At the same time, insurance in period 1 increases the value of shirking in period 1 and the mass of shirkers in period 1, but it increases the value of continuing in college in period 0. Thus it induces students to put in effort in period 0 and decreases their incentives to shirk in period 0. A direct result of these effects of offering insurance on shirking incentives is that the additional tax that must be collected in period 1 is higher, in general, relative to the tax that is collected in period 0. As a result, less generous insurance schemes are feasible for period 1 relative to the case where insurance is offered only in period 1, and the same or slightly less generous schemes are feasible for insurance offered in period 0 relative to the case where insurance is offered only in period 0. An exception is the top ability group, for whom insurance in period 1 significantly declines the incentive to shirk in period 0. This

**Figure 5: Feasible sets (in white) when insurance is provided**



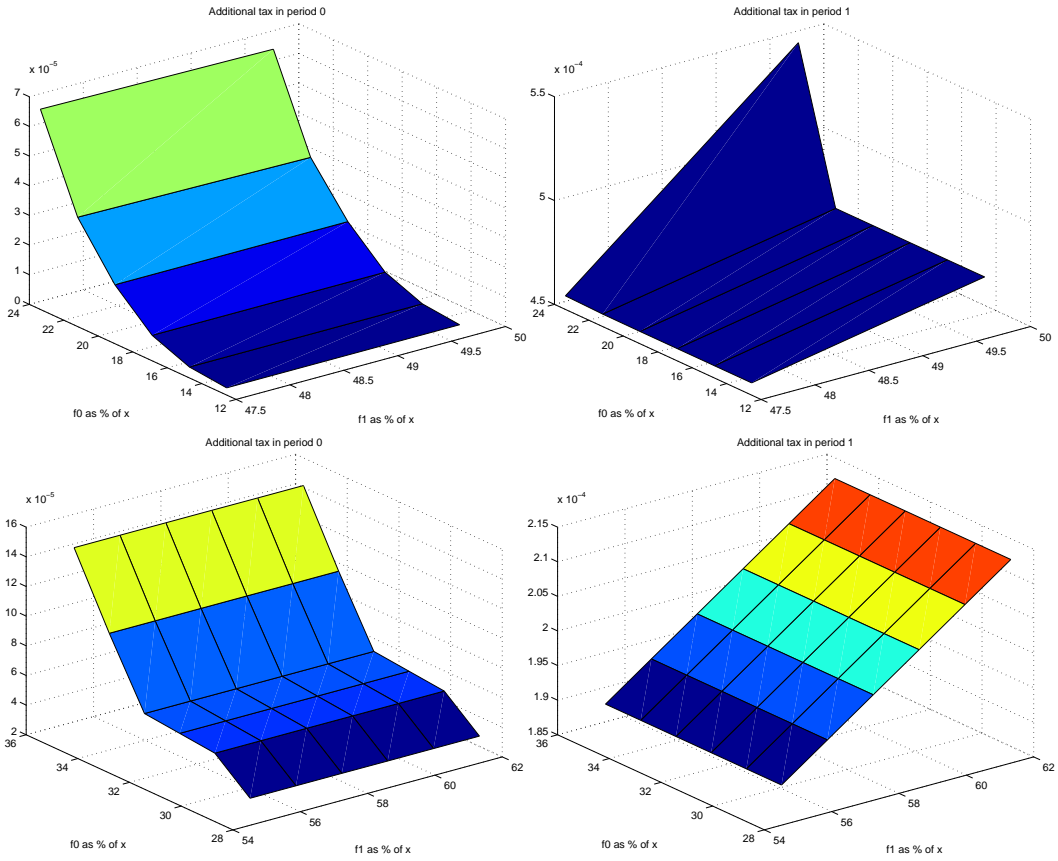
Note: The left panel is for ability level 3 and the right panel for the ability level 5.

effect along with the fact that the majority of leavers (and thus potential shirkers) in this ability group decide to leave college in period 0 induces more generous feasible indemnity levels in period 0 relative to the case where insurance is offered only in period 0. Figure 5 illustrates the feasible (in white) versus nonfeasible (in black) combinations of indemnities for the ability levels 3 and 5 and Figure 6 illustrates the additional tax that must be collected in order to cover the mass of shirkers in each of the periods when insurance is offered for some of the feasible combinations (including the optimal mix).

Even though more generous schemes are feasible earlier in college, the optimal contract dictates a higher indemnity to be offered in period 1 relative to period 0. It is desirable to offer more insurance later in college rather than early since it encourages effort in both periods 0 and 1. Welfare increases with offering insurance in both periods up to some combination  $(f_0, f_1)$  and declines significantly, if more insurance is offered, in particular if more is offered in period 0. Recall that higher levels of  $f_0$  may deter some students from enrolling in college. Figure 7 presents welfare for some of the feasible combinations of  $(f_0, f_1)$  (including the optimal mix) for the ability levels 3 and 5.

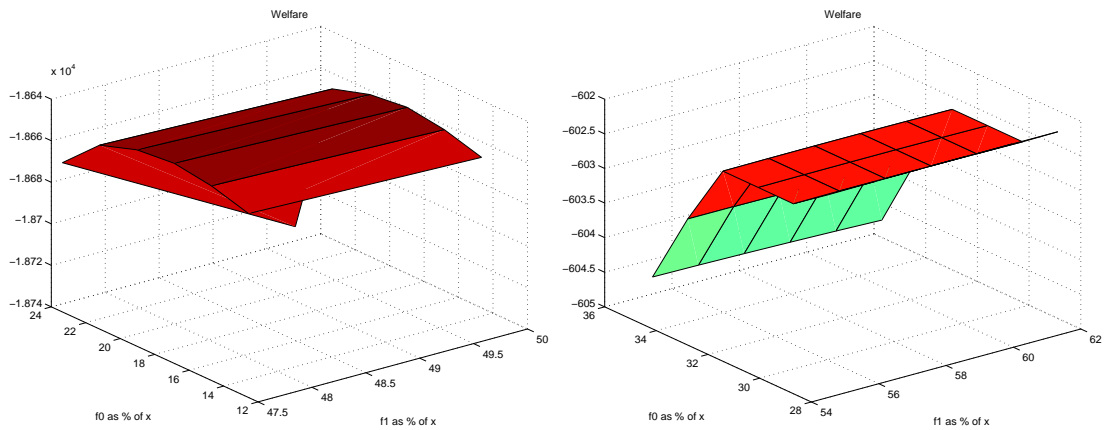
Table 7 presents the optimal mix of indemnity offered as well as the base premia and the two tax amounts that must be collected in both periods 0 and 1. As before, the optimal level of insurance in both periods combined increases in the ability level except for the lowest ability group, which is a result of the order of the distribution of  $\gamma$ . Results confirm that the shirking problem is more severe for the second ability group, who, on average, draws a

Figure 6: Additional tax collected when insurance is provided



Note: The top panel is for ability level 3 and the bottom panel for the ability level 5.

Figure 7: Welfare when insurance is provided



Note: The left panel is for ability level 3 and the right panel for the ability level 5.

higher effort in college, followed by the third ability group. In fact, the contract dictates that a combination of no shirking in period 1 and shirking in period 0 is optimal for these two ability groups. As mentioned, for these groups shirking in period 1 is too costly. Recall that there is a significant fraction of people in these ability groups who decide to leave college. In addition, most of them decide to leave college later rather than early during their college years (99.5% and 91.5% of leavers in these ability groups leave college in period 1). Thus, the shirking problem is more severe in period 1. In contrast, fewer people in the high ability groups leave college. In addition, they are more likely to leave later in college relative to the low ability groups (64% and 12% of leavers in the fourth and the fifth ability groups decide to leave college in period 1). As a result, the contract dictates that insurance that induces shirking in both period 0 and period 1 is optimal for ability group 4 and insurance that induces shirking only in period 1 is optimal for the top ability group.

Also note that the indemnity levels in the case where insurance is offered in both periods in college are in general lower than the levels offered in the case where insurance is offered only in one period in college. As mentioned, for the level of  $f_1^*$  offered in the case where  $f_0 = 0$ , more shirking may occur when  $f_0 > 0$  and insurance is more costly. At the same time, for the level of  $f_0^*$  offered in the case where  $f_1 = 0$  less or more shirking is possible when  $f_1 > 0$ . This effect is positive or negative depending on the timing of leaving from college. For instance, in the case of the top ability group (where the majority of leavers decide to leave college in period 0), insurance in period 1 induces the high ability college leavers to put in effort in period 0 rather than shirk. As a result, a much higher indemnity is offered in period 0 without inducing shirking in the case where insurance is offered in both periods relative to the case where insurance is offered only in period 0. The opposite is true, however, for the second ability group (where the majority of leavers decide to leave college in period 1). Insurance in period 1 is very costly and may limit the amount of insurance that can be offered to this ability group in period 0.

The optimal insurance induces an increase in enrollment rates from 88.6% to 94.02%.

**Table 7: Optimal shirking insurance: periods 0 and 1**

SAT scores	$\leq 700$	701 – 900	901 – 1100	1101 – 1250	$\geq 1251$
Indemnity $f_0^*$	0.0128	0.0041	0.0041	0.0065	0.01
$f_0^*$ as perc of 2x	50	11.65	9.45	11.4	13.62
$s_0^*$ as perc of 2x	61.9	5.1	3.5	2.95	1.91
Indemnity $f_1^*$	0.0113	0.0061	0.0065	0.0149	0.0305
$f_1^*$ as perc of 2x	44.04	17.3	23	26.45	41.55
$s_1^*$ as perc of 2x	54.6	7.6	8.5	6.85	5.81
$\tau_0^*$	0	1.187e-05	2.383e-05	4.809e-6	0
$\tau_1^*$	1.092e-4	0	0	1.386e-4	5.71e-05

**Table 8: Enrollment, leaving and completion rates: insurance**

SAT scores	$\leq 700$	701 – 900	901 – 1100	1101 – 1250	$\geq 1251$
Enrollment rates	92.5	88.8	94.3	97.6	100
Early leaving rates	0	4	2.91	0.75	0.97
Completion rates	44.7	66.6	70.8	78.5	86.7
Shirking rates	0.43	0.29	0.69	0.73	0.16

There are 0.61% students who shirk, on average. Out of everyone who enrolls, only 2.5% decide to leave college compared to 4.8% in the case where insurance is not offered. The average completion rate increases from 70.1% to 71.64% out of everyone who enrolls. The combination of these effects delivers that the percentage of high school graduates who acquire a college degree increases from 62% to 67.4%. Offering insurance increases the value of putting effort in college and thus induces more people to stay in college. This induces an increase in completion rates. While some of the marginal students who decide to enroll with insurance may decide to shirk and thus will counteract the positive effect on completion rates, this negative effect is secondary across all ability groups. Table 8 presents these rates across ability groups.

The different incentives to shirk across ability groups deliver different welfare consequences across these groups of students. We use an aggregate welfare measure equally weighted. There is a welfare gain of 3.03% on average in the optimal contract relative to the benchmark economy. Offering insurance in both periods is better than offering in-

insurance (that induces shirking) in period 1 or period 0 only. The average welfare gain from offering insurance in both periods relative to offering the contract only in period 0 is 2.91 % and relative to offering the contract only in period 1 is 1.15 %. Finally, offering insurance that induces a fair amount of shirking is better than offering insurance that does not induce shirking (offered in period 1). The welfare gain from offering insurance with shirking is 0.40% relative to the no-shirking insurance. Table 9 presents the welfare changes across the five ability groups in the case where the optimal contract is implemented relative to the benchmark case where no insurance is offered and relative to the case where the contract is offered only in period 0 or only in period 1 with shirking. We also present welfare gains from offering insurance in both periods in college relative to offering insurance in period 1 without inducing shirking. The group that benefits the most from the optimal contract is the lowest ability group, the one for whom the incentive to shirk is lowest. All ability groups present a higher welfare gain in the case where the optimal contract is offered in both periods during college relative to the case where insurance is offered during only one period in college. Our results suggest that offering insurance in period 1 is preferred to offering insurance in period 0, and offering insurance in both periods while recognizing a fair amount of shirkers is optimal accounting for that fact that loan administrators must collect enough tax in case of success in both periods to cover the mass of shirkers.

**Table 9: Welfare changes: insurance with shirking**

SAT scores	$\leq 700$	701 – 900	901 – 1100	1101 – 1250	$\geq 1251$
Rel to Benchmark	7.68	3.21	2.85	2.84	2.49
Rel to Period 0	5.76	3.08	2.8	2.78	2.49
Rel to Period 1	3.4	2.09	0.97	0.42	0.26
Rel to No-Shirking	3.4	0.47	0.32	0.25	0.52

## 7 Conclusion

We develop a theoretical framework that studies whether the student loan program can gainfully offer insurance against college failure risk. We conduct the analysis under two

important constraints on the provision of this sort of insurance. First, we assume that any proposed insurance scheme cannot redistribute resources from people with high probability of completing college to people with low probability of completing college. The government loan program does not permit students with superior post-college labor market outcomes to partially pay the college expenses of students with poor post-college labor market outcomes. Second, we assume that the insurance program must guard against moral hazard, that is, the possibility that provision of insurance against college failure risk may increase the risk of failure and adverse selection, that is the possibility that poor risks will attempt to pool with the good risks.

We argue that such an insurance scheme is administratively feasible and provide conditions under which these constraints leave open the possibility of some insurance against college failure risk. We calibrate the model to US data on college enrollment, early leaving and completion rates as well as the average indebtedness of program participants and quantify the effects on college enrollment, leaving, dropout rates, and welfare. Results suggest that it is optimal to offer more insurance later during college years rather than early. The optimal insurance offered in case of non-completion ranges from 10 to 40 percent of total college cost, depending on the ability of the student and the time when insurance is offered. The insurance scheme induces an increase in enrollment rate of 5.4 percentage points and an increase in college graduates of 5.4 percentage points. Although insurance draws in students with a high risk of failure, completion rates rise because fewer students drop out voluntarily from college. The leaving rate declines by more than three times. Insurance increases welfare by 3 percent on average.

## References

- AKYOL, A., AND K. ATHREYA (2005): “Risky Higher Education and Subsidies,” *Journal of Economic Dynamics and Control*, 29(6), 979–1023.
- CAMERON, S., AND C. TABER (2001): “Borrowing Constraints And the Returns to School-

- ing,” NBER working paper.
- CARD, D. (2001): “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems,” *Econometrica*, 69(5), 1127–1160.
- CARNEIRO, P., AND J. HECKMAN (2002): “The Evidence on Credit Constraints in Post Secondary Schooling,” *The Economic Journal*.
- COLLEGEBOARD (2001): “Trends in College Pricing,” .
- (2007): “College Bound Seniors: Total Group Profile Report,” .
- CUNHA, F., AND J. HECKMAN (2009): “The Economics and Psychology of Inequality and Human Development,” *Journal of the European Economic Association*, 7(2-3), 320–364.
- HECKMAN, J. (1999): “Doing it Right: Job Training and Education,” *Public Interest*, 135, 86.
- HENDRICKS, L., AND T. SCHOELLMAN (2009): “Student Abilities During the Expansion of US Education, 1950-2000,” Working paper.
- HOXBY, C. (2003): *The economics of school choice*. University of Chicago Press.
- IONESCU, F. (2009): “Federal Student Loan Program: Quantitative Implications for College Enrollment and Default Rates,” *Review of Economic Dynamics*, 12(1), 205–231.
- L. BERKNER, S. H., AND E. F. CATALDI (2002): “Descriptive Summary of 1995-96 Beginning Postsecondary Students: Six Years Later,” *National Center for Education Statistics 2003151*.
- M. AMADOR, I. W., AND G. ANGELETOS (2006): “Commitement vs Flexibility,” *Econometrica*, 74(2), 365–396.
- MANSKI, C. (1983): “College Choice in America,” Harvard University Press, 79 Garden Street, Cambridge, MA 02138-9983.

- (1992): “Parental Income and College Opportunity. DSC Report Series,” .
- MELGUIZO, T. (2008): “Quality Matters: Assessing the Impact of Attending More Selective Institutions on College Completion Rates of Minorities,” *Research in Higher Education*, 49(3), 214–236.
- MURPHY, K., AND F. WELCH (1992): “The structure of wages,” *The Quarterly Journal of Economics*, 107, 285–386.
- NCES (2004): “The Condition of Education 2004 ,” *Washington, DC: U.S. Government Printing Office*, (077), U.S. Department of Education.
- RESTUCCIA, D., AND C. URRUTIA (2004): “Intergenerational Persistence of Earnings: The Role of Early and College Education,” *The American Economic Review*, 94(5), 1354–1378.
- STINEBRICKNER, T., AND R. STINEBRICKNER (2007): “The Effect of Credit Constraints on the College Drop-Out Decision A Direct Approach Using a New Panel Study,” NBER Working paper.
- TITUS, M. (2006): “Understanding college degree completion of students with low socioeconomic status: The influence of the institutional financial context,” *Research in Higher Education*, 47(4), 371–398.
- WILLIS, R. (1986): “Wage Determinants,” *Ashenfelter, O. and Layard, R., Handbook of Labor Economics*, North Holland, Amsterdam.
- WILLIS, R., AND S. ROSEN (1979): “HEducation and Self-Selection,” *Journal of Political Economy*, 87(5), 7–36.