WORKING PAPER NO. 10-1
INSURING COLLEGE FAILURE RISK

Satyajit Chatterjee
Federal Reserve Bank of Philadelphia
and
Felicia Ionescu
Colgate University

November 2009

Insuring College Failure Risk*

Satyajit Chatterjee[†]
Federal Reserve Bank of Philadelphia

Felicia Ionescu[‡]
Colgate University

November 2009

Abstract

Participants in student loan programs must repay loans in full regardless of whether they complete college. But many students who take out a loan do not earn a degree (the dropout rate among college students 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 so, we develop a model that accounts for college enrollment, dropout, and completion rates among new high school graduates in the US and use that model to study the feasibility and optimality of offering insurance against college failure risk. We find that optimal insurance raises the enrollment rate by 3.5 percent, the fraction acquiring a degree by 3.8 percent and welfare by 2.7 percent. These effects are more pronounced for students with low scholastic ability (the ones with high failure probability).

Keywords: College Risk; Government Student Loans; Optimal Insurance

JEL Codes: D82; D86; I22;

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.

*The authors thank Orazio Attanasio, Lutz Hendricks, Narayana Kocherlakota, Luigi Pistaferri, Victor Rios-Rull, and Viktor Tsyrennikov for helpful comments and Matt Luzzetti for excellent research assistance. Comments from participants at the NBER-EFACR group, the Society of Economics Dynamics, Econometrics Society and Midwest Macroeconomic Meetings and seminar participants at Cornell University and the University of Connecticut are also gratefully acknowledged. The views expressed here are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers/.

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

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

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 (2009a)).[1]

However, there is considerable financial risk in taking out a student loan because many students do not complete college. Using the 1990 Panel Study of Income Dynamics (PSID), Restuccia and Urrutia (2004) document that 50 percent of people who enroll do not complete college. Using the NCES data and surveys, we find that 37 percent and 35 percent of students enrolled in 1989-90 and 1995-96, respectively, do not possess a degree and are not enrolled in college five years after their initial enrollment.

The financial risk implied by these facts is evident in the Survey of Consumer Finances (SCF). For the five surveys conducted between 1992 and 2004, the percentage of non-students with a student loan who report not having either a 2- or 4-year college degree is 47 percent, on average. Furthermore, non-students with loans but without a degree have a significantly higher (education) debt burden. Table 1 reports the ratio of median education debt to median income among non-students with student loans, 10 or more years after first taking out the loan. As is evident, students without degrees have a significantly higher debt burden than degree holders.

Table 1: College debt burden by completion status

	10+ years since taking out the loan	
Survey year	w/ degree	w/o degree
1992	0.09	0.12
1995	0.06	0.14
1998	0.12	0.15
2001	0.12	0.10
2004	0.14	0.24

The financial risk of taking out a student loan but being unable to complete college may discourage some people from taking out a loan and enrolling in college. Thus, even though prospective students may not 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.[2]

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

[1] For detailed evidence on how financial aid affects students' college-going behavior, see Dynarski (2003) and Hoxby (2004). The former study presents evidence that financial factors represent an important determinant of both enrollment and persistence. The latter provides a comprehensive perspective on the issue of college choice, examining it from both an individual and institutional point of view. Also, for an extensive analysis of the college financial system's weaknesses and strengths, see Kane (1999).

[2] Heckman (1999) has pointed out that the erosion of average real wages between 1980 and 1990 could have been mitigated (in an accounting sense) if more people had acquired college degrees. Specifically, for the 1990 workforce of 120 million, 5.4 million more would have to become college equivalents to reverse the 1980-1990 erosion of real wages, 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.

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, 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.[3]

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 a high probability of completion to people with a 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 adverse selection (the possibility that poor risks will attempt to pool with good risks). As we verify later, moral hazard is not an issue in this context because insurance against college failure risk increases the value of exerting effort in college.

In the theoretical section of the paper, we develop a simple model of a student's enrollment and college effort decisions. The model postulates the necessary heterogeneity in student characteristics in order to be consistent with the diversity of enrollment and effort decisions we see in reality and the importance generally assigned to ability heterogeneity and self-selection into college attendance and completion by researchers (see, for instance, Venti and Wise (1983)). The heterogeneity is in a student's utility cost of putting effort into college and his or her outside option, neither of which is 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, leaving, and completion rates as well as the average college costs of program participants, distinguishing between students of different scholastic ability levels as measured by SAT scores. 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 45 percent of total college cost. The insurance scheme induces an increase in the enrollment rate of 3.5 percentage points and an increase in college graduates of 3.8 percentage points. Although insurance draws in students with a high risk of failure, completion rates rise because fewer students voluntarily leave college. Insurance increases welfare by 2.7 percent on average.

[3]The borrower is permitted to discharge her loan only if a repayment effort over 25 years does not fully cover all obligations.

There is a rich literature on higher education, with important contributions focusing on college enrollment and completion. Studies that take a quantitative-theoretical approach have given a prominent role to the risk of college failure. These include studies by Akyol and Athreya (2005); Caucutt and Kumar (2003); Garriga and Keightley (2007); Ionescu (2009b); Restuccia and Urrutia (2004). But these studies do not generally consider the possibility of providing insurance against this risk. One exception is Ionescu (2009b) who studied the effects of alternative bankruptcy regimes for student loans. She shows that individuals with relatively low ability and low initial human capital levels are affected to a greater degree by the risk of failure and the option to discharge one's debt under a liquidation regime helps alleviate some of this risk.[4] Also, with the exception of Garriga and Keightley (2007) none of these studies recognize that students may choose to drop out.[5]

Empirical research on college behavior, however, calls for a careful modeling of college dropout behavior. Manski and Wise (1983) argue that college students learn over time about what college means and given this learning some choose to drop out. In addition, they suggest that college preparedness is more important than college aspiration for college completion. Furthermore, Stinebrickner and Stinebrickner (2008) show that most of the attrition among students from low-income families cannot be attributed to short-term credit constraints.[6] In a companion paper, Stinebrickner and Stinebrickner (2009) provide evidence on the relative importance of the most prominent alternative explanations for dropout behavior and find that learning about ability plays a particularly important role in this decision. Among other possible factors of importance, they find that students who find school to be unenjoyable are unconditionally much more likely to leave. But this effect seems to arise to a large extent because these same students also tend to receive poor grades. In our model, dropout behavior will arise for similar reasons.

Our paper is related to studies that focus on merit-based policies. Our insurance arrangement can be interpreted as being merit based: as we show later in the paper, the insurance premium is lower for higher ability types and the amount of insurance offered is higher as well. However, unlike merit-based aid, our insurance arrangement has no aid or grant component – it is self-financed with respect to each individual who participates, in expectation. Caucutt and Kumar (2003) analyze various types of college subsidies and conclude that merit-based aid that uses any available signal on ability increases educational efficiency with little decrease in welfare. Gallipoli, Meghir, and Violante (2008) examine the partial and general equilibrium effects of wealth-based and merit-based tuition subsidies on the distribution of education and earnings. In

[4]Although insurance against college failure risk is not the focus of their paper, Akyol and Athreya (2005) observe that the heavy subsidization of higher education directly mitigates the risk of college failure by reducing the college premium.

[5]Garriga and Keightley (2007) model college as a multi-period risky investment with endogenous enrollment, time-to-degree, and dropout behavior. The focus of their paper is on the effects of broad-based tuition subsidies and merit-based education policies on college enrollment and completion behavior rather than insurance against college failure risk.

[6]The authors use unique longitudinal data that have been collected specifically for this type of purpose at Berea College. Despite the fact that the direct costs to students at Berea are approximately zero, the authors document that 50 percent of students do not graduate.

related work, Redmon and Tamura (2007) use a Mincer model of human capital with ability differences to characterize the optimal length of schooling by ability class and the importance of school district composition for growth and distribution.

The key contribution of this paper is to construct a theory consistent with the reality of college enrollment, leaving and completion behavior as well as returns to education and use it to design an insurance scheme against the risk of failing at college, recognizing adverse selection and moral hazard. In addition, we map the model to the data and quantify the effects of alternative insurance arrangements on enrollment, completion and welfare. 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 are 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 $t = \{0, 1, 2, \ldots\}$. 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 are drawn from a distribution $H(y)$. At the time of the enrollment decision, the student knows θ but not the realization of y.

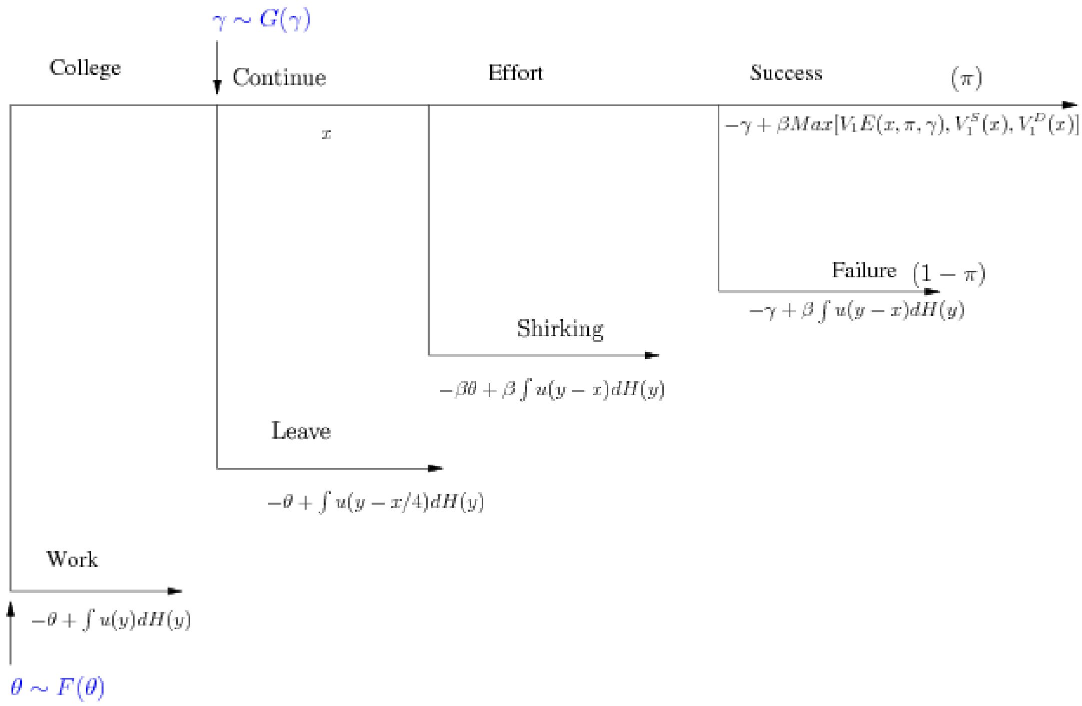
If the individual chooses to enroll in college, she learns the cost of making effort in college. Effort, e, is a binary variable that can take values 0 (no effort) or 1 (effort).[7] The cost of making an effort is denoted γ and the student draws $\gamma \geq 0$ from a distribution $G(\gamma)$. After she learns γ the student decides whether to continue on in college. If she chooses to leave, she incurs the cost of effort θ in the low-paid job and draws her (life-time) earnings y in period 1. She also incurs some partial college expenses ϕx, where $0 < \phi < 1$.[8] At the time of choosing whether to continue in college, the student knows γ and θ 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

[7]The assumption that effort is binary is essentially without loss of generality. Given the large college premium in earnings it is safe to assume that if a student finds it optimal to exert any effort in college, he or she would want to exert the maximum effort possible.

[8]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



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 as 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 in which 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. If she does not continue, she gets an earnings draw y from the distribution $H(y)$ and 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 π. If she succeeds in completing, she draws her life-time earnings y from the distribution $M(y)$ and 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 individuals 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, θ 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: $U(c) : R \to R^{++}$ 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 stochastic 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 π.*

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 γ, 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 π 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 π. 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 π. □

It is perhaps worth noting that the threshold γ will be zero for sufficiently low probability of success π. 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 $\overline{\pi}_1 > 0$ such that $V_1(x, \overline{\pi}_1, 0) = 0$. For all $\pi < \overline{\pi}_1$, $V_1^E(x, \pi, \gamma) - V_1^D(x) < 0$ for all $\gamma \geq 0$. Therefore, the threshold $\gamma_1(x, \pi)$ is 0 for all $\pi \leq \overline{\pi}_1$.

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 θ, 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 θ 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. □

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 θ. 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 π and θ.*

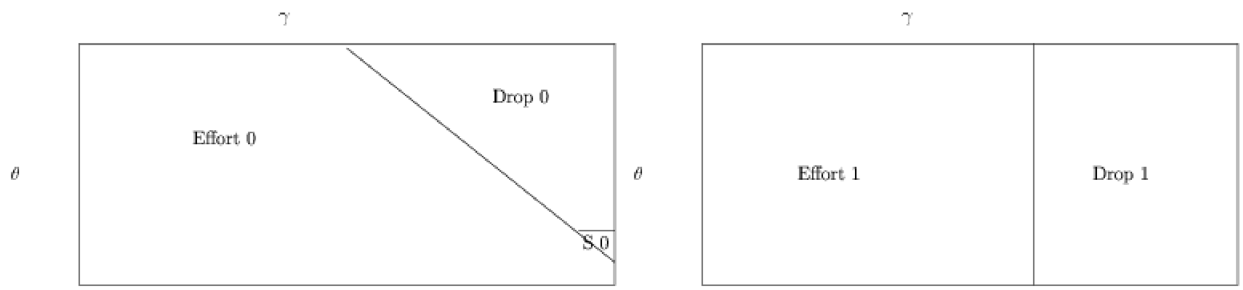
Proof. Consider the function $V_0(x, \pi, \gamma, \theta) = V_0^E(x, \pi, \gamma) - \max[V_0^D(x, \theta), V_0^S(x, \theta)]$ which is continuous for all $(\pi, \gamma, \theta) \in [0, 1] \times [0, \infty) \times [0, \infty)$ and strictly increasing in π (by Assumption 2), strictly decreasing in γ and strictly increasing in θ. 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 γ 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 π 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 θ, 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 θ. □

These propositions can be conveniently seen in Figure 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, θ, and the effort level required in college, γ.

Figure 2: Choices in periods 0 and 1



Propositions 3.1 and 3.3 give us two thresholds for γ. 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 θ, $\gamma_0(x, \pi, \theta) < \gamma_1(x, \pi)$ and for sufficiently high value of θ, $\gamma_0(x, \pi, \theta) > \gamma_1(x, \pi)$.*

Proof. We will evaluate $V_0(x, \pi, \gamma, \theta)$ at the value of γ 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 θ sufficiently close to 0, $V_0(x, \pi, \gamma_1(x, \pi), \theta) < 0$. Hence, for θ 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 θ sufficiently large $V_0(x, \pi, \gamma_1(x, \pi), \theta) > 0$. Hence for θ sufficiently large, $\gamma_0(x, \pi, \theta) > \gamma_1(x, \pi)$. □

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 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 θ.

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 γ. However, if the student's probability of success is sufficiently low, she may choose not to enroll because regardless of the value of γ 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 π.*

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 θ. 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 θ 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

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

Observe that the above change is non-negative because the positive Δterm contributes Δ while the negative Δ 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 θ.

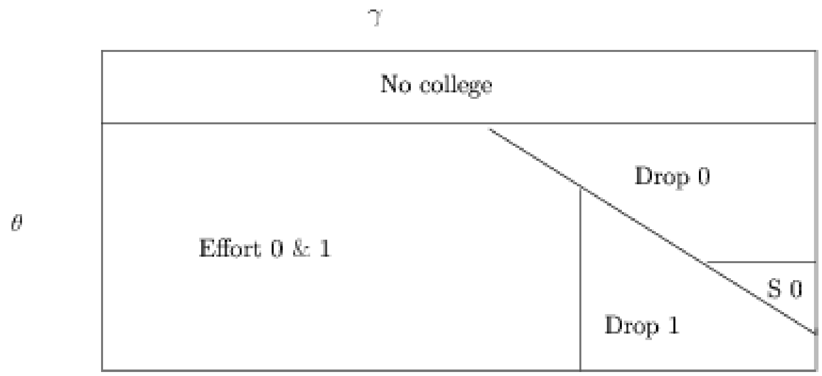
Since $V_C(x,\pi,\theta)$ is increasing in θ, if enrolling is optimal for some θ, enrolling must also be optimal for any $\hat{\theta}$ greater than θ. Therefore, there must be a cut-off value $\theta_C(x,\pi) \geq 0$ such that for all $\theta > \theta_C(x,\pi)$ the student will find it optimal to enroll 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 π observe that $V_0^E(x,\pi,\gamma)$ is strictly increasing in π and, therefore, $V_C(x,\pi,\theta)$ is strictly increasing in π. It follows that the cut-off $\theta_C(x,\pi)$ cannot be strictly increasing in π. □

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 period 0, but fail to complete their courses satisfactorily with probability $1-\pi$. Fourth, among students who successfully complete their courses in period 0, 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, θ and γ.

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 π. 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

Figure 3: Choices in college



observable index a that varies positively with probability of success π. 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 section 5, 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 π. 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 in college – are 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,

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

where $\tilde{G}(\gamma) = \min\{1, \frac{G(\gamma)}{G(\gamma_0(x,\theta,\pi(a)))}\}$ is the distribution of γ 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 π and the assumption that $\pi(a)$ is increasing in a. □

Next we show that average earnings are increasing in scholastic ability. By average earnings of a scholastic group a we mean

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

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

Proof. Follows from Proposition 3.5, which established that $\theta_C(x,\pi)$ is decreasing in π and therefore $\theta_C(x,\pi(a))$ is decreasing in a and Proposition 3.6, which delivered that $n(a)$ is decreasing in a, and Assumption 3, which implies $\int ydM(y) > \int ydH(y)$. □

These propositions relied on the assumption that a affected π only. It is possible that a also affects other primitives, for instance, the distribution from which the effort cost γ is drawn, the distribution from which earnings y are 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 a high probability of success (high ability) to students with a 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 into 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 $-\pi \cdot s_1 + (1-\pi) \cdot f_1 = 0$. Ignoring the $-\gamma$ term, expected utility given these transfers is then

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

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 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 $f_1 \in [0, 2x]$. □

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 β 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 makes dropping out better than 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 π and the premia collected will fail to cover loss claims.[9]

[9] It is not certain that these students will find it optimal to shirk. The reason is that insurance also increases the value of

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 π. In contrast, less restrictive insurance schemes may induce shirking and raise the probability of failure above π since shirkers fail with probability 1. The endogeneity of the failure probability makes the general insurance problem difficult. 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

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

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

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

putting in effort in college.

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 (θ,γ) 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:

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

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 the 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 $\bar{f}_1$. □

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. □

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" premia 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 premia that need 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 γ above which an enrolled student will not put in effort in college in period 0 (i.e., she will either drop out or shirk). This cut-off solves

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

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 θ 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

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

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 γ above which the student does not put effort in college in period 1. This cut-off solves

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

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 the feasibility of f.

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

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

and

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

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 (1) is the measure of enrolled students who put in effort in period 0 and succeed. Each of them pays the additional premium $\tau_0^*(f)$. The term on the rhs of (1) is the measure of enrolled students who do not put in effort in college *and* shirk. Each of them collects f_0 from the insurance scheme. For feasibility, the two sides must balance. Similarly, the term multiplying $\tau_1^*(f)$ on the lhs of (2) is the measure of students who put in effort in period 1 and succeed (as before $\tilde{G}$ is the distribution of γ conditional on the set of γ for which students put in effort in period 0). Each of them pays the additional premium $\tau_1^*(f)$. The term on the rhs of (2) 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. Φ 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 Φ 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: Choices when insurance is provided

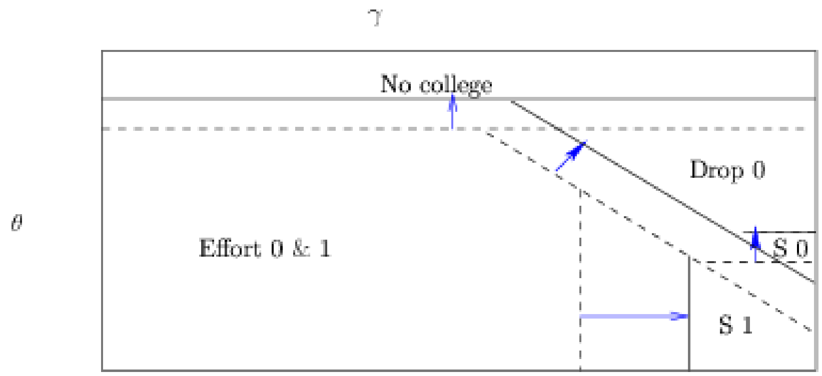


Figure 4 indicates the effects of optimal insurance. Insurance increases the value of going to college and, thus, shifts up the the γ_0 and γ_1 loci. Thus, insurance increases the fraction of students putting in effort in both periods. If optimal insurance requires $f_0 > 0$, then it shifts down the θ_0 locus – of the students who choose not to put in effort in college in period 0, a bigger fraction choose to continue on in college and shirk. Both effects work to lower dropout rates in period 0. Dropout rates also decline in period 1 because γ_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 the non-completion rate is ambiguous because it encourages some students who were dropping out to put in effort (this is the positive effect) and others who were dropping out to shirk (the negative effect). Of course, optimal insurance raises the enrollment 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.

Note that the insurance friction here is entirely about adverse selection. Optimal insurance never encourages anyone who was putting in effort in college to stop putting in effort. Indeed, it encourages people who were not putting in effort to put in effort. The friction is simply that some students who were choosing to drop out may choose to continue on in college without putting in effort (there is no change in their college effort decision). Thus, insurance attracts students whose failure probability is 1. This is an extreme form of adverse selection.

Some additional comments are worth making. First, we are implicitly assuming that once a student fails college, he or she never attempts college again. If we were to relax this assumption, the insurance arrangement would need to specify that once a student avails herself of insurance, she cannot re-enroll in college without re-paying the indemnity with interest.

Second, we are abstracting from the adverse effects on the private returns to college education that may stem from policy-induced increases in college completion rates.[10] On the other hand, we are also abstracting from the myriad social benefits of a more educated populace.

Finally, the following caveat should be kept in mind regarding the optimality of the insurance arrangement. Because higher education is subsidized by federal and state governments, changes in enrollment and completion rates induced by insurance will change the level of subsidy being received by the higher education sector. The welfare costs of this additional subsidy are being ignored here.

5 Mapping the Model to Data

The first issue that must be dealt with is that students vary in their probability of success π. Furthermore, the insurance arrangements discussed in the previous section assume that each student's π is observable to student loan administrators. So the first task is to pool students with respect to some observable index of the probability of success in college. We use SAT scores for this purpose. In particular, we classify students in 5 groups. Table 2 gives the distribution in 1999 of students who took the SAT.

Table 2: Distribution of SAT scores

SAT scores	$0-699$	$700-900$	$901-1100$	$1101-1250$	$1251-1600$
Fraction	0.079	0.224	0.342	0.205	0.15

In what follows, we will consider only the four top groups. We will denote these groups by the index $i \in \{1, 2, 3, 4, \}$.

There are 4 parameters and 4 distributions in the model. Among the parameters are 2 preference parameters σ and β and 2 college parameters x and π. Among the distributions are distributions for the (unobserved) heterogeneity $F(\theta)$ and $G(\gamma)$ and the distributions of earnings of non-college and college workers $H(y)$ and $M(y)$. We assume that all students have the same preference parameters and draw from the same distribution of the "outside option" $F(\theta)$ but we allow the parameters x and π and the distributions $G(\gamma)$, $H(y)$ and $M(y)$ to depend on i. Naturally, we expect $\pi(i)$ to increase with i. We also expect the distribution $G(\gamma)$ to depend on i because the utility cost of exerting effort in college is, plausibly, more likely to be lower for a student with a higher SAT score. We also expect x to depend on i because students with higher SAT scores tend to go to more selective colleges and these colleges tend to have higher tuition.[11] This tendency for x

[10] Card and Lemieux (2001) as well as Bound, Lovenheim, and Turner (2009) find evidence of congestion effects in higher education: an increase in the number of people seeking higher education tends to be associated with a decline in educational attainment.

[11] We do not explicitly analyze the matching of students of varying ability to colleges of varying selectivity, but our quantitative work recognizes the fact that students with similar scholastic abilities tend to sort into similar colleges. For details

to increase with i is partly offset by the tendency of more selective colleges to provide more financial aid. Finally, if scholastic ability is correlated with ability more broadly (as seems plausible), we also expect $H(y)$ and $M(y)$ to depend on i. In particular, we would expect students with higher SAT scores to be more likely to draw a higher y.

5.1 Preference Parameters, Earnings Distributions and College Costs

We assume that the utility function is given by

$$U(c) = \begin{cases} (c+\epsilon)^{1-\sigma}/(1-\sigma) & \text{if } c > 0 \\ \epsilon^{1-\sigma}/(1-\sigma) & \text{if } c \leq 0 \end{cases}$$

where ϵ is a small positive number. Thus the utility function is defined over the real line but is effectively CRRA with coefficient of relative risk aversion of σ for $c >> 0$. We set $\sigma = 2$ and $\beta = 0.97$, both conventional values in quantitative macroeconomics.

In the theory, y is the person's lifetime earnings. We calibrate the lifetime earnings distributions using earnings data from the CPS for 1969-2002 for synthetic cohorts. There are 5000 observations in each year's sample, on average. We distinguish between two education groups: those with at least 12 years but less than 16 years of completed schooling and those with at least 16 years of completed schooling. The former corresponds to the non-college group and the latter to the college group. For each education group, we calculate the mean real earnings of heads of households who are 25 years old in 1969, 26 years old in 1970, . . . , 58 years old in 2002.[12] The mean present value of life-cycle earnings for each group is simply the sum of the mean earnings at each age.[13] For the non-college group mean life-time earnings is \$1.07 million and for the college group it is \$1.69 million. These estimates imply a college premium of 58 percent. Micro-studies find that the increase in lifetime earnings from each additional year in college is between between 8% and 13% (see Willis (1986) and Card (2001)). Since the average college graduate has more than 4 years of college education (some students do post-graduate schooling), our calibration of the college premium is roughly consistent with the high end of this range of estimates.[14]

To estimate the variation of lifetime earnings around these mean values, we assume that the life-time earnings

on the importance of individual characteristics coupled with college characteristics for college attendance and completion, see Bound, Lovenheim, and Turner (2009), Hastings, Kane, and Staiger (2006), Hoxby (2004) and Light and Strayer (2000).

[12] To increase the number of observations in each age group, we consider five-year bins. That is, by age 25 in 1969 we mean heads of household who are between 23 and 27 years old (both inclusive) in that year. Real values are calculated using the CPI for 1999

[13] Ignoring discounting overestimates life-time earnings and ignoring earnings beyond age 58 underestimates it.

[14] Restuccia and Urrutia (2004) use a 10% rate of return, which corresponds to a lifetime college premium of about 1.5.

of an individual in education group k are given by $z(\mu_{25}^k + \mu_{26}^k + \cdots + \mu_{58}^k)$, where z is a random variable with mean 1 and variance $\sigma_z^2(k)$ and μ_n^k is the mean earnings in education group k at age n. Then, $\sigma_z(k)$ is simply the (common) coefficient of variation of earnings at any age n in education group k. We set $\sigma_z(k)$ equal to the mean coefficient of variation in earnings across all ages in education group k. This construction implies that the standard deviation of y is \$ 0.8 million for the college group and \$ 0.5 million for the non-college group.

The above calibration of the mean and standard deviation of lifetime earnings for the two education groups is for each group as a whole. Within each group, we permit the distribution of lifetime earnings of individuals to vary systematically with scholastic ability (see Cunha and Heckman (2009), Hendricks and Schoellman (2009)). We use the data set High School and Beyond (HS&B) to group students by the four ability groups $i \in \{1, 2, 3, 4\}$ and compute the mean earnings for each group of those students who are five years out from the year they acquired their highest degree and are employed full-time. We use these mean earnings to compute the mean earnings of each ability group relative to the overall mean earnings of the education group in question and then apply these relative mean earnings factors to the mean earnings in the CPS data for the corresponding education group. This yields $(\mu_i^C(y), \text{i} = 1,2,3,4) = (1.66, 1.74, 1.84, 1.91)$ and $(\mu_i^{NC}, \text{i} = 1,2,3,4) = (1.05, 1.11, 1.17, 1.21)$.[15] We assume that the standard deviation of earnings for each ability group is the same as for the group as a whole. Finally, in order to compute the relevant expected utility values, we assume that all earnings distributions are Normal.

The cost for college was \$20,706 per year for private universities and \$8,275 per 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 total college costs are \$49,508 in 1999 dollars (College Board (2001)). We consider heterogeneous costs of college. Using the same enrollment-weighted procedure, we estimate college costs across ability groups using data from the Princeton Review on college rankings in terms of average SAT scores of accepted students and data from USA Today on college costs (tuition and room and board). We estimate college costs for the 4 groups of ability levels to be: \$35,200, \$37,000, \$56,400, and \$73,400 (in 1999 dollars). Thus, we find that high ability students enroll in more expensive colleges (more selective colleges tend to be more expensive). We set college costs (in millions) $(2x_i, \text{i}=1,2,3,4) = (0.0352, 0.0370, 0.0564, 0.0734)$.

[15] We use the HS&B because the B&B data set (which reports earnings for more years) covers only college graduates while the BPS data set covers both high school and college graduates but reports earnings only upon graduation. Since earnings differentials due to ability are likely to manifest themselves gradually over time, using earnings information from some years out is preferable. We normalize the units in which earnings are measured in the model so that 1 unit means \$1 million.

5.2 Completion Probabilities and Distributions of Disutility from Effort

To calibrate π_i, we use the Beginning Postsecondary Student Longitudinal Survey (BPS 1995/96), which collects data on intensity of college attendance and completion status of post-secondary education programs for students who enrolled in 1995.

We consider only students who enroll without delay in either 2- or 4-year colleges following high school graduation. Because we do not have part-time enrollment in the model, we consider students who enroll exclusively full-time in their first academic year and enroll full-time in their first and last months of enrollment in future academic years.[16] The survey records the fraction of students (for each ability group) who, in 2001, report having earned a bachelor's degree. This is the *degree completion rate* and for our universe of students comes out to be $(c_i, \text{i=1,2,3,4}) = (0.601, 0.72, 0.825, 0.871)$.[17] These rates do not identify π_i because the universe includes students who do not put effort in college; for instance, it includes students who drop out shortly after enrolling and therefore never earn a degree. To identify π_i, we first locate students who, in 2001, report not having earned a bachelor's degree and who report having last enrolled in academic years 1995-96 or 1996-97. This group is our empirical analog of students who drop out or fail in period 0 or drop out at the start of period 1. We refer to this group as *leavers* and their fraction (in our universe of students) comes out to be $(l_i \text{ i=1,2,3,4}) = (0.088, 0.056, 0.025, 0.013)$.[18] The complement set is our empirical analog of students who are in good standing at the start of period 1 and who put in effort in college. Therefore, we obtain $(\pi_i \text{ i=1,2,3,4}) = ((0.601/(1-0.088), 0.72/(1-0.056), 0.825/(1-0.025), 0.871/(1-0.013)) = (0.659, 0.7627, 0.8462, 0.8825)$. Observe that π is increasing in SAT scores, which justifies our initial thought that SAT scores are an observable proxy for π.

The calibration of the distributions $F(\theta)$ and $G_i(\gamma)$ is achieved via moment matching. The moments we target are enrollment and leaving rates for the four ability groups. We use the National Education Longitudinal Study (NELS:88) to collect information on the college enrollment choices of students who were high school seniors in 1992. We consider a student to be enrolled in college if he or she enrolled without any delay after high school and was enrolled in either a 2-year or 4-year colleges in October 1992. The enrollment rates by our four ability groups comes out to be $(e_i \text{ i=1,2,3,4}) = (0.795, 0.894, 0.943, 0.957)$.

We assume that F is distributed normal with mean μ_θ and standard deviation σ_θ and the $G_i(\gamma)$ is distributed

[16]Since students can enroll full-time but drop out shortly thereafter, "exclusively full-time enrollment in the first academic year" simply means that the student is enrolled full-time for the months he or she is actually enrolled. For later academic years, we weaken the full-time requirement to apply to only the first and last months of enrollment. This allows students to go part-time for short stretches of time.

[17]We did not want the college performance of students with very low and very high SAT scores to overly affect the performance of their respective groups (the $700-900$ group and the $1250-1600$ group). We employed a 5% Winsorization with respect to SAT scores to reduce the sensitivity of group performance to outliers.

[18]These statistics also reflect a 5% Winsorization.

exponential with mean μ_{γ_i}. These distributional assumptions imply that there are 6 parameters to be constrained by 8 moments. The problem reduces to finding the vector of parameters $\alpha = (\mu_\theta, \sigma_\theta, \mu_{\gamma_{i=1,2,3,4}})$ that solves

$$\min_{\alpha} \left(\sum_{i=1}^{4} w_i((e_i - e_i(\alpha))^2 + v_i(l_i - l_i(\alpha))^2 \right),$$

where $e_i(\alpha)$ and $l_i(\alpha)$ are the corresponding model rates and w_i and v_i are the weights assigned to these rates.

Table 3: Enrollment and leaving rates: model and data

SAT scores	$700 - 900$	$901 - 1100$	$1101 - 1250$	≥ 1251
Enrollment rates : Data	0.795	0.894	0.943	0.957
Enrollment rates : Model	0.77	0.908	0.949	0.961
Leaving rates: Data	0.088	0.056	0.025	0.013
Leaving rates: Model	0.088	0.043	0.025	0.013

Table 3 gives the outcome of this moment matching exercise. As is evident, the match between data and model moments is quite good. We find the distributions $F(\theta) \sim (0.39, 0.21)$ and $G_1(\gamma) \sim (0.066)$, $G_2(\gamma) \sim (0.065)$, $G_3(\gamma) \sim (0.057)$, and $G_4(\gamma) \sim (0.046)$. Note that means of the γ distributions decline with ability. This is consistent with our interpretation of γ as the utility cost associated with school work. High-ability students seem to bear fewer costs (i.e., find the work more enjoyable) than low-ability students.

6 Insurance Against Failure Risk

In this section we report the results regarding insurance for each of the 4 ability groups. We follow the structure of the analysis in Section 4. For each ability group (i.e., for each π) we consider the best possible insurance when (i) effort is observable, (ii) effort is not observable and the insurance must respect the no-shirking constraint, and (iii) effort is not observable and shirking is tolerated.

6.1 Full Insurance

First, we consider the case where effort is observable. The model delivers that the level of insurance that equates marginal utilities across states, f_i^*, is $0.076, 0.104, 0.143, 0.172$ for $i = 1, ...4$. These levels are higher than the cost of college, $2x_i$, for all ability levels i (they represent 216.5%, 280.8%, 253.6%, and 234.5% of college costs by ability groups). Thus our calibrated economy satisfies Assumption 5. So, when effort is observable, it is optimal to insure students of all ability groups up to the full cost of college.

6.2 No-shirking Insurance

Recall from Proposition 4.2 that an optimal no-shirking insurance must offer $f_0 = 0$ in period 0 and up to $\bar{f}_{i\,1}$ in period 1, where $\bar{f}_{i\,1}$ satisfies $\int U(y - 5x_i/4)dH_i(y) = \beta \int (y - 2x_i + \bar{f}_{i\,1})dH_i(y)$ (here H_i is the non-college distribution of earnings of ability group i). An important observation is that when this level of insurance is provided, there is a positive mass of students who are indifferent between shirking and dropping out in each ability group. Our assumption is that if a student is indifferent between shirking and dropping out, she shirks. Given that, we consider giving an indemnity of 1% less than the level that makes shirking just as good as dropping out. Thus, we offer $\bar{\bar{f}}_{i\,1} = 0.99\bar{f}_{i\,1}$ in case of failure and the premium that is paid in case of success is $\bar{\bar{s}}_{i\,1} = (1 - \pi_i)\bar{\bar{f}}_{i\,1}/\pi_i$.

Table 4 presents the indemnity offered, $\bar{\bar{f}}_{i\,1}$, by ability groups, as well as the premium paid in case of success, $\bar{\bar{s}}_{i\,1}$, as percentages of the cost of college. The indemnity offered increases in ability, with the top ability group receiving more than two times more indemnity than the bottom ability group. However, given that the college cost increases in the ability level, each ability group is forgiven a roughly constant fraction of their college cost in the case where failure occurs. The bottom/highest ability group is forgiven 23.3%/24.6% of their college cost. The insurance, however, is more expensive for the low-ability groups relative to the high-ability groups: the premium is 12.1% of the college cost for the bottom ability group and only 3.3% of the college cost for the top ability group.

Table 5 displays how enrollment, leaving and completion rates change with the insurance. Since insurance increases the value of putting in effort in college, given θ there is less chance a student will want to drop out of college. Thus, there is a tendency for leaving rates to go down and completion rates to go up. On the other hand, there is a selection effect working in the opposite direction. Because insurance increases the value of putting in effort in college, it also increases enrollment. The new enrollees are students with low values of θ. Since the $\gamma_0(x, \theta, \pi)$ locus is increasing in θ, the new enrollees are more likely to drop out in period 0. For the first three ability groups, the first effect dominates and insurance causes leaving rates to fall and completion rates to rise. For the top ability group, the second effect is decisive. For this group, insurance encourages everyone to enroll and there is a sufficiently large increase in the share of "low θ" students so that leaving rates rise and completion rates fall.

Table 5 also displays the welfare gain from insurance, namely, the percentage increase in welfare with insurance relative to the no-insurance (baseline) model. As we might expect, the insurance is most valuable to students with a high probability of failure and, indeed, the welfare gains decline with rising ability.[19]

[19] These gains are in the nature of social welfare gains where the social welfare function treats students with different θ values symmetrically.

Table 4: No-shirking insurance

SAT scores	$700-900$	$901-1100$	$1101-1250$	≥ 1251
Indemnity $\bar{\bar{f}}_1$	0.0082	0.0084	0.0134	0.018
Percentage of 2x	23.34	22.74	23.7	24.55
Premium $\bar{\bar{s}}_1$	0.0043	0.0026	0.0024	0.0024
Percentage of 2x	12.08	7.61	4.31	3.27

Table 5: Enrollment, leaving and completion rates: no-shirking insurance

SAT scores	$700-900$	$901-1100$	$1101-1250$	≥ 1251
Enrollment rates with insurance	0.848	0.924	0.965	1
Enrollment rates : data	0.795	0.894	0.943	0.957
Leaving rates with insurance	0.032	0.030	0.020	0.015
Leaving rates: data	0.088	0.056	0.025	0.013
Completion rates with insurance	0.638	0.740	0.829	0.870
Completion rates: data	0.601	0.720	0.825	0.871
Welfare gains in percentage	2.83	2.38	2.06	1.86

6.3 Optimal Insurance

We consider the general insurance case where $f_i \in [0, x] \times [0, 2x_i]$. For comparison purposes, we first show the results if insurance is offered only in period 1.

The first task is to determine the set of feasible insurance schemes for each ability group. When insurance is offered only in period 1, an insurance arrangement $f = (0, f_1)$, $f_1 \in [0, 2x_i]$, is feasible if there exists a $\tau^*_{i\,1}(f)$ such that equation (2) is satisfied for ability group i. Obviously, any $f_{i\,1} < \bar{f}_{i\,1}$ is feasible because there is no shirking and $\tau_1 = 0$ will trivially satisfy the feasibility condition. To determine feasibility for $f_1 \geq \bar{f}_i$, we divide $[\bar{f}_{i\,1}, 2x_i]$ into a fine grid and for each grid point attempt to find a τ that satisfies (2). Our procedure is to iterate on τ_1. For iteration k, we set τ_1^k to the value that satisfies (2) given the decision rules corresponding to τ_1 from iteration $k-1$ (i.e., τ_1^{k-1}). We start the iterations with $\tau_1^0 = 0$. If this iterative process converges we classify that particular grid point as feasible. If the process diverges, we classify it as infeasible.

We find that the feasible indemnity levels $f_1 \in [\bar{f}_1, 2x]$ differ across ability groups. These sets turn out to be $\emptyset$, $[23.8, 29.8]$, $[24, 34.8]$, $[25.9, 50.5]$ (numbers are given in % of the college cost, $2x_i$) for i = 1,2,3,4. No insurance including and beyond $\bar{f}_{1\,1}$ is feasible for the lowest ability group. For the other three ability groups, insurance levels beyond $\bar{f}_{i\,1}$ are feasible. More f are feasible for higher ability levels.

These sets highlight the adverse selection problem. In the bottom ability group the probability of success π is low. A low π means that γ_1, the threshold above which a person does not put in effort in college, is low.

Table 6: Optimal shirking insurance: periods 0 and 1

SAT scores	$700-900$	$901-1100$	$1101-1250$	$1251-1600$
f_0^* as percentage of $2x$	29.83	10.27	11.26	13
s_0^* as percentage of $2x$	15.44	3.2	2.05	1.73
f_1^* as percentage of $2x$	15.99	18.38	21.01	30.65
s_1^* as percentage of $2x$	8.28	5.72	3.82	4.1
τ_0^* as percentage of $2x$	0.50	0.045	0.0062	0.0042
τ_1^* as percentage of $2x$	0	0	0	0.125

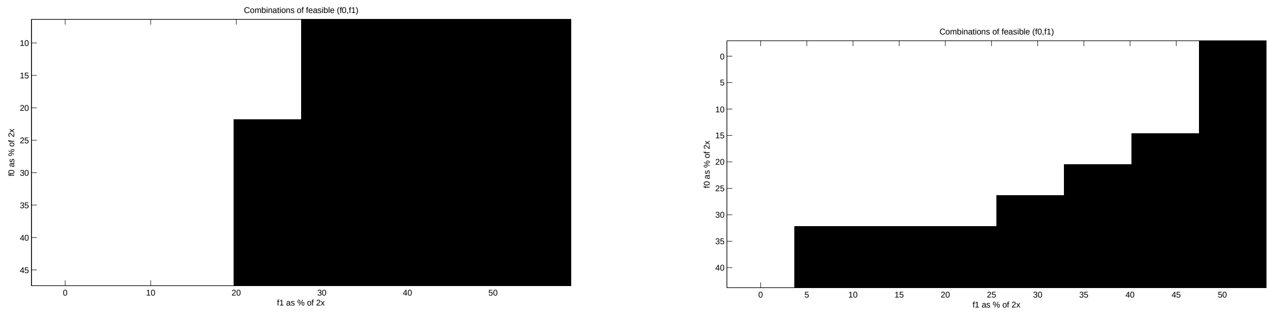
Furthermore, the mean of the G distribution for the lowest ability group is the highest. These two factors combine to make the mass of students who drop out in period 1 the highest for the lowest ability group. When insurance beyond $\bar{f}_{1\,1}$ is offered, *all* these students shirk. Thus there is a large jump in the measure of shirkers. This requires that τ_1 be increased significantly above zero to balance (2). A higher τ_1 decreases the number of students who put in effort in college and increases the number of students who wish to shirk. This makes the jump in τ_1 in the next iteration even higher still. This process of higher and higher jumps in τ_1 means that no insurance beyond $\bar{f}_{1\,1}$ can be offered. In contrast, the successive increases in τ_1 get smaller (and eventually converge to 0) for the other three ability groups – owing to the fact that γ_1 threshold is higher and the mean of the G distribution is lower.

Even though insurance higher than or equal to $\bar{f}_{i\,1}$ is feasible for the top 3 ability groups, we find that it is not optimal to offer such insurance. Thus the optimal insurance scheme, even if we allow for shirking in period 1, is to offer the best no-shirking insurance. Although insurance provides benefits for students who put in effort in college, the fact that students have to pay more than the actuarially fair insurance price ($\tau_1 > 0$) makes the net benefit of insurance at or beyond $\bar{f}_{i\,1}$ ($i = 2, 3, 4$) negative.

We turn now to the full insurance problem with shirking when insurance is offered in both periods. The calculation of feasible f is a natural extension of the method described above. For each ability group, we start with $(\tau_0, \tau_1) = (0, 0)$ and iterate on equations (1) and (2) simultaneously. If convergence is achieved, the f is classified as feasible. We find that a higher f_0 is associated with a lower f_1: if more insurance is offered in period 0, less can be offered in period 1. The reason is that period 0 insurance encourages more people to put in effort in college in period 0 and, if successful, to drop out in period 1. Thus, it increases the mass of potential shirkers in period 1 and therefore increases the cost of providing insurance beyond the "no-shirking" level. As examples, Figure 5 shows the sets of feasible f (shown in white), Figure 6 shows the associated τ_0 and τ_1, and Figure 7 presents welfare for feasible combinations of (f_0, f_1) (including the optimal mix) for ability levels 2 and 4.

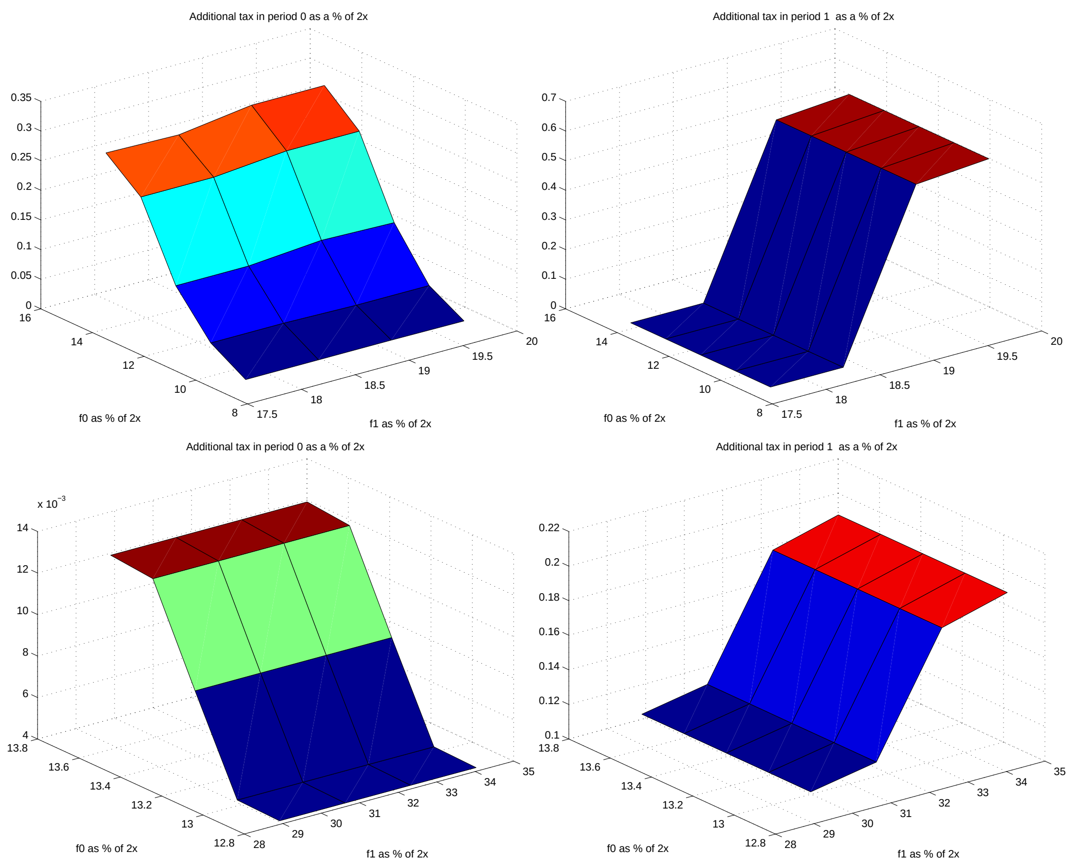
Table 6 presents the optimal mix of indemnity offered (f_0^*, f_1^*) as well as the base premia, (s_0^*, s_1^*) and (τ_0^*, τ_1^*).

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



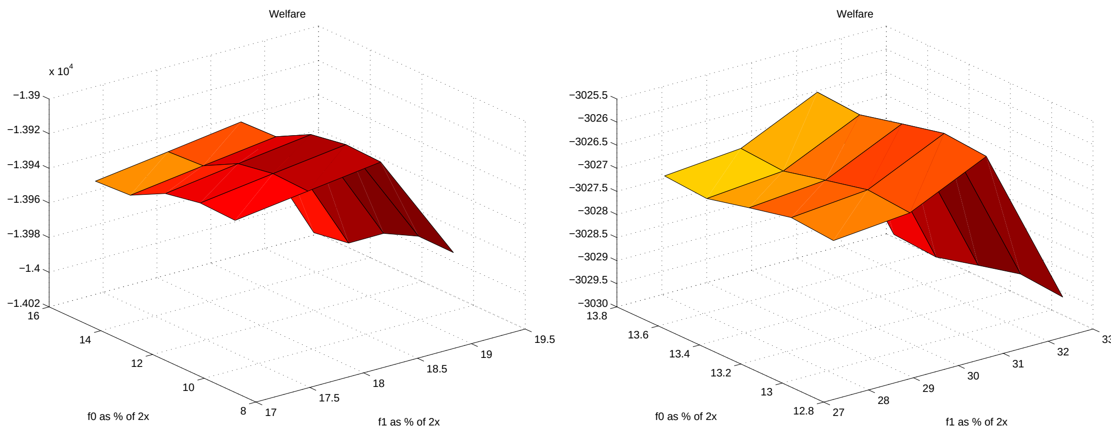
Note: The left panel is for ability level 2 and the right panel for ability level 4.

Figure 6: Additional tax collected when insurance is provided



Note: The top panel is for ability level 2 and the bottom panel for ability level 4.

Figure 7: Welfare when insurance is provided



Note: The left panel is for ability level 2 and the right panel for ability level 4.

It is optimal to offer significant amounts of insurance in periods 0 and 1. Since any positive insurance in period 0 induces shirking, the optimal insurance scheme tolerates some shirking in period 0 for every ability group. In period 1, for ability groups $i = 1,2,3$ the insurance offered in period 1 is just short of the level that would induce shirking. This is similar to the situation when insurance is offered in period 1 only, except that the "no-shirking" insurance levels are now lower. The reason is that students who succeed in college in period 0 owe the period 0 insurance premium. All else remaining the same, this reduces the value of putting in effort in college in period 1 and therefore it lowers the γ threshold above which it is better to drop out. Thus the "no-shirking" insurance levels for period 1 are lower than they would be if no insurance is offered in period 0. As in the case when insurance is offered in period 1 only, offering insurance beyond the no-shirking level in period 1 is too costly for the first three ability groups. The exception to this is the top ability group. For this group, the measure of potential shirkers in period 1 is low enough that it is optimal to go beyond the no-shirking insurance level.

Insurance offered is generally increasing in ability. This is true for insurance offered in period 1 and is true for insurance offered in period 0 for the top three ability groups. Adverse selection becomes less important as ability rises and, therefore, more generous insurance can be offered. The exception to this general rule is the lowest ability group for which the insurance offered in period 0 is quite high (higher than what is offered for any of the other ability groups). This happens because failure probability for this group is high and insurance in period 0 is more valuable than insurance in period 1.

Table 7 gives the enrollment, leaving and completion rates for optimal insurance. A comparison with Table 5 indicates that optimal insurance has virtually the same effects as insurance in period 1 only. There is an increase in enrollment and completion rates relative to the data for the bottom three ability groups. The top ability group behaves differently with respect to leaving and completion but this is due to the fact that there are more "low θ" students post insurance (the enrollment probability is 1).

Table 7: Enrollment, leaving and completion rates: insurance

SAT scores	$700-900$	$901-1100$	$1101-1250$	$1251-1600$
Enrollment rates with ins	0.869	0.924	0.966	1
Enrollment rates : data	0.795	0.894	0.943	0.957
Leaving rates with ins	0.027	0.029	0.020	0.012
Leaving rates: data	0.088	0.056	0.025	0.013
Completion rates with ins.	0.636	0.740	0.830	0.870
Completion rates: data	0.601	0.720	0.825	0.871
Shirking rates	0.187	0.46	0.06	0.35

In the aggregate, optimal insurance induces an increase in enrollment rates from 89.9% to 93.4%. On average, 0.57% students shirk. Out of everyone who enrolls, only 2.36% decide to leave college compared to 4.11% in the case where insurance is not offered. The average completion rate increases from 74.9% to 76.1% out of everyone who enrolls. The combination of these effects delivers the result that the percentage of high school graduates who acquire a college degree increases from 67.3% in the benchmark economy to 71.1%. 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. Although some of the marginal students who decide to enroll and stay in college with insurance may decide to shirk and thus will counteract the positive effect on completion rates, this negative effect is secondary.

Table 8: Welfare changes: insurance with shirking

SAT scores	$700-900$	$901-1100$	$1101-1250$	$1251-1600$
Relative to baseline model	3.54	2.66	2.28	2.01
Relative to no-shirking insurance	0.73	0.29	0.23	0.15

Table 6.3 displays the welfare gains from optimal insurance across ability groups. Two comparisons are presented. The first line displays the welfare gain relative to the baseline model. As one would expect the gain is largest for the lowest ability group and the gains decline with ability. The next line displays the gains relative to the no-shirking insurance arrangement. The gains are much smaller, indicating that the no-shirking insurance arrangement captures most of the welfare gains. In the aggregate, there is a welfare gain of 2.7% on average in the optimal contract relative to the baseline economy.

7 Conclusion

A large fraction of students who enroll in college do not earn a degree. Many of these students borrow money to finance their (failed) college education. We assume that students are cognizant of the fact that borrowing to go to college is a risky endeavor. The focus of our paper is to examine – theoretically and quantitatively – if the risk of failing to complete college (college failure risk) can be, at least partially, insured.

We conduct the analysis under two constraints on the provision of failure insurance. First, we assume that any insurance scheme cannot redistribute resources from students with a high probability of completing college to students with a low probability of completing college. Second, the insurance program must guard against adverse selection: the possibility that poor risks will attempt to pool with the good risks when insurance is offered.

We develop a model of student enrollment and effort decisions. Our model is consistent with a diversity of behavior on the part of students. We develop the notion of optimal insurance against college failure risk, taking into account the two constraints noted above. Our model predicts that some amount of insurance against failure risk is desirable and can be offered. Also, the optimal insurance scheme may tolerate some amount of adverse selection (the pooling of bad risks with good ones). We calibrate our model of student enrollment and effort decisions to match data on US college enrollment, leaving and completion rates. Using the calibrated model, we compute the optimal insurance and quantify the effect of optimal insurance on these rates as well as on welfare. We find that optimal insurance increases enrollment rates by 3.5 percentage points and increases college completion rates by 1.2 percentage points. Although insurance draws in students with a high risk of failure, the completion rate rises because fewer students drop out voluntarily from college. On average, welfare increases by 2.7 percent. We also present results broken down by ability groups. Students with relatively low scholastic ability and a high failure probability benefit the most from failure insurance. Since these students are typically from low-income backgrounds and most in need of loans to finance the expense of a college education, our results suggest that insurance against college failure risk will be particularly useful to students from low-income backgrounds.

References

Akyol, A., and K. Athreya (2005): "Risky Higher Education and Subsidies," *Journal of Economic Dynamics and Control*, 29(6), 979–1023.

Bound, J., M. Lovenheim, and S. Turner (2009): "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources," NBER Working Paper.

Cameron, S., and C. Taber (2001): "Borrowing Constraints and the Returns to Schooling," NBER Working Paper.

Card, D. (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69(5), 1127–1160.

Card, D., and T. Lemieux (2001): *Dropout and Enrollment Trends in the Post-war Period: What Went Wrong in the 1970s?* University of Chicago Press for NBER.

Carneiro, P., and J. Heckman (2002): "The Evidence on Credit Constraints in Post Secondary Schooling," *The Economic Journal*, 112(482), 704–735.

Caucutt, E., and K. Kumar (2003): "Higher Education Subsidies and Heterogeneity: A Dynamic Analysis," *Journal of Economic Dynamics and Control*, 27(8), 1459–1502.

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.

Dynarski, S. (2003): "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion," *American Economic Review*, 93(1), 279–288.

Gallipoli, G., C. Meghir, and G. Violante (2008): "Equilibrium Effects of Education Policies: A Quantitative Evaluation," Working Paper, New York University.

Garriga, C., and M. Keightley (2007): "A General Equilibrium Theory of College with Education Subsidies, In-School Labor Supply, and Borrowing Constraints," *Federal Reserve Bank of St. Louis Working Paper*, pp. 207–051.

Hastings, J., T. Kane, and D. Staiger (2006): "Gender and Performance: Evidence from School Assignment by Randomized Lottery," *American Economic Review*, 96(2), 232–236.

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, Iowa State University.

Hoxby, C. (2004): *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It.* National Bureau of Economic Research.

Ionescu, F. (2009a): "Federal Student Loan Program: Quantitative Implications for College Enrollment and Default Rates," *Review of Economic Dynamics*, 12(1), 205–231.

——— (2009b): "Risky Human Capital and Alternative Bankruptcy Regimes for Student Loans," Working Paper, Department of Economics, Colgate University.

Kane, T. (1999): *The Price of Admission: Rethinking How Americans Pay for College*. Brookings Institution Press.

Light, A., and W. Strayer (2000): "Determinants of College Completion: School Quality or Student Ability?," *Journal of Human Resources*, 35(2), 299–332.

Manski, C., and D. Wise (1983): "College Choice in America," Harvard University Press.

Redmon, C., and R. Tamura (2007): "Ability, Length of Schooling, Growth and the Distribution of Human Capital," .

Restuccia, D., and C. Urrutia (2004): "Intergenerational Persistence of Earnings: The Role of Early and College Education," *American Economic Review*, 94(5), 1354–1378.

Stinebrickner, R., and T. Stinebrickner (2008): "The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study," *American Economic Review*, 98(5), 2163–2184.

Stinebrickner, T., and R. Stinebrickner (2009): "Learning about Academic Ability and the College Drop-out Decision," NBER Working Paper.

Venti, S., and D. Wise (1983): "Individual Attributes and Self-selection of Higher Education: College Attendance versus College Completion," *Journal of Public Economics*, 21(1), 1–32.

Willis, R. (1986): "Wage Determinants," *Handbook of Labor Economics*, North Holland, Amsterdam.