View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Scholarship@Western

Western University
Scholarship@Western

Centre for Human Capital and Productivity. CHCP Working Papers

Economics Working Papers Archive

2010

2010-1 The Nature of Credit Constraints and Human Capital

Lance J. Lochner

Alexander Monge-Naranjo

Follow this and additional works at: https://ir.lib.uwo.ca/economicscibc

Part of the Economics Commons

Citation of this paper:

Lochner, Lance J., Alexander Monge-Naranjo. "The Nature of Credit Constraints and Human Capital." CIBC Centre for Human Capital and Productivity. CIBC Working Papers, 2010-1. London, ON: Department of Economics, University of Western Ontario (2010).

The Nature of Credit Constraints and Human Capital

by

Lance J. Lochner and Alexander Monge-Naranjo

Working Paper # 2010-1 **February 2010**

CIBC Working Paper Series

Department of Economics
Social Science Centre
The University of Western Ontario
London, Ontario, N6A 5C2
Canada

This working paper is available as a downloadable pdf file on our website
http://economics.uwo.ca/centres/cibc/

The Nature of Credit Constraints and Human Capital *

Lance J. Lochner
University of Western Ontario and NBER

Alexander Monge-Naranjo
Pennsylvania State University

February 10, 2010

Abstract

We develop a human capital model with borrowing constraints explicitly derived from government student loan programs and private lending under limited commitment. Two key implications of our analysis are: (i) binding constraints may not depress investment; and (ii) a positive relationship between investment and ability is unlikely to arise in standard exogenous constraint models but arises more generally in our framework. Our model also helps explain a number of important empirical observations in the U.S. higher education sector since the early 1980s: (i) a strong and stable positive correlation between ability and college attendance for all income and wealth backgrounds; (ii) the rising importance of family income as a determinant of college attendance; (iii) the increase in the share of undergraduates borrowing the maximum from government student loan programs; and (iv) the dramatic rise in student borrowing from private lenders. In our framework, all of these are natural responses to the rising costs and returns to college (with stable real government loan limits) observed in recent decades. In contrast, the standard exogenous constraint model cannot simultaneously explain observations (i) and (ii) under standard assumptions about preferences; it is also silent on the rise in private lending. Finally, by incorporating both public and private lending, our framework offers new insights regarding the interaction of government and private student lending as well as the responsiveness of private student credit to economic and policy changes.

*For their comments, we thank Pedro Carneiro, Martin Gervais, Tom Holmes, Igor Livshits, Jim MacGee, Victor Rios-Rull, participants at the 2008 Conference on Structural Models of the Labour Market and Policy Analysis, and seminar participants at the University of British Columbia, Univeristy of Carlos III de Madrid, Indiana University, University of Minnesota, Simon Fraser University, University of Western Ontario, and University of Wisconsin. Lochner acknowledges financial research support from the Social Sciences and Humanities Research Council of Canada. Monge-Naranjo acknowledges financial support from the National Science Foundation.

1 Introduction

The human capital literature has consistently appealed to credit constraints to explain differences in schooling investments by family income.[1] The literature, however, has paid little attention to the nature of those constraints, i.e. the underlying institutions and incentive problems associated with providing credit to young individuals with little collateral to pledge while in school.[2] Instead, nearly all studies of human capital with imperfect credit markets assume that individuals face an arbitrary exogenous limit on borrowing (or arbitrary differences in interest rates based on family income).[3] Such simple assumptions are at odds with the actual operation of public and private sources of credit for education. Existing government student loan (GSL) programs explicitly link credit to educational expenditures, while private lenders extend credit to students based on their prospects of repayment and projected future earnings.

In this paper, we model schooling investment decisions when access to credit is determined by key features of public and private lending. We incorporate the fact that credit from GSL programs is restricted by a pre-specified upper limit *and* a 'tied-to-investment' constraint that does not allow borrowing to exceed educational expenditures (net of any grants or subsidies). To model private lending, we build on recent work on credit constraints that arise endogenously when lenders have limited mechanisms for enforcing repayment, a natural assumption given the inalienability of human capital.[4] Under standard enforcement mechanisms, the costs of default are higher for individuals with greater earnings capacity. As a result, private lenders are willing to extend more credit to individuals that invest more in their skills and/or exhibit higher ability. Altogether, access to credit in our model is explicitly linked to schooling expenditures and observable factors that affect the expected future earnings of borrowers.

Relative to a standard model with exogenous constraints, our framework provides new and empirically relevant insights on the effects of ability and family resources on schooling decisions. Incorporating the endogenous nature of credit restrictions fundamentally changes the intertemporal trade-off that defines human capital investments. Ability has stronger positive effects on education investments, and schooling is more sensitive to government policies. Calibrating our

[1] Becker's seminal Woytinsky Lecture (1967) provides an important early theoretical treatment of human capital investment when borrowing opportunities are limited. Hansen and Weisbrod (1969) represents an early empirical analysis of educational attainment gaps by family income; although, their primary focus is on the redistributive effects of education subsidies. Manski and Wise (1983) emphasize borrowing constraints specifically as an explanation for their estimated family income – schooling gaps.

[2] Notable recent exceptions include Andolfatto and Gervais (2006), who focus on optimal intergenerational transfers (in the form of social security and education subsidies) under limited commitment and Ionescu (2008, 2009), who studies default in federal student loan programs.

[3] Studies assuming variable or heterogeneous interest rates include Becker (1975), Cameron and Taber (2004) and Card (1995). Studies assuming a fixed limit on borrowing include Belley and Lochner (2007), Caucutt and Kumar (2003), Cordoba and Ripoll (2009), Hanushek, Leung, and Yilmaz (2003), and Keane and Wolpin (2001).

[4] The literature on endogenous credit constraints has mostly focused on risk-sharing and asset prices in endowment economies (e.g. Alvarez and Jermann 2000, Fernandez-Villaverde and Krueger 2004, Krueger and Perri 2002, Kehoe and Levine 1993, and Kocherlakota 1996) or firm dynamics (e.g. Albuquerque and Hopenhayn 2004, Monge-Naranjo 2009). Our punishments for default are similar to those in Livshits, MacGee, and Tertilt (2007) and Chatterjee, et al. (2007) in their analyses of bankruptcy.

model to the U.S. economy, we show that it can explain the rising importance of family resources for college attendance and the dramatic rise in private student borrowing observed over the past few decades as natural responses to the well-documented increases in the costs of and returns to college.

We highlight two important cross-sectional implications of our model. First, credit constrained individuals do not necessarily under-invest in human capital. When private loans are unavailable, students constrained only by the GSL's 'tied-to-investment' constraint always invest the unconstrained optimal amount. When both public and private loans are available, poor low ability youth may actually *over-invest* in human capital.[5] In either case, investment in human capital is no less than the unrestricted amount and is non-decreasing in family wealth. These results are important in light of the countless empirical studies that are based on the premise (from exogenous constraint models) that borrowing constraints always reduce investment.

Second, our model better predicts the observed ability – college attendance relationship compared to the standard model. We show that under empirically plausible assumptions about the consumption intertemporal elasticity of substitution (IES), exogenous constraint models predict a *negative* relationship between investment and ability among constrained students.[6] This is strongly at odds with empirical studies, which universally indicate that cognitive ability has a strong positive effect on schooling.[7] Our model performs much better. The link between credit and ability and schooling in our model implies a positive education – ability relationship under empirically relevant assumptions.

Our model quantitatively accounts for a number of major changes in college attendance and borrowing patterns in the U.S. between the early 1980s and the early 2000s. In particular, we show that the model is consistent with four facts: (1) a positive ability – education relationship for all family income groups over time (see Section 3); (2) the increased importance of family income as a determinant of college attendance from the early 1980s to the 2000s (Belley and Lochner 2007); (3) a sharp increase in the fraction of undergraduates borrowing the maximum available from GSL programs since the early 1990s (Berkner 2000 and Titus 2002); and (4) a sharp increase in the students borrowing from private lenders since the mid-1990s (College Board 2005). Calibrating our model to the U.S. in the early 1980s, we show that these changes can be explained as equilibrium responses to the observed increase in the returns to and costs of college since the early 1980s (and stable GSL limits).

Our results suggest that in the early 1980s, the GSL provided adequate credit so that few stu-

[5]Over-investment — investing so much that the marginal return is below the (financial) marginal cost — can arise as a way to increase consumption during school. When only the tied-to-investment constraint binds, additional investments (at the margin) can be financed fully from the GSL, and increases in investment expand private credit that can be used to enhance current consumption. While over-investment is theoretically possible, our quantitative analysis indicates that it is not empirically relevant given relatively low current GSL limits.

[6]Specifically, we show that for an $IES \leq 1$ investment is decreasing in ability for constrained borrowers; however, most estimates of the IES are less than one (Browning, Hansen and Heckman 1999).

[7]For example, see Cameron and Heckman (1998), Carneiro and Heckman 2002), Belley and Lochner (2007), as well as evidence provided in Section 3.

dents would have needed to turn to private creditors. College attendance was strongly increasing in ability and largely independent of family resources. The rising college costs and returns over time have encouraged more recent students to invest and borrow more, with many exhausting their GSL loans and borrowing substantially from private lenders. Although private lenders have responded to increases in schooling by offering more credit, our results suggest that many students with low family resources are now constrained and unable to invest as much as they would like. Our simulations imply a weaker ability – investment relationship for constrained youth than ability – college attendance patterns in the 1997 Cohort of the National Longitudinal Survey of Youth (NLSY97) would suggest. However, our model performs noticeably better than the current alternative in this regard: the exogenous constraint model predicts a *negative* ability – schooling relationship at the bottom of the family income/wealth distribution where youth are most constrained. Additionally, the exogenous constraint model offers little insight regarding the significant changes in the composition of student borrowing; although, it generically predicts an increase in total borrowing and in the fraction constrained.

We analyze a number of policy issues that cannot be studied without explicitly endogeneizing access to credit. For instance, our framework lends itself naturally to an analysis of the interaction between GSL and private credit as well as the response of private credit to government policies. We simulate the effects of changes in the GSL program (upper loan limits and repayment schedules) and to private loan enforcement regulations. Most interestingly, we show that expansions of public credit do not fully crowd out private lending: increases in GSL limits lead to higher levels of total credit and raise human capital investment among youth constrained by those limits. Additionally, we show that changes in GSL credit tend to have a relatively greater impact on investment among the least able, while changes in private loan enforcement tend to impact investment more among the most able. Clearly, not all credit expansions are equivalent.

Finally, endogenous borrowing constraints make human capital investment more sensitive to government education subsidies. Any policy that encourages investment is met with an increase in access to credit, which further encourages the investment of constrained students. This 'credit expansion effect', absent with fixed constraints, can be quite large. In our quantitative analysis, investment responds as much as 50% more than in the exogenous constraint model.

The paper is organized as follows. Section 2 characterizes the existing sources of credit in the U.S. Section 3 reports U.S. evidence on the relationship between ability, family income, and college attendance. Section 4 uses a general two-period model to analytically compare the cross-sectional implications for borrowing and investment under alternative assumptions about credit markets. Section 5 extends our framework to a multi-period life-cycle and presents our calibration and baseline quantitative analysis. Section 6 simulates the effects of increased returns and costs to college as observed from the early 1980s to early 2000s. This section also contains a number of policy experiments. Section 7 concludes.

2 Available Sources of Credit

In this section, we briefly review the main sources of credit in the U.S. for college education, including government and private student loans.

2.1 Government Student Loan Programs

Federal GSL programs are an important source of finance for higher education in the U.S., accounting for 71% of the federal student aid disbursed in 2003-04. The largest program is the Stafford Loan program, which awarded nearly $50 billion to students in the 2003-04 academic year.[8] A second program, the Parent Loans for Undergraduate Students (PLUS), awarded $7 billion to parents of undergraduate students during the same period. Also, on a much smaller scale, the Perkins Loan program disbursed $1.6 billion to a small fraction of students from very low-income families.[9]

GSL programs generally have three important features. First, lending is directly tied to investment. Students (or parents) can only borrow up to the total cost of college (including tuition, room, board, books, supplies, transportation, computers, and other expenses directly related to schooling) less any other financial aid they receive in the form of grants or scholarships. Thus, GSL programs do not finance non-schooling consumption expenses. Second, GSL programs set cumulative and annual upper loan limits on the total amount of credit available for each student.[10] Third, loans covered by GSL programs typically have extended enforcement rules compared to unsecured private loans.

Table 1 reports loan limits (based on the dependency status and class of the student) for Stafford and Perkins programs for the period 1993-2007. Dependent students could borrow up to $23,000 from the Stafford Loan Program over the course of their undergraduate careers. Independent students could borrow roughly twice that amount, although most traditional undergraduates do not fall into this category. Qualified undergraduates from low income families could receive as much as $20,000 in Perkins loans, depending on their need and post-secondary institution. It is important to note, however, that amounts offered through this program have typically been less than mandated limits.[11] Student borrowers can defer loan re-payments until six (Stafford) to nine (Perkins) months after leaving school.

[8] The Stafford program offers both subsidized and unsubsidized loans, with the latter available to all students and the former only to students demonstrating financial need. The government waives the interest on subsidized loans while students are enrolled; it does not do so for unsubsidized loans. Prior to the introduction of unsubsidized Stafford Loans in the early 1990s, Supplemental Loans to Students (SLS) were an alternative source of unsubsidized federal loans for independent students.

[9] See The College Board (2006) for details about financial aid disbursements and their trends over time.

[10] Since 1993-94, the PLUS loan program no longer has a fixed maximum borrowing limit; however, parents still cannot borrow more than the total cost of college net of other financial aid.

[11] Parents that do not have an adverse credit rating can borrow up to the cost of schooling from the PLUS program, with repayment typically beginning within 60 days of loan disbursement. Dependent students whose parents do not qualify for PLUS loans (due to a bad credit rating) are able to borrow up to the independent student loan limits.

Table 1: Borrowing Limits for Stafford and Perkins Student Loan Programs (1993-2007)

	Stafford Loans		
	Dependent Students	Independent Students*	Perkins Loans
Eligibility Requirements	Subsidized: Financial Need Unsubsidized: All Students		Financial Need
Undergraduate Limits:			
First Year	$2,625	$6,625	$4,000
Second Year	$3,500	$7,500	$4,000
Third-Fifth Years	$4,000	$8,000	$4,000
Cum. Total	$23,000	$46,000	$20,000
Graduate Limits:			
Annual	$18,500		$6,000
Cum. Total**	$138,500		$40,000

Notes:
* Students whose parents do not qualify for PLUS loans can borrow up to independent student limits from Stafford program.
** Cumulative graduate loan limits include loans from undergraduate loans.

Cumulative Stafford Loan limits, in real terms, were nearly identical in 2002-03 to what they were twenty years earlier. (We focus on these years since the individuals we study below from the 1979 and 1997 Cohorts of the National Longitudinal Surveys of Youth, NLSY79 and NLSY97 respectively, made their college attendance decisions around these two periods.) While the government nominally increased loan limits (especially for upper-year college students) in 1986-87 and 1993-94, inflation has otherwise eroded these limits away.[12] The relative stability of real GSL limits combined with a near doubling of tuition costs in recent decades (College Board 2005), has pushed recent student borrowing up against these upper limits for many undergraduates. Indeed, the fraction of all undergraduate borrowers that borrowed the maximum limit from the federal Stafford Student Loan Program nearly tripled from only 18% in 1989-90 to 52% in 1999-2000. Among traditional dependent undergraduates, the fraction increases to nearly 70% of all borrowers in 1999-2000 (Berkner 2000 and Titus 2002).

An important aspect of GSL loans is that they are more strictly enforced relative to typical unsecured private loans. Except in very special circumstances, these loans cannot be expunged through bankruptcy. If a suitable re-payment plan is not agreed upon with the lender once a

[12]From 1982-83 to 2002-03, Stafford borrowing limits for undergraduates declined by 44% for first-year students and 25% for second-year students, while they increased by about 20% for college students enrolled in years three through five. For most of this period, loan limits for independent undergraduates remained about twice the amounts available to dependent students. Stafford loan limits for graduate students declined by about 35% in real terms from 1986–87 to 2006–07, roughly the time NLSY97 respondents would have began attending graduate school.

borrower enters default, the default status will be reported to credit bureaus and collection costs (up to 25% of the balance due) may be added to the amount outstanding.[13] Up to 15% of the borrower's wages can also be garnisheed. Moreover, federal tax refunds can be seized and applied toward any outstanding balance. Other sanctions include a possible hold on college transcripts, ineligibility for further federal student loans, and ineligibility for future deferments or forbearances.

2.2 Private Lending

Historically, private financing of higher education has been relatively uncommon and mostly restricted to students at elite institutions or those enrolled in professional schools (e.g. business, law and medicine), whose post-graduation earnings (and ability to repay) are expected to be high. As late as the mid-1990s, few private lenders offered loans to students outside the GSL programs (e.g. in 1995-96, total non-federal student loans amounted to only $1.3 billion).

But, much has changed since then. By 2004-05, the amount of student borrowing from private lenders had risen to almost $14 billion, which was nearly 20% of all student loan dollars distributed. Even if private loans are most prevalent among graduate students (especially in professional schools) and undergraduates at high-cost private universities (Wegmann, Cunningham and Merisotis 2003), the rise in borrowing from private student lenders outside the Stafford and Perkins Loan Programs indicates that the GSL limits are no longer enough to satisfy many students' demands for credit.[14] The rising importance of private lending is even more pronounced than these figures suggest, since they do not include student borrowing on credit cards, which also increased considerably over this period (see College Board 2005).

As we show below, the design of private lending programs is broadly consistent with the problem of lending under limited repayment incentives. Private lenders directly link credit to educational investment expenditures and indirectly to projected earnings. All private student loan programs require evidence of post-secondary school enrollment, offering students credit far in excess of what is otherwise offered in the form of more traditional uncollateralized loans. While many private student lending programs are loosely structured like federal GSL programs, they vary substantially in their terms and eligibility requirements. Most notably, some private lenders clearly advertise that they consider the school attended, course of study, and the grades of students in determining loan packages.[15] Finally, private lenders seem to react quickly to changes in economic conditions that affect the broader credit market and the ability of students to meet their future repayment obligations. The New York Times (Glater 2008) reports that in response to the recent credit crisis

[13] Formally, a borrower is considered to be in default once a payment is 270 days late.

[14] Private student loans generally charge higher interest rates than Stafford or Perkins loans and are, therefore, typically taken after exhausting available credit from GSL programs.

[15] For example, MyRichUncle states on its website (www.myrichuncle.com) that it "believes that success in school is indicative of your willingness and ability to repay your loans...taking into account your GPA, school, and course of study." The financial aid help website Finaid.org discusses the growing practice of peer-to-peer student lending, which lets students "...provide some background information on why they need the money. Often this information is structured, providing information about the degree program, year in school, name of the college and GPA."

in the U.S., a number of private lenders discontinued lending to students at community colleges and lower quality four-year institutions, while they continued to lend to students at higher quality schools where graduates were expected to earn more after school.

Until very recently, enforcement of private student loans was regulated by U.S. bankruptcy code. In filing for bankruptcy under Chapter 7, former students could discharge all private student loan obligations after leaving school.[16] Court and filing fees amounting to as much as a couple thousand dollars must also be paid. Other less explicit costs associated with bankruptcy filing are also likely to be important. For example, bankruptcy shows up on an individual's credit report for ten years, limiting future access to credit. Bankruptcy may spill over into other domains as well (e.g. banks, mortgage companies, landlords, and employers often request credit reports from potential customers or employees). Finally, U.S. bankruptcy requires "good faith" attempts to meet debt obligations, which may make it difficult for former students to expunge their debts if current income levels are high. Livshits, MacGee, and Tertilt (2007) argue that punishments associated with Chapter 7 bankruptcy are well-approximated by a temporary period of both wage garnishments and exclusion from credit markets. We follow their approach below.

3 Ability, Family Resources, and College Attendance

The empirical literature on borrowing constraints and human capital has largely focused on the relationship between family income and college attendance. Recent studies also consider the importance of cognitive ability in determining schooling outcomes. In this section, we discuss the empirical relationship between family income, cognitive ability, and college attendance in the U.S. during the early 1980s and in the early 2000s. We document two important facts using data from the NLSY79 and NLSY97. First, in both the early 1980s and the early 2000s, there is a strong positive relationship between college attendance and cognitive ability or achievement (as measured by scores on the Armed Forces Qualifying Test, AFQT) for youth from all levels of family income and wealth.[17] Second, for recent student cohorts, there is a much stronger relationship between family income (or wealth) and college attendance. Indeed, in the early 1980s, there was a only weak link between family income and college-going.

Using data for the 1980s (NLSY79), a number of empirical studies have found that family income played little role in college attendance decisions. Cameron and Heckman (1998, 1999) find that after controlling for family background, AFQT scores, and unobserved heterogeneity, family income has little effect on college enrollment rates. Carneiro and Heckman (2002) also estimate

[16] More generally, borrowers filing under Chapter 7 must surrender any non-collateralized assets (above an exemption) in exchange for discharging all debts; however, most school-leavers considering bankruptcy have few if any assets. Since the 'Bankruptcy Abuse Prevention and Consumer Protection Act of 2005', individuals can no longer discharge student loans, public or private, through bankruptcy.

[17] AFQT scores are widely used as measures of cognitive achievement by social scientists and are strongly correlated with post-school earnings conditional on educational attainment. See, e.g., Cawley, *et al.* (2000). Appendix A provides additional details on the AFQT.

small differences in college enrollment rates and other college-going outcomes by family income after accounting for differences in family background and AFQT. Cameron and Taber (2004) and Keane and Wolpin (2001) explore different features of the NLSY79 data and also argue that credit constraints had little effect on educational outcomes in the early 1980s.

Using data for the late 1990s and early 2000s (NLSY97), Belley and Lochner (2007) show that family income has become a much more important determinant of college attendance.[18] Youth from high income families in the NLSY97 are 16 percentage points more likely to attend college than are youth from low income families, conditional on AFQT scores, family composition, parental age and education, race/ethnicity, and urban/rural residence. This is roughly twice the effect observed in the NLSY79. The NLSY79 do not contain data on wealth; however, the combined effect of family income and wealth in the NLSY97 are substantially greater than the effects of income alone. Comparing youth from the highest family income and wealth quartiles to those from the lowest quartiles yields an estimated difference in college attendance rates of nearly 30 percentage points after controlling for ability and family background.

Despite changes in the relationship between family resources and college attendance, the relationship between ability and schooling has remained strong over time. Figure 1 shows college attendance rates by AFQT quartiles and either family income or family wealth quartiles in the NLSY79 and NLSY97.[19] For all family resource levels in both NLSY samples, we observe substantial increases in college attendance with AFQT. The differences in attendance rates between the highest and lowest ability quartiles range from 47% to 68% depending on the family income or wealth quartile. The figure reveals an equally strong positive ability – college attendance relationship for youth from low and high income/wealth families. In the NLSY97 data, the college attendance gap between the highest and lowest ability quartiles from both the lowest family income and wealth quartiles is 47%, compared to a 37% gap for those from both the highest family income and wealth quartiles.[20]

Of course, AFQT scores may be correlated with other family background variables that influence college attendance decisions conditional on family resources. In Lochner and Monge-Naranjo (2008), we use the NLSY79 and NLSY97 to estimate the effects of AFQT on college attendance by family income or wealth quartile after controlling for gender, race/ethnicity, mother's education,

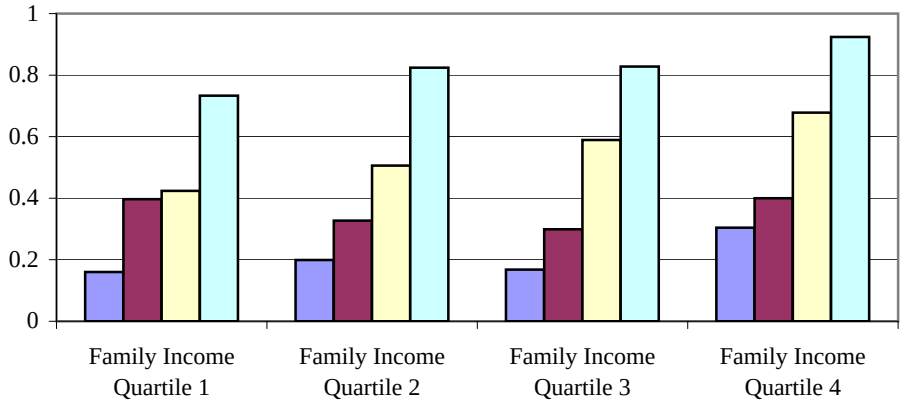
[18] Ellwood and Kane (2000) argue that college attendance differences by family income were already becoming more important by the early 1990s. Using data on youth of college-ages in the 1970s, 1980s, and 1990s (from the Health and Retirement Survey), Brown, Seshadri, and Scholz (2007) estimate that borrowing constraints limit college-going; however, they do not examine whether constraints have become more limiting in recent years. While Stinebrickner and Stinebrickner (2007) find little effect of borrowing constraints (defined by the self-reported desire to borrow more for school) on overall college dropout rates for a recent cohort of students at Berea College, they find substantial differences in dropout rates between those who are constrained and those who are not. They do not study the effects of borrowing constraints on attendance.

[19] See Appendix A for a detailed description of the data and variables used here.

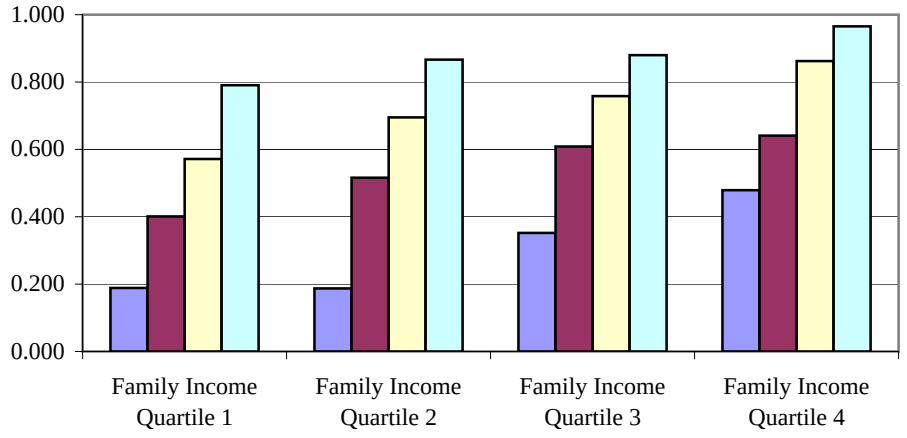
[20] We observe similar patterns in the NLSY97 for age 20 enrollment in four-year colleges/universities conditional on attendance at any post-secondary institution. Among youth from the lowest wealth quartile, the enrollment rate in four-year schools (conditional on post-secondary enrollment) is 34% higher for the most able relative to the least able. Among the highest wealth quartile, the difference is 32%. For the lowest family income quartile, the same high - low ability gap is 41%, while it is 52% for the highest income quartile.

Figure 1: College Attendance by AFQT and Family Income or Wealth (NLSY79 and NLSY97)

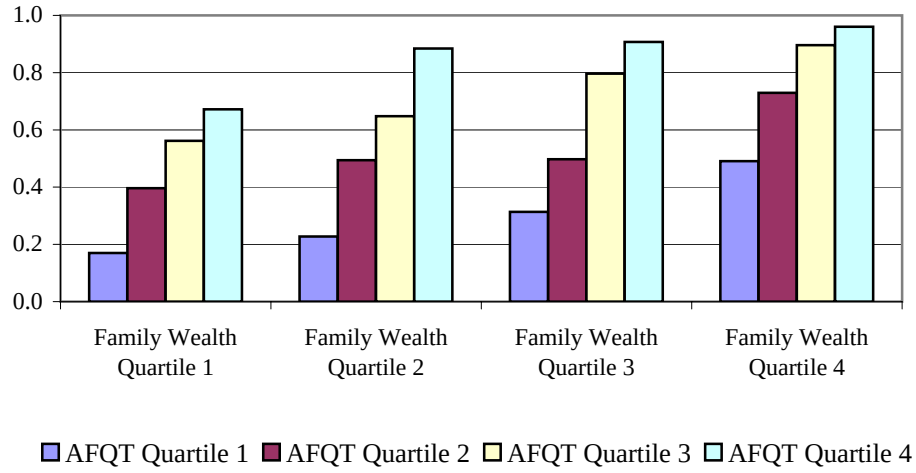
(a) Attendance by AFQT and Family Income (NLSY79)



(b) Attendance by AFQT and Family Income (NLSY97)



(c) Attendance by AFQT and Family Wealth (NLSY97)



intact family during adolescence, number of siblings/children under age 18, mother's age at child's birth, urban/metropolitan area of residence during adolescence, and year of birth. These estimates confirm the general patterns observed in Figure 1: Cognitive ability has strong positive effects on college attendance for all family income and wealth quartiles in both NLSY samples. Below, we demonstrate that these findings are strongly at odds with the predictions of a standard exogenous constraints model.

4 Modeling Student Credit

As discussed in Section 2, both public and private student lenders behave quite differently from the standard assumption that credit is limited by a single invariant upper limit. Most importantly, they link credit to the investment and ability of the borrower. In this section, we use a two-period model to show how this link shapes the behavior of human capital investment. Incorporating key features of public and private lending yields predictions about cross-sectional investment patterns that are qualitatively consistent with the empirical patterns discussed above, while the standard model of exogenous borrowing constraints does not.

4.1 Preferences and Human Capital Production Technology

Consider two-period-lived individuals who invest in schooling in the first period and work in the second. Their preferences are

$$U = u(c_0) + \beta u(c_1), \tag{1}$$

where c_t is consumption in periods $t \in \{0, 1\}$, $\beta > 0$ is a discount factor, and $u(\cdot)$ is the period utility function. We assume $u(\cdot)$ is strictly increasing, strictly concave, twice continuously differentiable and satisfies $\lim_{c \searrow 0} u'(c) = +\infty$.

Each individual is endowed with financial assets $w \geq 0$ and ability $a > 0$. Initial assets capture all familial transfers while ability reflects innate factors, early parental investments and other characteristics that shape the returns to investing in schooling. We take (w, a) as fixed and exogenous to focus on schooling decisions that individuals make largely on their own; however, our central results generalize naturally to an intergenerational environment in which parents endogenously make transfers to their children.[21]

Labor earnings at $t = 1$ are equal to $af(h)$, where h is schooling investment and $f(\cdot)$ is a positive, strictly increasing, strictly concave, twice continuously differentiable function that satisfies $\lim_{h \searrow 0} f'(h) = +\infty$ and $\lim_{h \nearrow \infty} f'(h) = 0$. Note that both a and h increase earnings

[21] In an online appendix, we derive equivalent analytical results in three common models of parental transfers: (i) an 'altruistic' model (i.e. parents directly value the utility of their children); (ii) 'warm glow' preferences (i.e. parents directly value the resources transferred to their children); and (iii) a 'paternalistic' model (i.e. parents directly value the human capital investment of their children). In the last model, we need to impose a few additional mild conditions.

and are complementary with each other.[22]

Human capital investment, h, is in units of the consumption good.[23] Individuals can borrow d of these units (or save, in which case $d < 0$) at a gross interest rate $R > 1$. Given w, a, h and d, consumption in each of the periods is

$$c_0 = w + d - h, \tag{2}$$

$$c_1 = af(h) - Rd. \tag{3}$$

4.2 Unrestricted Allocations

Young individuals maximize utility (1) subject to (2) and (3). In the absence of financial frictions, this maximization can be separated into two steps. The first is to choose human capital investment h to maximize the present value of net lifetime income, $-h+af(h)/R$, which leads to the condition

$$af'\left[h^U(a)\right] = R. \tag{4}$$

Optimal unrestricted investment, $h^U(a)$, equates the marginal return on human capital with the return on financial assets and is, therefore, strictly increasing in ability, a, and independent of initial assets, w.

The second step is to smooth consumption, i.e. borrow $d^U(a,w)$ units to satisfy the Euler equation:

$$u'\left(w + d^U(a,w) - h^U(a)\right) = \beta R u'\left(af\left[h^U(a)\right] - Rd^U(a,w)\right). \tag{5}$$

Unconstrained borrowing is strictly decreasing in wealth and increasing in ability. Ability increases borrowing for two different reasons: (i) more able individuals wish to finance a larger investment and (ii) for any given level of investment, more able individuals earn higher net lifetime income and wish to consume more in the first period. Because of (ii), unrestricted borrowing increases more steeply in ability than does unrestricted human capital investment. The following lemma formalizes this result and is used below to determine who is credit constrained.

Lemma 1 *$h^U(a)$ is strictly increasing in a, and $d^U(a,w)$ is strictly increasing in a and strictly decreasing in w. Moreover, $\frac{\partial d^U(a,w)}{\partial a} > \frac{dh^U(a)}{\partial a} > 0$ and $-1 < \frac{\partial d^U(a,w)}{\partial w} < 0$.*

See Appendix B for all proofs and other analytical details related to this section.

[22] We implicitly assume a constant elasticity of substitution between ability and investment equal to one. This specification is consistent with most empirical studies, which generally incorporate ability in the intercept of log wage/earnings regressions and with standard theoretical models of human capital (e.g. the widely used Ben-Porath (1967) model). In an online appendix, we generalize a few key results below to the more general case of a CES production function in both ability and human capital, $\tilde{f}(a,h)$.

[23] Our model is isomorphic to one in which foregone earnings for any given investment amount, h, are independent of ability. In an online appendix, we extend our model to allow the cost of investment to depend generally on ability. We show that our main conclusions here hold under fairly general and empirically relevant assumptions.

4.3 Exogenous Borrowing Constraints

Credit constraints are typically introduced in models of human capital by imposing a fixed and exogenous upper bound on the amount of debt.[24] Following this approach, assume that borrowing is restricted by the exogenous constraint:

$$d \leq d_0, \tag{EXC}$$

where $0 \leq d_0 < \infty$ is fixed and uniform for all agents. We use the superscript X for allocations in this model.

For each ability a, a threshold level of assets $w^X_{\min}(a)$ defines who is constrained ($w < w^X_{\min}(a)$) and who is unconstrained ($w \geq w^X_{\min}(a)$). Constrained persons have high ability relative to their wealth since $w^X_{\min}(a)$ is increasing in ability (see Appendix B). Individuals constrained by (EXC) have exhausted their possibilities to bring future resources to the early (investment) period. Their human capital investment $h^X(a, w)$ must strike a balance between increasing lifetime earnings and smoothing consumption and is uniquely determined by

$$u'\left(w + d_0 - h^X(a, w)\right) = \beta u'\left(af\left[h^X(a, w)\right] - Rd_0 af'\left[h^X(a, w)\right]\right),$$

the equality between the marginal cost of investing (reducing current consumption) and its marginal benefit (net return in terms of future consumption).

The next proposition highlights four empirically relevant implications of this model. Most importantly, the implied relationship between constrained investment and ability in part (iv) depends on the consumption intertemporal elasticity of substitution (IES), $-u'(c) / [cu''(c)]$.

Proposition 1 *Consider individuals with wealth $w < w^X_{\min}(a)$, so (EXC) binds. Then: (i) $h^X(a, w) < h^U(a)$; (ii) $h^X(a, w)$ is strictly increasing in w; (iii) the marginal return on human capital investment, $af'\left[h^X(a, w)\right]$, is strictly greater than R and strictly decreasing in w; and (iv) if the $IES \leq 1$, then $h^X(a, w)$ is strictly decreasing in ability, a.*

Results (i)-(iii) are well-known (Becker 1975) and central to the empirical literature on credit constraints. For instance, Cameron and Heckman (1998, 1999), Ellwood and Kane (2000), Carneiro and Heckman (2002), and Belley and Lochner (2007) empirically examine if youth from lower income families acquire less schooling conditional on family background and ability (results (i) and (ii)). Lang (1993), Card (1995), and Cameron and Taber (2004) explore the prediction that the marginal return on human capital investment exceeds the return on financial assets (result (iii)).

The most interesting result is part (iv). The relationship between ability and investment for constrained individuals is determined by the balance of two opposing forces. On the one

[24] See, for example, Aiyagari, Greenwood, and Seshadri (2002), Belley and Lochner (2007), Caucutt and Kumar (2003), Cordoba and Ripoll (2009), Hanushek, Leung, and Yilmaz (2003), and Keane and Wolpin (2001). Instead, Becker (1975) assumes that individuals face an increasing interest rate schedule as a function of their investment. Becker's formulation yields similar predictions to those discussed here.

hand, there is an intertemporal substitution effect: more able individuals earn a higher return on human capital investment, so they would like to invest more. On the other hand, there is a wealth effect: more able individuals have higher lifetime earnings, which increases their desired consumption at all ages. Since constrained borrowers can only increase consumption during the initial period by investing less, the wealth effect discourages investment. With strong preferences for intertemporal consumption smoothing (i.e. IES$\leq$1), the wealth effect dominates and a negative ability – investment relationship arises.

The prediction of a *negative* relationship between ability and investment for an IES ≤ 1 is a serious shortcoming of the model.[25] Most estimates of the IES are less than one (see Browning, Hansen, Heckman 1999) and as discussed earlier, schooling is strongly increasing in ability even for youth from low-income families. Result (iv) not only implies perverse cross-sectional investment patterns, but it also implies that an increase in the market price of human capital can lead to aggregate reductions in investment among constrained individuals. This is because a change in the price of skill is analogous to increasing the ability of everyone in the economy.

4.4 Government Student Loan Programs

In this subsection, we consider GSL programs as the only source of credit. We then introduce private lending in the following subsection.

As described in Section 2, GSL programs possess three key features. First, lending is *tied to investment* and cannot be used to finance non-schooling related consumption goods or activities:

$$d \leq h. \tag{TIC}$$

This condition is equivalent to $c_0 \leq w$. Second, borrowing is constrained by a fixed upper limit $0 < d_{\max} < \infty$, so

$$d \leq d_{\max}. \tag{6}$$

Combining these two constraints yields actual credit limits from GSL programs:

$$d \leq \min\{h, d_{\max}\}. \tag{GSLC}$$

Third, the government has enhanced enforcement mechanisms to ensure repayment. To capture this feature, we assume that government loans are fully enforceable (an assumption implicit also in the previous model.)

To isolate the role of (TIC), first assume that it is the only constraint.[26] In this case, individuals are unconstrained as long as desired borrowing is less than or equal to desired investment. Because desired investment is increasing in ability, the (TIC) constraint is less stringent than

[25] An $IES \leq 1$ is only a sufficient condition. We further show in the online appendix that the result is even stronger if investment is in terms of foregone earnings that increase with ability.

[26] This would be the case if upper borrowing limits were non-existent or set very high (e.g. PLUS program for students' parents).

(EXC) for higher ability individuals but more stringent for those with low ability. When $d = h$, borrowing can finance investment (but no more), and early consumption equals initial wealth. The individual's problem reduces to choosing investment h to maximize $\{u(w) + \beta u[af(h) - Rh]\}$, which is equivalent to maximizing discounted net lifetime earnings. Therefore, optimal investment equals $h^U(a)$.

By itself, (TIC) does not lead to a conflict between smoothing consumption and maximizing net lifetime resources, because credit cannot be used for anything other than investment. Despite potentially large distortions in the intertemporal allocation of consumption, if (TIC) were the only constraint on borrowing, everyone would invest the unconstrained amount $h^U(a)$ regardless of ability and initial wealth. Empirical tests based on investment differences by family resources would fail to capture consumption distortions and would always conclude that borrowing constraints are non-binding or non-existent. *A simple lesson is that robust empirical tests for binding credit constraints should include information on the behavior of consumption over time.*

Now, consider the full GSL constraint (GSLC), denoting allocations in this model by the superscript G. To facilitate the exposition, we assume (throughout this section) that $d_{\max} = d_0$. Unconstrained individuals ($w \geq w^G_{\min}(a)$) possess relatively high assets relative to their ability.[27] The remaining population of constrained individuals falls into three categories: First, a low ability group is comprised of individuals constrained only by (TIC) and not by the maximum $d_{\max}$. They invest the unrestricted level $h^U(a)$ but would like to borrow to increase consumption while in school. Second, a more able group consists of individuals who borrow up to the maximum $d_{\max}$ and invest beyond that using some of their own available resources. For them, investment coincides with $h^X(a, w)$. A third group might emerge if $h^X(a, w)$ is decreasing in a (i.e. $IES \leq 1$.) This third group would be composed of very high ability youth who are constrained by both (6) and (TIC). We formalize this discussion as follows:

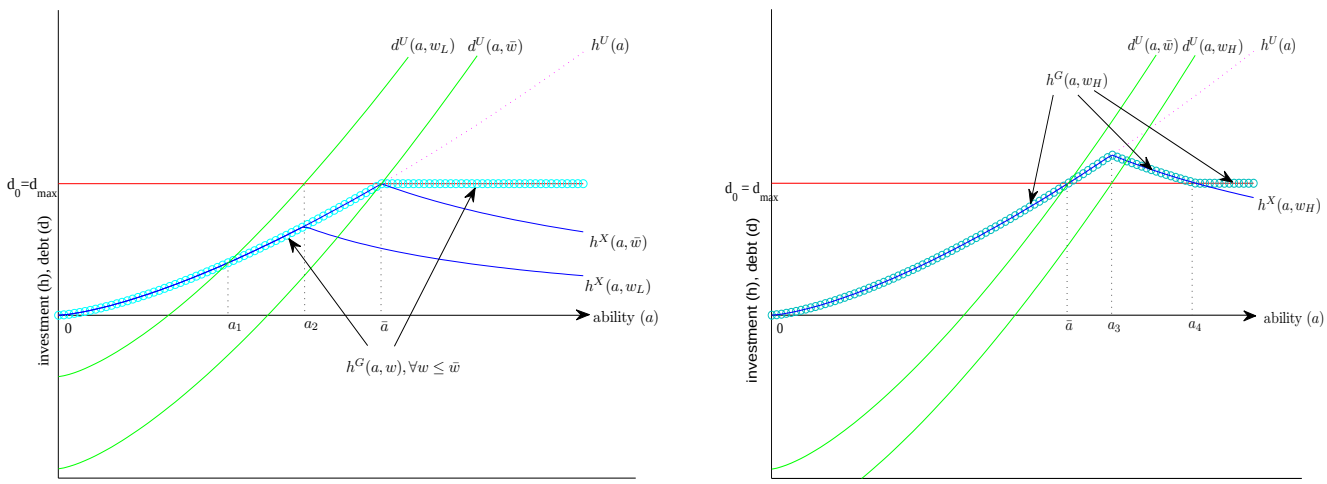
Proposition 2 *Assume that $u(\cdot)$ has IES ≤ 1. Let $d_{\max} = d_0 > 0$; let $\bar{a} > 0$ be defined by $h^U(\bar{a}) = d_{\max}$; and let $\hat{w} : [\bar{a}, \infty) \to R_+$ be defined by $h^X[a, \hat{w}(a)] = d_{\max}$, the (possibly infinite) wealth level that leads an exogenously constrained individual with ability a to invest $d_{\max}$. Then:*

$$h^G(a, w) = \begin{cases} h^U(a) & a \leq \bar{a} \text{ or } w \geq w^X_{\min}(a) \\ h^X(a, w) & a > \bar{a} \text{ and } w < \hat{w}(a) \\ d_{\max} & \text{otherwise.} \end{cases}$$

Figures 2(a) and (b) illustrate the behavior of $h^G(a, w)$, $h^X(a, w)$, and $h^U(a)$ for the empirically relevant case of IES ≤ 1. These figures also display unconstrained borrowing as a function of ability for different levels of wealth. Figure 2(a) displays investment and borrowing behavior for two low levels of wealth, $\bar{w}$ and $w_L < \bar{w}$, while Figure 2(b) illustrates investment behavior for a higher level of wealth $w_H > \bar{w}$.[28]

[27] In Appendix B, we show that the threshold $w^G_{\min}(a)$ is increasing in ability. We also show that when $d_{\max} = d_0$, $w^G_{\min}(a) \geq w^X_{\min}(a)$ and more persons are constrained by the GSL, because it imposes an additional constraint.

[28] Note that $\bar{w} \equiv w^G_{\min}(\bar{a})$ reflects the level of wealth below which agents of ability $\bar{a}$ are constrained, where $\bar{a}$ is the ability level at which unconstrained investment equals the upper limit on borrowing (i.e. $h^U(\bar{a}) = d_{\max}$).



(a) low wealth individuals

(b) high wealth individuals

Figure 2: d^U, h^U, h^X, and h^G for low and high wealth individuals

Because of the 'tied-to-investment' constraint, the implied investment – ability and investment – wealth relationships in the GSL model are more closely aligned with the empirical evidence than the simple exogenous constraint model. First, investment is equal to the unconstrained level $h^U(a)$ and increasing in ability for a larger range of lower ability and low/middle wealth individuals (e.g. individuals with wealth w_L and ability $a \in (a_2, \bar{a}]$ in Figure 2(a)). Second, among high ability individuals (i.e. $a > \bar{a}$), investment never falls below $d_{\max}$; this shrinks the range of abilities for which investment is negatively related to ability (e.g. individuals with ability $a > a_4$ in Figure 2(b)). Third, among high ability types, investment is weakly increasing in initial assets (e.g. individuals with ability $a \in (a_3, a_4)$ in Figure 2(b)).

4.5 GSL Programs and Private Lenders

Our complete model allows the coexistence of both private and public lenders. We assume that private lenders are competitive but face limited repayment incentives from students due to the inalienability of human capital and lack of other forms of collateral. We continue to assume full enforcement of repayment in GSL programs.

A rational borrower repays private loans if and only if the cost of repaying is less than the cost of defaulting. These limited incentives can be foreseen by rational lenders who, in response, limit their supply of credit to what will be re-paid.[29] Since penalties for default are likely to impose a larger monetary cost for borrowers with higher earnings and assets — only so much can be taken from someone with little to take — credit offered to an individual is directly related to his perceived future earnings. Because expected earnings are determined by ability and investment, private credit limits and investments are co-determined in equilibrium.

[29] Gropp, Scholz, and White (1997) empirically support this form of response by private lenders.

In the life-cycle model of Section 5, credit limits arise from temporary exclusion from credit markets and wage garnishments. Here, we derive a similar form of constraint by simply assuming that defaulting borrowers lose a fraction $0 < \tilde{\kappa} < 1$ of labor earnings.[30] In this case, optimal repayment behavior is quite simple: borrowers repay (principal plus interest on private debt d_p) if and only if the payment Rd_p is less than the punishment cost $\tilde{\kappa} af(h)$. As a result, credit from private lenders is limited to a fraction of post-school earnings:

$$d_p \leq \kappa af(h), \tag{7}$$

where $\kappa \equiv R^{-1}\tilde{\kappa}$. Private credit is directly increasing in both ability and investment. Moreover, ability may also indirectly affect credit through its influence on investment.

Students can borrow d_g from the GSL (subject to (GSLC)) and d_p from private lenders (subject to (7)). Because GSL repayments are fully enforced and do not affect incentives to repay private loans, total borrowing is constrained by

$$d_g + d_p \leq \min\{h, d_{\max}\} + \kappa af(h). \tag{8}$$

We use the superscript $G+L$ to highlight that both sources of credit are present. Note that our GSL-only model above is a special case with no private loan enforcement (i.e. $\kappa = 0$). One could similarly define a private lender-only economy setting $d_{\max} = 0$. We use the superscript L to refer to this special case.

The coexistence of both sources of credit reduces the incidence of constrained individuals – relative to economies with only one of these credit sources. The threshold $w_{\min}^{G+L}(a)$ of assets below which individuals are constrained is decreasing in $d_{\max}$ and κ, because increases in either of these parameters represent an expansion of total credit. Expanding either public or private credit would reduce the population of constrained individuals and change the investment behavior of those who remain constrained.

Lemma 2 *Let $h^{G+L}(a, w; d_{\max}, \kappa)$ denote the optimal investment for an individual with ability a and wealth w in an economy with $d_{\max} > 0$ and $\kappa > 0$. Then: (i) $w_{\min}^{G+L}(a) < \min\left\{w_{\min}^{G}(a), w_{\min}^{L}(a)\right\}$; (ii) For constrained individuals with abilities $a > \bar{a}$, the inequalities $\frac{\partial h^{G+L}(a,w;d_{\max},\kappa)}{\partial d_{\max}} > 0$ and $\frac{\partial h^{G+L}(a,w;d_{\max},\kappa)}{\partial \kappa} > 0$ hold.*

The two sources of credit have differential impacts on investment depending on ability. Among highly able youth constrained by the upper GSL limit and private constraints, increasing the GSL limit may increase investment more than one-for-one, since private credit expands with investment. The associated rise in private credit also yields an increase in consumption while in school. An increase in private credit (i.e. a higher κ) would also raise in-school consumption and investment.

[30] This is consistent with wage garnishments and penalty avoidance actions like re-locating, working in the informal economy, borrowing from loan sharks, or renting instead of buying a home, which are all costly to those who default.

Notice that result (ii) in this lemma applies only to higher ability persons with $a > \bar{a}$ (i.e. persons with $h^U(a) > d_{\max}$). Less able individuals are constrained by (TIC) and not by $d_{\max}$, so an expansion of the GSL limit has no affect on their behavior. Moreover, as we discuss below, an increase in κ might actually reduce their investments.

Unlike the models with exogenous or government constraints alone, it is possible that for the same level of familial resources w, a more able person is unconstrained while another with lower ability is constrained. That is, for large enough κ, the threshold $w_{\min}^{G+L}(a)$ may be decreasing in a, since punishment for default may be substantially more costly for the more able/higher earnings person. For the same reason, it is possible that individuals at the top of the ability distribution are always unconstrained (i.e. $w_{\min}^{G+L}(a) < 0$ for high a). These features are driven entirely by the presence of private lenders in the market.

There is an interesting interaction between GSL credit and private lending, depending on which of the GSL constraints binds, (6) or (TIC). Among the more able individuals for whom the upper GSL limit $d_{\max}$ binds, there is under-investment and investment is increasing in wealth (as in the previous models). For individuals in this group, the ability – investment relationship depends on the IES as well as the relative importance of the GSL and private lending. We show that if private lending is a relatively important source of funds, investment is increasing in ability for empirically relevant values of the IES less than one. Among lower ability individuals, for whom (6) is slack but (TIC) binds, investment behavior can be quite different. In the absence of private lenders, these individuals borrow and invest $h^U(a)$ as discussed earlier. With private lenders, constrained individuals actually *over-invest* in human capital (i.e. $h > h^U(a)$ and $af'(h) < R$) if (TIC) is the binding GSL constraint, since on the margin, total credit is increasing more than one-for-one with investment. This is because (i) additional marginal investments can be financed fully by the GSL, and (ii) additional investments raise earnings, which expands access to private credit and allows for greater consumption while in school. Over-investing is socially inefficient and produces a negative relationship between investment and wealth for those individuals. Furthermore, their investment may decline with more access to private credit (i.e. an increase in κ). In any event, in this situation, we show that a positive relationship between ability and investment arises under fairly weak conditions (e.g. a constant IES).

The following proposition summarizes the relationship between investment, ability, and wealth when GSL programs and private lending co-exist. To this end, define $\varrho(a) \equiv \frac{Rd_{\max}}{af(d_{\max})}$ ($\equiv 0$ if $d_{\max} = 0$), the fraction of post-school earnings someone of ability a can borrow from the GSL if they invest $h = d_{\max}$.[31]

Proposition 3 *Assume $d_{\max} > 0$ and $\kappa > 0$ and consider constrained individuals with $w < w_{\min}^{G+L}(a)$, so constraint (8) binds. Then, the following results hold: (1) If $a > \bar{a}$, then: (i) $h^{G+L}(a, w) < h^U(a)$, (ii) $h^{G+L}(a, w)$ is strictly increasing in w, (iii) $h^{G+L}(a, w)$ is strictly in-*

[31] When $a > \bar{a}$, $\varrho(a)$ is less than the elasticity of earnings with respect to human capital investment evaluated at $h = d_{\max}$, i.e. $\varrho(a) = \frac{\bar{a}}{a}\frac{f'(d_{\max})d_{\max}}{f(d_{\max})} < \frac{f'(d_{\max})d_{\max}}{f(d_{\max})}$.

creasing in a if either (a) the IES is bounded below by $\frac{1-\kappa R}{1-\varrho(a)}$ or (b) $\beta R \leq 1$, the IES is non-decreasing in consumption and bounded below by $\frac{1}{1-\varrho(a)}\left(\frac{1-\kappa(R+1)}{1+\kappa(\beta^{-1}-R)}\right)$. (2) If $a < \bar{a}$, then: (i) $h^{G+L}(a,w) > h^{U}(a)$, (ii) $h^{G+L}(a,w)$ is strictly decreasing in w, and (iii) $h^{G+L}(a,w)$ is strictly increasing in a if the IES is constant.

The size of the GSL program has complicated effects on the ability – investment relationship when private lending is also available. On one hand, a larger GSL limit $d_{\max}$ reduces the mass of individuals for which this constraint is binding (i.e. it increases $\bar{a}$). Assuming the weak condition of part (2)(iii) is met, this ensures a positive ability – investment relationship for a broader range of ability levels. On the other hand, a larger value of $d_{\max}$ increases the value of $\varrho(a)$, thus reducing the range of IES that ensures a positive ability – investment relationship for those higher ability individuals constrained by $d_{\max}$.[32]

Increasing private lending (i.e. κ) weakens the conditions in part (1) for a positive ability – investment relationship, allowing for a broader range of IES values. Upon inspection of condition (a) in part (1)(iii), if someone investing $h = d_{\max}$ can borrow more from private lenders than from the GSL program (i.e. $d_{\max} < \kappa a f(d_{\max})$), then there is a positive ability – investment relationship for a range of IES less than one. In general, the bound in (b) is weaker, so under additional mild conditions, a positive ability – investment relationship holds for still lower values of the IES.

5 Quantitative Analysis

We now explore the quantitative implications of our model of public and private lending for schooling. To this end, we extend our two-period model to a multi-period setting which we calibrate using data on college costs, labor earnings, and other features of the U.S. economy. We examine whether the model can reproduce the main empirical patterns reported in Sections 2 and 3. We also consider the effects of potential policy changes.

5.1 A Multiperiod Model

Consider individuals whose post-secondary life is the time interval $[S,T]$ which is divided into three stages: "Youth", $t \in [S,P)$, when individuals invest in school; "maturity," $t \in [P,R)$, when they work full-time; and "retirement," $t \in [R,T]$, when they consume from accumulated savings. S reflects the starting age of college education, P the entry to full-time participation in labor markets, R the start of retirement, and T the age of death.

Preferences are standard. As of any $t_0 \in [S,T]$, a consumption flow $c(t)$ generates utility

$$U(t_0) = \int_{t_0}^{T} e^{-\rho(t-t_0)} \left[\frac{c(t)^{1-\sigma}}{1-\sigma}\right] dt, \tag{9}$$

[32] If only private lending prevails (i.e. $d_{\max} = 0$), then $\varrho(a) = 0$ and only part (1) of Proposition 3 is relevant since $\bar{a} = 0$. In this case, both conditions for a positive ability – investment relationship admit a (potentially large) range of IES below one.

where $\sigma > 0$ is the inverse of the IES and $\rho > 0$ is a subjective discount rate.

We assume competitive financial markets. The market interest rate is $r > 0$, which may differ from ρ. Consumption and investment are restricted by the lifetime budget constraint

$$\int_S^T e^{-r(t-S)} c(t)\, dt + \int_S^P e^{-r(t-S)} x(t)\, dt \leq w + \int_P^R e^{-r(t-S)} y(t)\, dt, \tag{10}$$

where $w \geq 0$ indicates the individual's own financial wealth as of $t = S$.[33]

Individuals are endowed with a minimum human capital $h_0 \geq 0$ but they can invest to increase the human capital with which they enter the labor market at $t = P$.[34] An investment flow $x(t) \geq 0$ during the schooling period accumulates into a stock of human capital investment $h_I \equiv \int_S^P e^{-r(t-S)} x(t)\, dt$ at date P. We assume that the government matches every unit of privately financed investment with a subsidy rate of $s \geq 0$. Total human capital at the time of labor market entry is, therefore,

$$h = h_0 + (1+s)\, h_I. \tag{11}$$

We abstract from the timing of the flow $x(t)$ and focus on accumulated private investment (h_I) and total human capital (h).[35] As a normalization, the units of h, h_I, and h_0 are all in present value terms as of date $t = S$, the beginning of the schooling period.

Labor earnings at date t, $y(t)$, depend positively on individual ability a, schooling h, and experience $E(t-P)$ accumulated since labor market entry at date $t = P$:

$$y(t) = a h^{\alpha} E(t-P), \tag{12}$$

where $0 < \alpha < 1$ and $E(t-P) = \exp\left(\int_P^t g(z) dz\right)$ where $g(z)$ is the rate of growth of earnings at date z. For any two dates $t_0 < t_1$, define $\Phi_{[t_0,t_1]} \equiv \int_{t_0}^{t_1} e^{-r(t-t_0)} E(t-P)\, dt$ as the discounted value of earnings over the period t_0 to t_1 that comes from the accumulation of experience (discounted to date t_0). With this definition, the present value of lifetime labor income (as of date $t = S$) is $e^{-r(P-S)} \Phi_{[P,R]} a h^{\alpha}$.[36]

5.2 Unrestricted Allocations

Frictionless financial markets allow individuals to fully smooth consumption and maximize the present value of lifetime labor earnings net of investments costs, $-h_I + e^{-r(P-S)} \Phi_{[P,R]} a \left[h_0 + (1+s)\, h_I\right]^{\alpha}$.

[33] Since, for simplicity, we assume that human capital is produced from goods rather than time inputs, w is most easily thought of as the present value of family transfers during youth. We could equivalently assume that human capital investment only requires time inputs and that an individual's total 'initial wealth', w, reflects family transfers plus the total discounted value of earnings he could receive if he worked (rather than attended school) full-time during "youth". In this case, private investment costs reflect any earnings foregone for school. Our calibration below implicitly assumes both goods and time investments are perfectly substitutable and combines these costs to determine total investment in human capital.

[34] We have also estimated a version of the model allowing h_0 to depend on ability a. These more general estimates suggest that h_0 is about 25% higher for the top AFQT quartile relative to the bottom quartile; other parameter estimates were very similar to our baseline values. Most importantly, simulation results for the more general model were quite similar to those presented here.

[35] Implicitly, our analysis assumes that investments during youth are perfectly substitutable over time.

[36] Our results readily extend to the case where h_0 and/or $E(t-P)$ are increasing in a.

Individuals with ability $a \leq a_0 \equiv \frac{h_0^{1-\alpha}}{\alpha(1+s)e^{-r(P-S)}\Phi_{[P,R]}}$ do not find it worth investing above what is publicly provided and choose $h_I = 0$. Those with $a > a_0$ equate the marginal return on human capital investment with its private marginal cost of investing. The resulting unconstrained level of total human capital investment is

$$h^U(a) = \max\left\{h_0, \left[\alpha(1+s)ae^{-r(P-S)}\Phi_{[P,R]}\right]^{\frac{1}{1-\alpha}}\right\}, \tag{13}$$

and private investment is $h_I^U(a) = \left[h^U(a) - h_0\right]/(1+s)$. Both $h^U(a)$ and $h_I^U(a)$ are independent of the individual's wealth.

The optimal consumption path grows at the constant rate $\frac{r-\rho}{\sigma}$ over the life-cycle. Define the function $\Theta_{[t_0,t_1]} \equiv \int_{t_0}^{t_1} e^{\left[\frac{r-\rho}{\sigma}-r\right](t-t_0)}dt$ as the cumulation of present value factors along the unconstrained optimal path for consumption between any pair of dates t_0 and t_1. Under the unrestricted allocations, an individual with ability a and initial wealth w enters the labor market with a debt (in present value terms of $t = P$)

$$d^U(a,w) = \Phi_{[P,R]}\left(\frac{\Theta_{[S,P]}}{\Theta_{[S,T]}}\right)a\left[h^U(a)\right]^{\alpha} + e^{r(P-S)}\left(1 - \frac{\Theta_{[S,P]}}{\Theta_{[S,T]}}\right)\left(h_I^U(a) - w\right). \tag{14}$$

This function $d^U(a,w)$ shares the same essential properties as in the two-period model.

5.3 Borrowing Constraints

To introduce limitations in the amount of credit for post-secondary education, it is convenient to first describe utility at the time of labor market entry ($t = P$) for an individual with ability a, human capital h and financial liabilities d. If he fully repays all debts, the present value (as of $t = P$) of his net lifetime resources is $\Phi_{[P,R]}ah^{\alpha} - d$. Optimal consumption smoothing implies a discounted utility

$$V_P^R(a,h,d) = \Theta_{[P,T]}\frac{\left[\left(\Phi_{[P,R]}ah^{\alpha} - d\right)/\Theta_{[P,T]}\right]^{1-\sigma}}{1-\sigma}, \tag{15}$$

where $\Theta_{[t_0,t_1]}$ is defined above. The superscript R indicates full repayment and the subscript P indicates that it is discounted utility as of date $t = P$.

5.3.1 GSL Programs and Private Lending

Young individuals can borrow from GSL programs, d_g, and from private lenders, d_p. Credit from the GSL is tied to schooling-related expenses, subject to a maximum cumulative amount:

$$d_g \leq \min\{h_I, d_{\max}\}, \tag{16}$$

for some $0 < d_{\max} < \infty$. Here, government credit is linked to personal out-of-pocket investment expenses h_I rather than total human capital h. We assume that GSL loans entail a fixed repayment schedule $r(t; d_g)$ over the employment period. The repayment $\int_P^R e^{-r(t-P)}r(t;d_g)\,dt = d_g$ is fully enforced regardless of whether individuals default on private loans.

In addition to d_g, private loans d_p would lead to total borrowing $d = d_g + d_p$ during school. Private lenders restrict student credit due to their limited ability to punish default. We assume that they employ two punishments commonly assumed in the literature on consumer bankruptcy (e.g. Livshits, MacGee, and Tertilt (2007), Chatterjee, et al (2007)). First, defaulting borrowers are reported to credit bureaus, an action that disrupts, at least temporarily, their access to formal credit markets. This penalty inhibits consumption smoothing, which can be quite costly when the IES is low and the earnings profile is steep in experience. Second, defaulting borrowers must forfeit a fraction $\gamma \in [0, 1)$ of their labor earnings. The fraction γ encompasses direct garnishments from lenders and/or the costs of actions taken by borrowers to avoid direct penalties (e.g. working in the informal sector, renting instead of owning a house, etc.). Both penalties last for a period of length $0 \leq \pi < R - P$ that starts the moment default takes place. (The special case of $\pi = 0$ implies that punishments are negligible and private credit is non-existent.)

We make three additional assumptions that greatly simplify the analysis: (i) individuals can only default on private loans at the time of labor market entry; (ii) individuals that choose to repay their private student loans have access to perfect financial markets upon entry into the labor market; and (iii) individuals that default on private loans can access frictionless and fully enforceable credit markets after the punishment period. In short, we abstract from issues related to the optimal timing of default and the enforcement of post-school loans.[37] Assumptions (i) and (ii) help to isolate and focus on the impact of limited access to credit during school. For many parameter values, (iii) is not an assumption but an equilibrium outcome.[38]

Consider an individual that decides to default on private debt d_p at the time of labor market entry. With ability a, human capital h, and GSL liabilities d_g, his attainable utility is:

$$V_P^D(a, h, d_g, r(\cdot; d_g)) = \int_P^{P+\pi} e^{-\rho(t-P)} \frac{\left[(1-\gamma)\, a h^{\alpha} E\,(t-P) - r\left(t; d_g\right)\right]^{1-\sigma}}{1-\sigma} dt + e^{-\rho\pi} V_{P+\pi}^R\left(a, h, d_g^{P+\pi}\right)$$

The first term is the discounted utility during the punishment phase from P to $P + \pi$. During that time, consumption equals earnings net of garnishments (from private lenders) and net debt repayments (to GSL lenders.) The second term reflects the discounted utility acquired after the punishment phase. When entering this phase, the individual carries a liability with GSL lenders equal to $d_g^{P+\pi} = e^{r\pi}\left[d_g - \int_P^{P+\pi} e^{-r(t-P)} r(t, d_g) dt\right]$ but is cleared of all private debt. At the end of the punishment period, he is granted unrestricted access to financial markets and can optimally smooth consumption thereafter. His utility $V_{P+\pi}^R\left(a, h, d_g^{P+\pi}\right)$ as of that time is determined by equation (15) but for time $t = P + \pi$. Note that the value of default $V_P^D(a, h, d_g, r(\cdot; d_g))$ depends on the actual timing of GSL repayment $r\,(\cdot; d_g)$ but not on the amount of private debt d_p.[39]

[37] See Monge-Naranjo (2009) for a continuous time model in which default can take place in any period and the optimal contract must satisfy a continuum of participation constraints.

[38] For example, see Lochner and Monge-Naranjo (2002).

[39] The value of repayment does not depend on the timing of GSL repayment, since individuals that do not default can freely borrow and lend after school to fully smooth consumption. On the other hand, private debt is irrelevant in the case of default, since borrowers are cleared of all debts.

A borrower would repay a private debt d_p if and only if $V_P^R(a, h, d_g + d_p) \geq V_P^D(a, h, d_g, r(\cdot; d_g))$, a condition that implicitly defines the maximum private credit limit as a function of a, h, and d_g. The many different paths that GSL repayments $r(\cdot, d_g)$ can take affect this private credit limit. As shown in Appendix C, the faster an individual must repay GSL loans, the more costly it is to default and the more he can borrow in private loans. That shorter GSL repayment schedules allow private lenders to extend more credit is interesting given recent moves in the U.S. to extend GSL repayment periods through loan consolidation and other options. We consider the effects of different repayment periods further below.

For analytical tractability, we assume that GSL debtors must repay at least a constant fraction δ of their earnings during the punishment period to service their GSL debt. Therefore, consumption during the punishment period equals $ah^\alpha E(t-P)(1-\gamma-\delta)$ and the value of GSL debt balance at the end of the punishment period equals $e^{r\pi}\left(d_g - \delta\Phi_{[P,P+\pi]}ah^\alpha\right)$. The structure of repayments after $P+\pi$ is irrelevant, since individuals can then freely borrow and save. Moreover, if we assume that the GSL repayment rate is set such that individuals repay a constant fraction of their income (net of garnishments in the case of default) over their entire working lives, i.e. $\delta^* = \frac{(1-\gamma)d_g}{\left[\Phi_{[P,R]}-\gamma\Phi_{[P,P+\pi]}\right]ah^\alpha}$, then we obtain a simple closed-form private lending constraint:[40]

$$d_p \leq \kappa_1 \Phi_{[P,R]} ah^\alpha + \kappa_2 d_g, \tag{17}$$

where $0 \leq \kappa_1 \leq 1$ and $\kappa_2 \geq -1$ are both constants that depend on preferences (σ, ρ), the interest rate r, and enforcement parameters (γ, π). (See Appendix C for details.) It is important to note that κ_1 and κ_2 incorporate the effects of punishment for default and that they do not depend on government subsidies s or the minimum human capital level h_0. Both s and h_0 only affect private constraints through total human capital h and GSL borrowing d_g. In general, we find that $\kappa_2 > -1$ and therefore private credit does not decrease one-for-one with expansions of GSL credit. However, $\kappa_2 < 0$ implies a partial 'crowding out' of private credit with expansions in GSL programs.

In contrast to the two-period model, even if wage garnishments are not allowed ($\gamma = 0$), private lending can be sustained ($\kappa_1 > 0$) as long as defaulting individuals face a disruption in their ability to smooth consumption by being excluded from credit markets for some period (i.e. $\pi > 0$). It is only when $\pi = 0$ that the punishment for default is negligible and private credit dries up entirely (i.e. $\kappa_1 = \kappa_2 = 0$). In general, the amount of sustainable borrowing (as determined by κ_1 and κ_2) is higher with: (i) tougher punishments (higher values of γ and π); (ii) more patient individuals (lower discount rate ρ) because future punishments are more costly; (iii) a stronger desire to smooth consumption (lower IES, higher σ), and (iv) higher growth in earnings with experience.

[40] In practice, American students can generally extend their GSL repayment period up to 30 years through loan consolidation and other repayment plans.

5.3.2 The Behavior of Human Capital Investment

We adopt the private lending constraints defined by equation (17) as our baseline case. Therefore, optimal schooling investment decisions maximize initial discounted utility given by (9) subject to $h = h_0 + (1+s)\, h_I$, the budget constraint $\int_S^P e^{-r(t-S)} c\,(t)\, dt + h_I \leq w + e^{-r(P-S)}\,(d_p + d_g)$, and the credit constraints (16) and (17).

Investment behavior is analogous to that of the two-period model. Given our human capital production function, we can define $\varrho(a) \equiv \frac{d_{\max}}{\Phi_{[P,R]} a [h_0 + (1+s) d_{\max}]^{\alpha}}$. For $a > \bar{a}$, we have $\varrho(a) < \alpha e^{-r(P-S)} \left[1 + \frac{h_0}{(1+s) d_{\max}}\right]^{-1} < \alpha < 1$, which is useful in characterizing the ability – investment relationship. See Appendix C for further details and the proof for the following proposition.

Proposition 4 *Consider individuals with ability $a > a_0$ (i.e. $h_I^U\,(a) > 0$) and whose wealth w is below the threshold $w_{\min}^{G+L}(a)$, so constraints (16) and (17) bind. Then, the following holds: (1) If $a > \bar{a}$, then: (i) $h^{G+L}\,(a, w) < h^U(a)$, (ii) $h^{G+L}\,(a, w)$ is strictly increasing in w, (iii) $h^{G+L}\,(a, w)$ is strictly increasing in a if either* (a) $\kappa_1 \geq \frac{\Theta_{[S,P]}}{\Theta_{[S,T]}}$ *or* (b) $\sigma \leq \left[1 - \left(\frac{1+\kappa_2}{1-\kappa_1}\right) \varrho\,(a)\right] \left[1 - \kappa_1 \frac{\Theta_{[S,T]}}{\Theta_{[S,P]}}\right]^{-1}$ *hold. (2) If $a < \bar{a}$, then: (i) $h^{G+L}\,(a, w) > h^U(a)$, (ii) $h^{G+L}\,(a, w)$ is strictly decreasing in w, and (iii) $h^{G+L}\,(a, w)$ is strictly increasing in a.*

As with the two-period model, the strength of the link between lifetime earnings and private credit (κ_1) is important in generating a positive ability – investment relationship. Even for high values of σ (i.e. low values of the IES) the sufficient conditions of the proposition hold if κ_1 is high enough, e.g. higher than $\frac{\Theta_{[S,P]}}{\Theta_{[S,T]}}$, the ratio of present value consumption while in school relative to lifetime consumption. This is just a sufficient condition, and it is likely to hold if the schooling period is short relative to the lifespan or if the agent is patient. As in the two-period model, the effect of $d_{\max}$ is complex: on one hand, a higher $d_{\max}$ increases $\bar{a}$ and more individuals can directly finance $h^U\,(a)$ with GSL programs alone; on the other hand, a higher $d_{\max}$ increases $\varrho\,(a)$ and makes it more difficult for the sufficient condition in part (1)(iii)(b) to hold.

Aside from the important advantage of facilitating calibration, the multiperiod model embodies economic forces that are absent from the two period model. Most interestingly, κ_1 is increasing in σ because the cost of disrupting consumption smoothing is increasing in the curvature of the utility function. Indeed, the sufficient conditions for part (1)(b) can hold more easily when σ is high (IES low). This is the case under our calibration.

5.4 Parameter Values

We now discuss the parameter values used to study the quantitative implications of our model. We normalize time so that a unit interval represents a calendar year. All monetary amounts are denominated in 1999 dollars using the Consumer Price Index (CPI-U). As a measure of ability, we use quartiles of the AFQT distribution in our sample. This facilitates comparison with the empirical patterns discussed earlier in Section 3. Baseline parameter values, reported in Table 3,

Table 3: Baseline Model Parameters

Calibrated Parameters			GMM estimates		
Parameter	Value	To match:	Parameter	Value	Coefficient on:
S	19		g_0	0.03	Experience
P	26	US Demographics			
R	65		α	0.70	Schooling investment
T	80		h_0	160,312	Min. human capital
π	10	U.S. Legal environment			
$\rho = r$	0.05	See text			AFQT Quartiles:
σ	2	IES = 0.5	a_1	1.51	1
$d_{\max}$	35,000	GSL Loan Limits	a_2	1.55	2
γ	0.2	Garnishments & other costs	a_3	1.60	3
s	0.80	Subsidy school grades 10+	a_4	1.72	4

are chosen to match basic features of the U.S. economy, while others are estimated using data on earnings and educational attainment from the random sample of males in the NLSY79.

With our focus on college education, we assume that youth (investment period) begins at age $S = 19$; maturity (labor market participation) begins at age $P = 26$; and retirement runs from age $R = 65$ until death at age $T = 80$. These values roughly capture the demographics and the timing of college education and labor market decisions in the U.S.

We assume an annual interest rate $r = 0.05$ based on historical averages of the risk-less rate and the return to capital in the U.S. We also set $\rho = r$. Given our calibration strategy, reasonable variations of ρ and r, including differences between them, have little impact on our results. We set $\sigma = 2$, which implies an IES of 0.5 – an intermediate value in the estimates reported in Browning, Hansen, and Heckman (1999). Values of σ inside the interval $[1.5, 3]$ yield similar results.

We calibrate the length of the penalty period π based on the U.S. legal environment. According to U.S. bankruptcy code, individuals must wait for at least 7 years after filing for Chapter 7 before they qualify to file again, while default records remain in an individual's credit history for a period of 10 years. Thus, π should range between 7 and 10. In our baseline, we set $\pi = 10$, but $\pi = 7$ produces similar conclusions. Regarding the effective earnings 'lost' in the event of private loan default, regulations provide little direct guidance. For private unsecured loans, an explicit garnishment rule does not exist. Moreover, actual costs of default – either via direct penalties or via avoidance actions – extend beyond simple garnishments (e.g. individuals may end up sub-optimally employed, renting instead of owing a house, and paying sub-prime interest rates for short-term transactions, etc.) Finally, defaulting may involve other non-pecuniary costs as well as disruption in career possibilities. We set $\gamma = 0.2$ as the baseline fraction of lost earnings for individuals who default. This fraction is a bit higher than in Livshits, MacGee, and Tertilt (2007) and Chatterjee, et al. (2007), partly because we abstract from the benefits of financial markets in smoothing out temporary earnings and preference shocks. Finally, we assume $d_{\max} = 35,000$ based

on loan limits for Perkins and Stafford Loan Programs.[41] Below, we explore reasonable changes in this upper GSL loan limit as well as private loan enforcement parameters γ and π. Finally, we also report how the predictions of this model differ from an exogenous constraint model with a limit equal to $d_0 = 70,000$, a value we explain below.

5.4.1 Estimation of the earnings function

Data on wage income, education, age, and AFQT quartile from the NLSY79 (1979-2006) are used to estimate parameters of the labor earnings function. Our sample includes all men ages 19+ with at least 12 years of completed schooling from the random sample. We associate different levels of investment with different levels of reported schooling, calculating the total expenditures associated with each level of schooling separately by AFQT quartile. These costs include both foregone earnings and direct expenditures as discussed below. Consistent with the formulation of the model, we make no distinction for investment in time costs (foregone earnings) or purchased inputs. Implicitly, they are perfect substitutes in the production of human capital, an issue we discuss further in the online appendix. We also abstract from investment differences related to differences in college quality. While an interesting margin of choice, we leave this to future work.

Estimation of the labor earnings function proceeds in three separate steps:

Step 1: Estimating foregone earnings. Foregone earnings reflect the present value of average earnings relative to someone with 12 years of completed schooling, taking into account earnings during college. For someone with C years of college, we first calculate average earnings at ages 19 through $19 + C$ by AFQT quartile for all non-enrolled men with exactly 12 years of completed schooling (based on highest grade reported in the NLSY79). We then subtract predicted earnings at ages 19 through $19 + C$ for men currently enrolled in college to obtain our measure of foregone earnings.[42]

Step 2: Determining total costs of schooling and the government subsidy matching rate. Total schooling expenditures are set to zero for those with only 12 years of completed schooling. For those attending college, we add foregone earnings determined in Step 1 to direct costs to determine total schooling expenditures. Direct expenditures are based on current-fund expenditures per full-time equivalent student in all institutions of higher education (1999 Digest of Education Statistics, Table 342). Direct expenditures for the first two years of college are based on

[41] As discussed in Section 2, there are a number of different government loan limits depending on the type of loan, dependency status, and whether the student is an undergraduate or graduate student. Our choice of \$35,000 is higher than the limit for dependent undergraduates borrowing only from the Stafford program but lower than the limit for independent undergraduates or for graduate students.

[42] Let $\bar{y}_{12}(q, j)$ reflect average wage income for men with 12 years of schooling, AFQT quartile q, and age j. Let $\hat{y}_C(q, j)$ reflect predicted earnings for men with C years of completed college (i.e. highest grade completed less 12), AFQT quartile q, and age j. This prediction is based on a regression of earnings on AFQT quartile indicators, experience (= age -19), and experience-squared using a sample of men that are enrolled in college and whose age is between 19 and 26 (with age not exceeding $18 + C$). Then, foregone earnings for someone with AFQT quartile q and $C \geq 1$ years of college completed are calculated as $FE(q, C) = \sum_{x=0}^{C-1} (1 + r)^{1-x} [\bar{y}_{12}(q, 19 + x) - \hat{y}_C(q, 19 + x)]$ for interest rate $r = .05$.

2-yr school averages for academic years 1980-81 to 1984-85, while direct expenditures for 3+ years of college are based on 4-yr school averages for academic years 1980-81 to 1989-90.[43] These dates correspond to the years most students in our NLSY79 sample attended college. See Table D1 in Appendix D for measures of direct expenditures, foregone earnings, and total expenditures.

To calculate the subsidy rate s used in our analysis, we first compute marginal subsidy rates for each year of college (1-8 years) by AFQT quartile.[44] Since these rates differ somewhat by the number of years of schooling and AFQT quartile, we average over these values using the distribution of completed schooling in our NLSY79 sample. The resulting government subsidy matching rate is $s = 0.799$. In simulating the 'year 2000' economy below, we use a lower subsidy matching rate of $s = 0.710$, consistent with the observed rise in current fund revenue that came from tuition.[45]

Step 3: Estimating the parameters. With Step 2, we have imputed total investment expenditures $h(q, C)$ for each AFQT quartile q and years of completed schooling C.[46] We next estimate α, h_0, g, and ability parameters $a_1, ..., a_4$ using NLSY79 data on wage income, schooling, and age.

In the model, earnings at any experience $x = age - 26$ for someone who invested h_I are:

$$y(a, x) = aE(x; g)[h_0 + h_I\,(1 + s)]^{\alpha},$$

where $E(x; g)$ is a known function of experience and parameters g. We use $E(x; g) = exp(gx)$, so earnings are log linear in experience. Taking logs and assuming earnings are measured with error ε_i, we have the following specification for individual earnings as a function of AFQT quartile q_i and schooling C_i:

$$ln(y_i) = ln[a_{q_i}] + gx + \alpha\; ln[h_0 + h(q_i, C_i)] + \varepsilon_i. \tag{18}$$

The model's implied unconstrained investment for someone with ability a is given by expression (13) with $\Phi_{[P,R]} = \frac{e^{(g-r)(R-P)}-1}{g-r}$ given our assumption that earnings are log linear in experience.

We use GMM to estimate our parameters using moments based on both (18) and (13):

$$E\left\{[ln(y_i) - (ln[a_{q_i}] + gx + \alpha\; ln[h_0 + h(q_i, C_i)])]Z_i\right\} = 0$$
$$E\left\{h_0 + h(q_i, C_i) - h^U(a_{q_i})|q_i\right\} = 0,$$

[43] Direct expenditures for 3+ year of college are based on 3+ years of average expenditures at 4-year schools, while direct expenditures for 1-2 years are based on 1-2 years at two-year schools. Because average expenditures at 4-year institutions are higher than at 2-year institutions, this generates a noticeable jump in direct expenditures between 2 and 3 years of college.

[44] The marginal subsidy is computed as 0.77 × direct expenditures divided by total expenditures, where 0.77 reflects the ratio of current-fund revenue that does not come from tuition and fees averaged over academic years 1980-81 to 1989-90 (Digest of Education Statistics, 2003, Table 333).

[45] In 1995-96, the ratio of current-fund revenue that did not come from tuition and fees was 0.72, down from 0.77 in the 1980s (Digest of Education Statistics, 2003, Table 333).

[46] Since we include total expenditures in calculating $h(q, C)$, it reflects total private investment plus public subsidies (i.e. $h_I(a)(1 + s)$).

where Z_i includes indicators for each year of schooling from grades 12 to 20, experience x, and AFQT quartile indicators.[47] The first set of moments using the wage equation simply estimates parameters to best fit average earnings conditional on schooling, age, and AFQT quartile. Using only this set of moments is nearly identical to non-linear least squares estimation of equation (18). With the second set of moments, we also match average schooling expenditures by AFQT quartile with the unconstrained optimal levels as implied by the model.[48]

The strategy of estimating of (a_q, α, g, h_0) targeting unconstrained investments is consistent with evidence in the NLSY79 (e.g. Cameron and Heckman (1998, 1999) and Carneiro and Heckman (2002)) suggesting that most individuals were not constrained in their schooling investments at that time. However, it is important to note that this does not guarantee that simulations of our baseline model necessarily lead to these unrestricted investments. None of our assumptions about the credit environment (i.e. the GSL program and private lending under limited commitment) imply adequate credit for everyone. Therefore, one metric for evaluating our model is whether anyone is constrained in our baseline calibration. For this to be the case, the thresholds $w_{\min}^{G+L}(a)$ implied by the model should be low.

Finally, it is important to discuss the nature and correct interpretation of an individual's available resources w in our simulations. Because foregone earnings are an important part of investment expenditures in our calibration, w includes at least the amount he could earn if he began working immediately after high school. These amounts depend on ability, since foregone earnings depend on ability (see Appendix Table D1). The relevant range of available resources, therefore, begins at $\$36,000$ for the least able, $\$73,000$ for AFQT quartile 2, $\$76,000$ for AFQT quartile 3, and $\$79,000$ for the top quartile. Any available resources above these amounts must be interpreted as transfers from parents or others.

5.5 Baseline Simulations

We now report the model's main implications given our baseline parameterization. Figure 3 shows the wealth threshold $w^{G+L}(a)$ for our benchmark model. It also shows $w^{G}(a)$ and $w^{L}(a)$, the special cases when we shut down private or GSL credit, respectively. Individuals with ability-wealth pairs above and to the left of the thresholds are unconstrained, while those with pairs below and to the right are constrained. The x-marks indicate the point estimates for each ability quartile as reported in Table 3. Finally, the dotted horizontal lines reflect estimated potential earnings (PE) for these same ability levels.

[47] We do not attempt to address concerns about unobserved heterogeneity in estimating the wage equation (i.e. we assume ε_i is orthogonal to completed schooling conditional on AFQT).

[48] We could have estimated the parameters of the human capital production function using only the moments based on the wage equation. However, this produces fairly noisy estimates of most parameters, especially h_0. Since the model implies an optimal unconstrained investment that is quite sensitive to all parameter values, including the second set of moments provides much more precise and robust estimates. We do not lose much in terms of mean squared error (MSE) for the log wage equation when estimating the model using both sets of moments. This MSE only increases from 0.593 to 0.601 when the second set of moments are used.

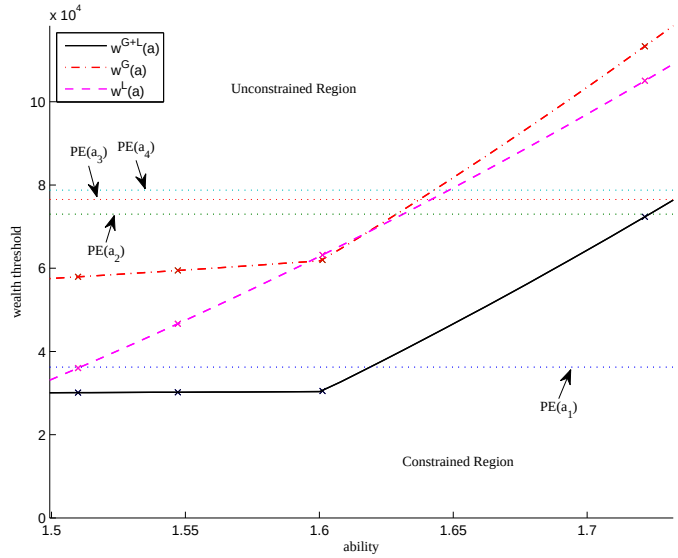


Figure 3: Thresholds for Unconstrained Allocations (Baseline)

The baseline model implies that individuals in the NLSY79 were unconstrained, justifying our estimation strategy of matching unconstrained investment with average investment in the data. For all estimated ability types a_q, the dotted $PE(a_q)$ lines lie above the corresponding wealth threshold $w^{G+L}(a_q)$. This implies that even youth who receive zero transfers (from their parents or other sources) can attain unconstrained consumption and investment allocations given available credit from the GSL and private lending. Regardless of individual resources, our baseline parametrization implies investments of \$8,000, \$22,300, \$44,600, and \$100,900 for AFQT quartiles 1, 2, 3 and 4, respectively.[49] Clearly, our model implies considerable differences in educational investment between the most and least able as observed in the NLSY79.

Figure 3 also reveals that ability quartiles 2 and 3 would be unconstrained by the GSL alone; thus, middle ability individuals would not need to borrow from private lenders. Lower ability individuals lie in the flat region of w^G and w^{G+L}, indicating that the GSL's tied-to-investment constraint (i.e. $d_g \leq h_I$) may bind. The fact that $w^G(a_1) < PE(a_1) < w^{G+L}(a_1)$ implies that the least able would be constrained (low consumption during school) under the GSL alone, but they receive enough credit from private lenders to enable full consumption smoothing. Among the most able, the upper GSL loan limit (i.e. $d_g \leq d_{\max}$) binds for those receiving no family transfers. They would under-invest without access to private lenders; however, private lenders provide enough credit to ensure unconstrained maximization.

Figure 4 reports total borrowing $d_g + d_p$ for each level of ability as a function of initial wealth minus potential earnings (i.e. family or outside transfers). Only youth from the top AFQT quartile

[49] These reflect total expenditures for post-secondary education and are very close to average total expenditures by AFQT quartile in the NLSY79 (from least to most able): \$8,800, \$29,000, \$47,400, \$107,700. See Table D1 in the appendix for a mapping between these amounts and years of college attendance.

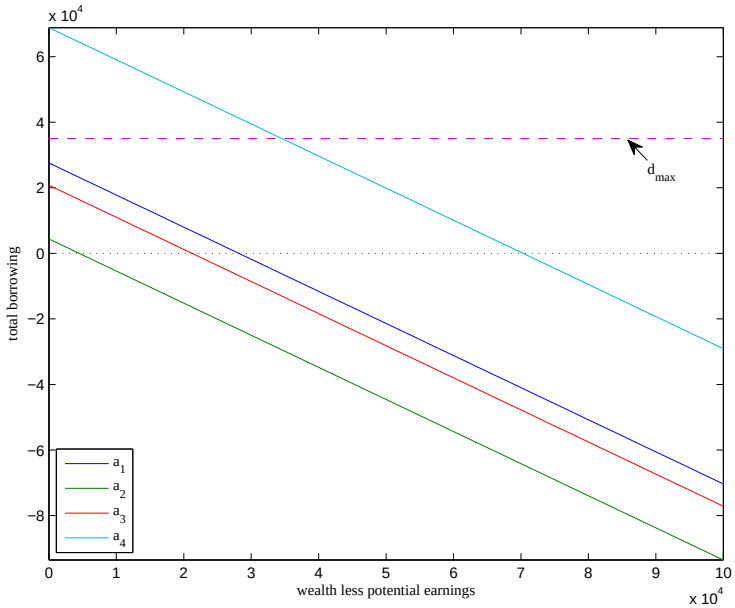


Figure 4: Total Borrowing as a Function of Ability and Wealth less Potential Earnings (Baseline)

with very low resources wish to borrow more than the upper GSL limit (reflected in the dashed horizontal line at $35,000).[50] Among the most able, roughly $35,000 in family transfers (received over ages 19-26) would be enough to ensure unconstrained consumption and investment without any need for private loans. All other youth wish to borrow less than the GSL maximum. As noted above, youth from the lowest ability quartile would like to borrow more than they invest, which the GSL does not accommodate. As a result, the least able receiving less than $20,000 in family transfers (cumulative over ages 19-26) would like to borrow small amounts from private lenders (e.g. credit cards) in order to smooth consumption. Youth in the interquartile range invest more than they wish to borrow from the GSL and do not run up against the GSL upper loan limit; they are fully unconstrained by the GSL regardless of parental transfers.

Altogether, our baseline model fits the '1980s facts' quite well. The prediction that investment is unconstrained for all ability levels is consistent with the evidence from the NLSY79, i.e. that investments are independent of the individual's wealth and strongly increasing in ability. The model further predicts that most NLSY79 respondents should borrow less than the GSL maximum. Only youth with high ability and low family transfers would borrow up to the GSL maximum and then some from private lenders. This is consistent with the fact that early private student loan programs in the 1980s were relatively unimportant and almost exclusively served students of elite institutions and professional schools.

[50]This group should be quite small given the strong correlation between ability and family resources observed in the NLSY79. For example, Table 2 in Belley and Lochner (2007) reveals that 70% of youth from the highest AFQT quartile have family income in the top half of the distribution.

6 Counterfactual Exercises

We now use our model to conduct two sets of counterfactual exercises. First, we simulate an increase in both the costs of and the returns to education as observed between the 1980s and early 2000s to see whether our model is consistent with the rising importance of family resources as a determinant of schooling and the increase in student borrowing from both the GSL and private lenders. Second, we conduct a number of policy experiments related to the financing of college education.

6.1 A Rise in the Costs of and Returns to Schooling

We simulate the effects of an increase in the costs of and returns to schooling — two major economic changes that took place between the early 1980s and early 2000s. We model an increase in the wage returns to education by assuming that α increases by 0.01 to 0.71, which leads to a modest increase in the college – high school log wage differential.[51] We model the rise in net tuition costs by assuming that the government subsidy rate, s, falls from 0.799 to 0.71. As discussed above, this reduction reflects the increased importance of tuition and fees as a fraction of total current-fund revenue for public and private universities in the U.S. Our simulations capture the observed stability of maximum GSL loans in real terms by assuming that $d_{\max}$ remains unchanged at $\$35,000$. We refer to the baseline parameterization as the "1980s economy" and to the counterfactual parameterization as the "2000s economy."

The model suggests that the higher returns to investment led to an increase in the amount of available private credit. However, the demand for credit rose even more such that the $w^{G+L}(a)$ thresholds increased substantially relative to their 1980 levels as shown in Figure 5. A much larger set of wealth-ability pairs lies in the constrained region in the 2000s. The model suggests that many youth receiving little or moderate transfers from their parents are likely to be borrowing constrained in the more recent period. Finally, the kink in the threshold that was present in the 1980s economy disappears completely in the 2000s economy. Only the steep region of the threshold remains, indicating that the only potentially binding constraint in the GSL is $d_g \leq d_{\max}$. This rules out the possibility of over-investment.

In the 2000s economy, wealth becomes an important determinant of human capital investment. Figure 6 shows total investment, $h_I(1+s)$, as a function of available resources, w. The solid lines represent investment for the estimated ability levels by AFQT quartile; dotted vertical lines indicate potential earnings and delineate the empirically relevant regions of w for each ability quartile.

[51] There is some disagreement in the literature regarding the underlying cause for the increase in estimated college – high school wage differentials. Some argue that much of the rise is due to a rising 'return to ability', while others argue that most of the rise is due to an increase in the actual 'return to school'. See Cawley, et al. (2000) and Taber (2001) for detailed discussions of the empirical difficulties and evidence. Changing α more closely reflects the latter, but we increase α less than the amount needed to fully account for the rise in the college – high school log wage differential. An increase in the 'return to ability' is equivalent to shifting the ability distribution upwards in our framework, which produces qualitatively similar effects to those discussed here.

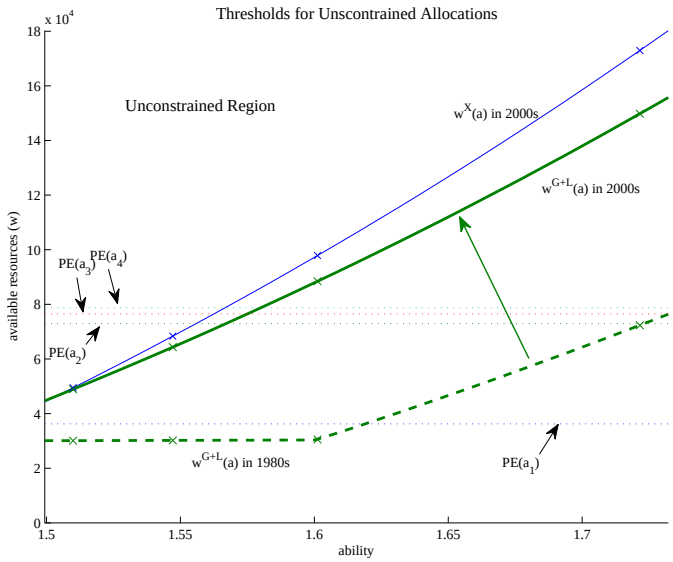


Figure 5: Thresholds for Unconstrained Allocations

Consistent with the predictions of Proposition 4, constrained investment is steeply increasing in wealth until it reaches the unconstrained level. Constraints are binding for a wide range of wealth levels. Most notably, the top ability individuals can only reach unconstrained investments and consumption if their parents give them at least \$70,000 during college.

Credit available from the GSL is no longer sufficient in the 2000s economy. As shown in Figure 7, the model predicts a significant expansion in the set of individuals borrowing beyond the maximum $d_{\max}$ from the GSL. Private lending expands to the point that it is comparable to or greater than GSL borrowing for youth with low-to-medium parental transfers. Among the most able, borrowing from private lenders is as much as \$50,000 for a large range of wealth (and parental transfer) levels. Private lenders are willing to provide the extra credit in the 2000s, because the increased return to investment raises earnings and the cost of default. Interestingly, borrowing is not monotone in wealth, because constrained wealthy individuals consume and invest more. The latter expands private credit.

The endogeneity of credit limits is important. To see this, compare our baseline model with an exogenous constraint model. Figure 5 includes the threshold $w^X(a)$ assuming $d_0 = 70,000$ the exogenous limit that yields the same wealth threshold for the lowest ability quartile in the '2000s economy'.[52] The same set of low ability individuals are constrained in either model, but the steeper $w^X(a)$ curve implies that more higher ability individuals are constrained under exogenous constraints. The gap between the two thresholds is increasing in ability, since private credit endogenously increases with ability.

The endogeneity of borrowing limits implies an 'extensive margin' effect on the ability – in-

[52] That is, $w^{G+L}(a_1) = w^X(a_1)$. This exogenous constraint level is also consistent with the 1980s, since it does not bind for any estimated ability levels.

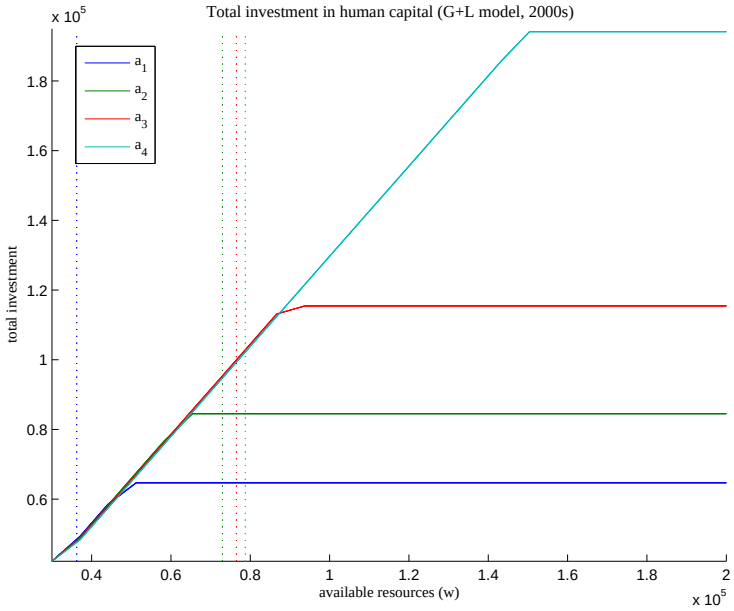


Figure 6: Total investment in human capital (2000s)

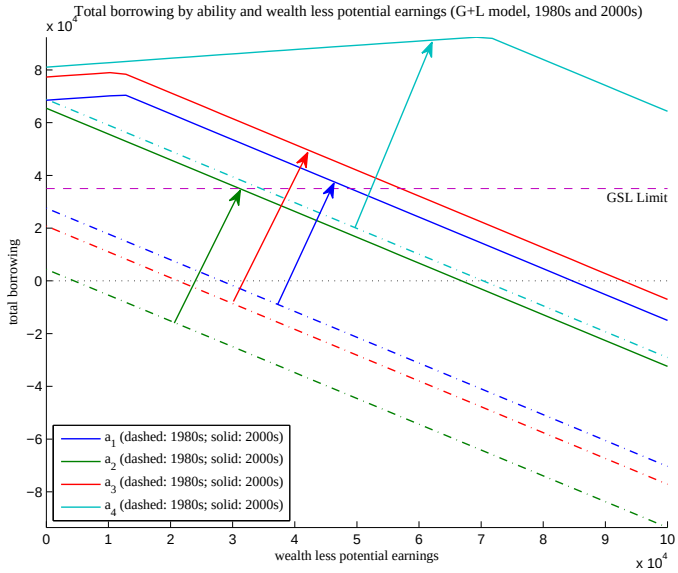


Figure 7: Total borrowing by ability and wealth less potential earnings (1980s and 2000s)

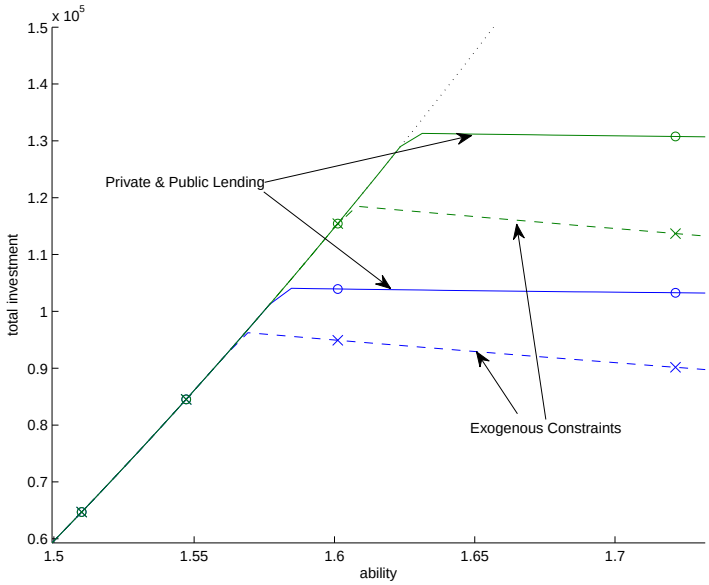


Figure 8: Total investment in human capital for $w = 80,000$ and $w = 100,000$ (2000s)

vestment relationship, since more able persons attain unrestricted investments. There is also an 'intensive margin' effect among those that are constrained. Figure 8 compares the relationship between ability and human capital investment (for two wealth levels) implied by exogenous constraints and our baseline model with endogenous GSL and private credit constraints. The effects of endogenous constraints on the extensive margin are evident in the wider range of abilities for which unconstrained investment is observed. The effects on the intensive margin for those that are constrained is reflected in the different slopes between the solid and dashed lines at higher abilities. As expected from Proposition 1 and $\sigma > 1$, the exogenous constraint model predicts that constrained investment is decreasing in ability. With exogenous constraints, low-income youth from the top AFQT quartile would invest 5% *less* than youth from the third quartile. In contrast, our baseline model predicts that constrained investments are essentially flat in ability.[53]

It is noteworthy, however, that our model delivers the observed positive ability – investment relationship at the bottom of the family income distribution where family transfers are likely to be negligible. Comparing youth receiving no family or other transfers (i.e. $w = PE(a)$), the most able invest more than double the least able. This is because potential earnings (i.e. resources available to those receiving zero transfers) are increasing in ability. More generally, total investment is increasing in ability for any given level of transfers, $w - PE(a)$.

In sum, our model predicts that the increase in costs and returns to schooling have led to a rise in borrowing from both the GSL and private lenders. The model further predicts what while more youth have become constrained, private lenders have expanded credit opportunities in response to the higher earnings associated with a college education. These patterns are consistent with

[53]We can easily generate steeper ability – investment profiles for our baseline model using higher values for γ and π.

Table 4: Effects of Lending Policy Changes on Human Capital Investment (in % terms)

	'Year 2000' Baseline	Private Lending Parameters: $\pi =$ 0	7	15	$\gamma =$ .1	.3	GSL Parameters: $d_{\max} =$ 0	50,000	$M = 15$
$h^{G+L}(a_1, w)$:									
$w = PE(a_1)$	48,239	-12.6	-12.6	19.5	-12.6	34.1	-86.7	34.1	34.1
$w = 50,000$	64,702	-34.8	-11.0	0.0	-25.3	0.0	-61.8	0.0	0.0
$w = 80,000$	64,702	-9.0	0.0	0.0	0.0	0.0	-3.2	0.0	0.0
$w = 100,000$	64,702	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$h^{G+L}(a_2, w)$:									
$w = PE(a_2)$	84,529	-41.0	0.0	0.0	-10.5	0.0	-36.6	0.0	0.0
$w = 80,000$	84,529	-31.8	0.0	0.0	-0.3	0.0	-25.9	0.0	0.0
$w = 100,000$	84,529	-5.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$w = 120,000$	84,529	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$h^{G+L}(a_3, w)$:									
$w = PE(a_3)$	99,966	-48.0	-10.0	11.7	-21.0	15.5	-41.9	15.5	15.5
$w = 80,000$	104,485	-46.5	-9.7	10.5	-20.3	10.5	-40.1	10.5	10.5
$w = 100,000$	115,447	-32.4	0.0	0.0	-6.8	0.0	-23.6	0.0	0.0
$w = 120,000$	115,447	-13.2	0.0	0.0	0.0	0.0	-1.3	0.0	0.0
$h^{G+L}(a_4, w)$:									
$w = PE(a_4)$	102,213	-50.3	-10.6	12.3	-22.1	28.2	-40.3	17.5	28.3
$w = 80,000$	103,819	-49.8	-10.5	12.2	-21.9	27.9	-39.7	17.2	27.9
$w = 100,000$	129,693	-43.0	-9.0	10.4	-18.8	23.9	-32.0	13.9	22.2
$w = 120,000$	155,581	-38.4	-8.0	9.2	-16.8	21.3	-26.8	11.6	18.4

Notes: Unconstrained investments, $h^U(a)$, are \$64,702, \$84,529, \$115,447, and \$194,164.

the evidence on family income – college attendance patterns in the NLSY79 and NLSY97, the increased fraction of youth constrained by upper GSL limits, and the expansion of private credit discussed in Section 2. While the model does not necessarily deliver a strong positive relationship between ability and schooling conditional on available resources for constrained youth, it performs noticeably better than the exogenous constraint model.

6.2 Policy Experiments

We next consider the response of human capital investments to three types of changes in the economy: (i) changes in the enforcement institutions underlying private lending; (ii) changes in the extent of GSL programs; and (iii) changes in government subsidies. In all exercises, our point of departure is the 2000s economy where some agents are constrained. We report the response for the lowest resources available by ability (i.e. potential earnings, $PE(a)$), and for other levels of w. For the lowest ability quartile, we report the results for lower values of w, because their potential earnings are substantially lower.

Changes in the enforcement of private lending. Columns 2-6 of Table 4 show the percentage change in human capital investment (relative to the 2000s economy benchmark in column 1) for each ability quartile and different levels of available resources w. Column 2 presents the case of $\pi = 0$, when private lending collapses to zero and the GSL is the only source of credit. The elimination of private lending leads to sizeable reductions in investment, as much as 50% for bright youth from poor families. Columns 3 and 4 show that variations in π closer to our benchmark value of 10 years lead to more modest responses in human capital investments. Except for the most able, a punishment period of 15 years would lead to unconstrained investments for all wealth levels; top ability students from poor backgrounds would remain constrained but would invest considerably more than under the benchmark. The punishment period would need to be extended to near retirement (i.e. $\pi \approx R-P$) before the most able with no familial income transfers would be unconstrained.

The next two columns of Table 4 show that a reduction in γ to 0.1 would reduce investment by as much as 25% for the poorest youth of different ability levels, while increasing γ to 0.3 would lead to unconstrained investment for all but the highest ability quartile. Although the latter would substantially increase investment among the most able (by nearly 30% for the very poor), γ needs to rise above 0.45 before everyone is unconstrained. Of course, simultaneously increasing π and γ would more easily ensure unrestricted investment in human capital for everyone.

Changes in the GSL program. The remaining columns of Table 4 consider changes to GSL programs. First consider eliminating the GSL program altogether ($d_{\max} = 0$). This policy change would severely restrict investment among the poorest and least able. However, the effects are fairly large for all poor youth regardless of ability. Comparing these results against those with only government lending (i.e. $\pi = 0$) suggests that the GSL is more important for investment among the least able, while private lending is more important for all other ability groups. This is because contrary to the GSL, private lenders base credit directly on ability. Next, we consider a modest expansion in the GSL program, increasing the upper limit to $d_{\max} = \$50,000$. Such a policy would disproportionately benefit the least able poor, but it would also help low income youth of high ability. As with an increase in γ to 0.3, this GSL expansion enables unconstrained investment for the bottom three-quarters of the ability distribution, while effects are comparatively weaker for the most able.

The last column of Table 4 reports the impact of changing the GSL repayment period. Recall that our baseline model allows individuals to spread their GSL re-payments over their entire working careers. Here, we consider reducing the repayment period to 15 years after the completion of school (see appendix C for details). This change effectively increases the cost of default by reducing resources available for consumption during the period of exclusion from financial markets. Interestingly, such a policy would have nearly identical effects on private lending constraints and human capital accumulation as increasing γ to 0.3. Therefore, our baseline calibration closely mimics a model with a shorter GSL repayment period and lower γ.

Table 5: Response of Investment to Increasing Subsidy Rate to 1980s Level (2000s Economy)

	2000s benchmark levels		% Changes from benchmark	
	h^{G+L}	h^X	h^{G+L}	h^X
$h(a_1, w)$:				
$w = PE(a_1)$	$48,239$	$49,823$	8.7	6.7
$w = 50,000$	$64,702$	$64,702$	10.0	7.2
$w = 80,000$	$64,702$	$64,702$	66.8	61.9
$w = 100,000$	$64,702$	$64,702$	66.8	66.8
$h(a_2, w)$:				
$w = PE(a_2)$	$84,529$	$84,529$	21.1	12.2
$w = 80,000$	$84,529$	$84,529$	32.4	22.0
$w = 100,000$	$84,529$	$84,529$	55.7	49.9
$w = 120,000$	$84,529$	$84,529$	55.7	55.7
$h(a_3, w)$:				
$w = PE(a_3)$	$99,966$	$91,481$	6.9	5.7
$w = 80,000$	$104,485$	$95,382$	6.8	5.6
$w = 100,000$	$115,447$	$115,447$	20.3	7.6
$w = 120,000$	$115,447$	$115,447$	43.9	28.0
$h(a_4, w)$				
$w = PE(a_4)$:	$102,213$	$89,252$	7.0	5.8
$w = 80,000$	$103,819$	$90,625$	6.9	5.8
$w = 100,000$	$129,693$	$112,766$	6.5	5.5
$w = 120,000$	$155,581$	$134,984$	6.3	5.3

Response to education subsidies. Finally, consider the effects of reducing the government subsidy rate s to its 1980s level (in our benchmark 2000s economy). As Table 5 demonstrates, a higher subsidy rate leads to substantial increases in investment with the largest responses among wealthier, unconstrained youth. Investment among constrained youth responds less, because they also want to consume more while in school. Overall, a universal subsidy to investment amplifies inequality in earnings.

Table 5 also compares the investment responses for our model (h^{G+L}) with those for an exogenous constraint model (h^X) with $d_0 = 70,000$. Since private credit expands with investment in our framework, investment responses are always greater for constrained individuals than under exogenous constraints. The main differences are for the middle ability groups, where the effects are as much as 50% higher in our model compared to the exogenous constraint model.

With respect to the impact of these policies on welfare (not shown here), we make two remarks. First, impacts on welfare tend to be smaller than on human capital investment, because borrowers only benefit from the difference between the returns and costs of additional human capital. Second, impacts on welfare (across different policies or individuals) need not correlate highly with impacts on investment, because consumption is also an important margin of response to credit constraints.

7 Conclusions

GSL programs and private lending under limited commitment link the borrowing opportunities of young individuals with their cognitive ability and investments in human capital. We show that this link shapes the intertemporal trade-off between investment and consumption for those that are credit constrained and is important for understanding college attendance and borrowing patterns in recent decades. Most notably, the link is important for explaining the positive ability – schooling relationship for youth from low-income families and the rapid expansion in private student lending in recent decades. Conventional wisdom and numerous empirical studies presume that borrowing constraints always inhibit investment; however, we show that this is not the case if what constrains youth is the GSL's tied-to-investment constraint (i.e. their borrowing is restricted by their level of investment). Finally, we show that schooling is more sensitive to government policies when credit depends on investment behavior: policies that increase schooling also expand private credit opportunities, which further increases schooling among constrained youth.

A calibrated version of our model reinforces existing empirical findings that American youth were not constrained during the 1980s but suggests that many youth may be constrained today. This change is explained by rising college costs, even faster rising returns to education, and largely unresponsive GSL programs. Consistent with the evidence, our model predicts that these forces make family resources a more important determinant of higher education, cause more individuals to exhaust their government borrowing opportunities, and lead to an expansion in private student lending.

Our framework enables us to study the effects of changes in government student loan programs

on private lending. We show that expansions of government lending are only partially offset by reductions in private lending, so total student credit is increasing in GSL limits. In contrast, efforts to extend GSL repayment periods lead to contractions in private lending, since they reduce the costs associated with private loan default. These private credit responses, in turn, affect educational investment decisions. We also study the effects of changes in private loan enforcement or bankruptcy regulations on schooling in our framework. We show that expansions in private loan enforcement capabilities increase human capital investment, especially among the more able, while expansions in government credit tend to favor the least able.

Finally, our framework can serve as a natural starting point for future empirical work and policy analysis that incorporates dimensions ignored here. An obvious next step is to introduce uncertainty and learning about the returns to investment, opening the door to default in equilibrium. Default may serve as insurance against adverse outcomes, and loan contracts with private lenders and the GSL must strike a balance between ensuring repayment and providing insurance against unexpected outcomes. We have also abstracted from school quality and labor supply decisions while in school. Both are likely to be important margins of response in the face of credit constraints and deserve further attention. With reliable data on schooling, borrowing, earnings, and loan repayment (an admittedly tall order), estimation of models that explicitly incorporate government and private lending should provide important new insights on the nature of endogenous borrowing constraints, who is constrained, and the effects of higher education policies and economic changes on private credit offerings and, ultimately, individual schooling and borrowing decisions.

Appendices

A NLSY79 and NLSY97 Data

The NLSY79 is a random survey of American youth ages 14-21 at the beginning of 1979, while the NLSY97 samples youth ages 12-16 at the beginning of 1997.[54] Since the oldest respondents in the NLSY97 recently turned age 24 in the 2004 wave of data, we analyze college attendance as of age 21 in both samples.

Individuals are considered to have attended college if they *attended* at least 13 years of school by the age of 21.[55] For the 1979 cohort, we use average family income when youth are ages 16-17, excluding those not living with their parents at these ages. In the NLSY97 data, we use household income and net wealth reported in 1997 (corresponding to ages 13-17), dropping individuals not living with their parents that year.[56] We use AFQT as a measure of cognitive ability. It is a composite score from four subtests of the Armed Services Vocational Aptitude Battery (ASVAB) used by the U.S. military: arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operations. These tests are taken by respondents in both the NLSY79 and NLSY97 during their teenage years as part of the survey process. We categorize individuals according to their family income, family net

[54] Our sample and variables are explained in detail in Belley and Lochner (2007).

[55] Schooling attainment by age 22 is used if it is missing or unavailable at age 21 (fewer than 10% of all respondents in both surveys).

[56] Family income includes government transfers (e.g. welfare and unemployment insurance), but it does not subtract taxes. Net wealth is the value of all assets (e.g. home and other real estate, vehicles, checking and savings, and other financial assets) less loans and credit card debt.

wealth (in NLSY97), and AFQT score quartiles.[57]

B Proofs and Other Aspects of the Two-Period Model

The set of constrained individuals. For each ability level a, the various forms of credit constraints define a threshold wealth level below which the agent is constrained (and above which he is not). We now characterize those thresholds.

Exogenous Constraints: The threshold $w^X_{\min}(a)$ is defined by $d^U\left(a, w^X_{\min}(a)\right) = d_0$, and therefore it is increasing in a. Consumption smoothing implies that $w^X_{\min}(a) \geq h^U(a) - d_0$ (the minimum wealth needed to finance $h^U(a)$ given maximum borrowing) and that $w^X_{\min}(a)$ is steeper than $h^U(a)$ as a function of a. To see this, implicit differentiation leads to $\frac{dw^X_{\min}(a)}{da} = \frac{\partial d^U\left(a, w^X_{\min}\right)}{\partial a} / \frac{\partial d^U\left(a, w^X_{\min}\right)}{\partial w} > \frac{\partial d^U\left(a, w^X_{\min}\right)}{\partial a} > \frac{dh^U(a)}{da} > 0$.

GSL Programs: The threshold $w^G_{\min}(a) \equiv \max\{w^X_{\min}(a), \tilde{w}_{\min}(a)\}$, where $\tilde{w}_{\min}(a)$ is defined by $h^U(a) = d^U(a, \tilde{w}_{\min}(a))$. It is increasing in a because $d^U(\cdot, w)$ is steeper than $h^U(\cdot)$. To see that $w^X_{\min}(a)$ is steeper than $\tilde{w}_{\min}(a)$, use implicit differentiation to obtain $\frac{d\tilde{w}_{\min}(a)}{da} = \frac{dw^X_{\min}(a)}{da} + \frac{\partial h^U}{\partial a} / \frac{\partial d^U}{\partial w} < \frac{dw^X_{\min}(a)}{da}$.

GSL Programs Plus Private Lenders: The threshold $w^{G+L}_{\min}(a)$ is defined by $d^U\left(a, w^{G+L}_{\min}(a)\right) = \kappa a f\left(h^U(a)\right) + \min\left\{h^U(a), d_{\max}\right\}$. An instructive special case is when $d_{\max} = 0$ and only private lending is available in the economy. In this case, the threshold $w^L_{\min}(a)$ is defined by $d^U(a, w^L_{\min}(a)) = \kappa a f\left[h^U(a)\right]$, which increases at a slower rate in a than does $w^X_{\min}(a)$. Indeed, $w^L_{\min}(a)$ may even be decreasing in a if κ is large enough. Both of these facts can be seen from $\frac{dw^L_{\min}}{da} = \frac{dw^X_{\min}}{da} + \kappa\left(f(h^U) + R\frac{\partial h^U}{\partial a}\right) / \frac{\partial d^U}{\partial w} < \frac{dw^X_{\min}}{da}$ because $\frac{\partial d^U}{\partial w} < 0$. In the general case when both private and GSL credit is available, direct inspection reveals that $w^{G+L}_{\min}(a) < \min\left\{w^G_{\min}(a), w^L_{\min}(a)\right\}$. As with $w^L_{\min}(a)$, the threshold $w^{G+L}_{\min}(a)$ can be decreasing in a and may even be negative.

Proof of Lemma 1. Implicit differentiation of (4) yields $\frac{dh^U(a)}{da} = -\frac{f'[h^U(a)]}{af''[h^U(a)]} > 0$. Using expression (5), define

$$F \equiv u'\left[w + d - h^U(a)\right] - \beta R u'\left[af\left[h^U(a)\right] - Rd\right] = 0.$$

From the implicit function theorem $\frac{\partial d^U(a,w)}{\partial a} = -\frac{\partial F}{\partial a} / \frac{\partial F}{\partial d}$, then

$$\frac{\partial d^U(a,w)}{\partial a} = \frac{\partial h^U(a)}{\partial a} + \beta R \frac{u''\left[af\left[h^U(a)\right] - Rd\right] f\left[h^U(a)\right]}{u''\left[w + d - h^U(a)\right] + \beta R^2 u''\left[af\left[h^U(a)\right] - Rd\right]} > \frac{\partial h^U(a)}{\partial a} > 0,$$

where we have used $af'\left[h^U(a)\right] = R$. Similarly,

$$\frac{\partial d^U(a,w)}{\partial w} = -\frac{u''\left[w + d - h^U(a)\right]}{u''\left[w + d - h^U(a)\right] + \beta R^2 u''\left[af\left[h^U(a)\right] - Rd\right]} = -\frac{1}{1 + \beta R^2 \frac{u''[af[h^U(a)] - Rd]}{u''[w + d - h^U(a)]}}.$$

Since the denominator is greater than one, the argument is complete. ■

Proof of Proposition 1. From the FOC define

$$F \equiv -u'(w + d_0 - h) + \beta a f'[h] u'[af(h) - Rd_0] = 0.$$

The second order condition implies $\partial F / \partial h < 0$, which, combined with implicit differentiation, implies that $sign\left\{\frac{\partial h}{\partial w}\right\} = sign\left\{\frac{\partial F}{\partial w}\right\}$ and $sign\left\{\frac{\partial h}{\partial a}\right\} = sign\left\{\frac{\partial F}{\partial a}\right\}$. First, we have $\frac{\partial h}{\partial w} > 0$ since $\frac{\partial F}{\partial w} = -u''(w + d_0 - h) > 0$. Second,

$$\begin{aligned}\frac{\partial F}{\partial a} &= \beta f'[h] u'[af(h) - Rd_0]\left\{1 + af(h)\frac{u''[af(h) - Rd_0]}{u'[af(h) - Rd_0]}\right\} \\ &< \beta f'[h] u'[af(h) - Rd_0]\left\{1 - 1/\eta[af(h) - Rd_0]\right\},\end{aligned}$$

[57] Since AFQT percentile scores increase with age in the NLSY79, we determine an individual's quartile based on year of birth. AFQT percentile scores in the NLSY97 have already been adjusted to account for age differences.

where the first results from direct derivation, and the second from $u' > 0$, $u'' < 0$, $f' > 0$, $d_0 \geq 0$, and the definition of IES$\equiv \eta(\cdot)$. If $\eta(c) \leq 1 \ \forall \ c > 0$, the right-hand-side (RHS) of the last line is non-positive and $\frac{\partial F}{\partial a} < 0$. ■

Proof of Proposition 2. Using the FOC for the exogenous constraint model,

$$\hat{a}(w) \equiv \sup\{\hat{a} : u'(w) \geq \beta \hat{a} f'[d_{\max}] u'[\hat{a} f(d_{\max}) - R d_{\max}]\},$$

which in principle could be $+\infty$. If $u(c) = c^{1-\sigma}/(1-\sigma)$, then $\hat{a}(w)$ is finite and given by $\hat{a} : w(\beta f'[d_{\max}])^{\frac{1}{\sigma}} = (\hat{a})^{\frac{\sigma-1}{\sigma}} f(d_{\max}) - R d_{\max} (\hat{a})^{\frac{-1}{\sigma}}$. If $\sigma > 1$ (IES < 1), the RHS is strictly increasing and unbounded, so $\hat{a}(w)$ is finite. The rest is direct upon examination of optimality conditions under the three different cases. ■

Proof of Lemma 2. Part (i) is from direct inspection based on the thresholds as derived above. For part (ii), use the FOC for a constrained person with $a > \bar{a}$ (i.e. $d_g = d_{\max}$, $d_p = \kappa a f(h)$ and $h > d_{\max}$) to define

$$F(h, d_{\max}, \kappa) \equiv (\kappa a f'(h) - 1) u'[w + d_{\max} + \kappa a f[h] - h] + \beta a f'(h)(1 - \kappa R) u'[a f(h)(1 - \kappa R) - R d_{\max}].$$

For constrained agents, with $a > \bar{a}$, we have that $u'(c_0) > \beta R u'(c_1)$ and $a f'(h) < R$. It is straightforward to verify that $\frac{\partial F}{\partial d_{\max}} > 0$, and $\frac{\partial F}{\partial \kappa} > 0$, and therefore, implicit differentiation implies the state results.■

Proof Proposition 3. Part (1): If $a > \bar{a}$, the FOC is given by

$$F = u'(c_0)\left[\kappa a f'(h) - 1\right] + \beta u'(c_1) a f'(h)(1 - \kappa R) = 0,$$

where $c_0 = w + \kappa a f(h) + d_{\max} - h$ and $c_1 = a f(h)(1 - \kappa R) - R d_{\max}$. Moreover, notice that $\frac{\partial c_0}{\partial h} = \kappa a f'(h) - 1 < 0$, and $\frac{\partial c_1}{\partial h} = a f'(h)(1 - \kappa R) > 0$. To prove (i) notice that if the agent is constrained, then $u'(c_0) > \beta R u'(c_1)$. Therefore, $F = 0$ implies $\left[1 - \kappa a f'(h)\right] < \frac{a f'(h)}{R}(1 - \kappa R) \implies a f'(h) > R$, i.e. there is under-investment. To prove (ii), notice that $\text{sign}\{\frac{dh}{dw}\} = \text{sign}\{\frac{dF}{dw}\}$ and that $dF/dw = u''(c_0)\left[\kappa a f'(h) - 1\right] > 0$. To prove (iii), first define for any $a \geq \bar{a}$ and $h \geq d_{\max}$ the fraction of labor earnings needed to pay back the maximum debt from the GSL: $\varrho(a, h) \equiv \frac{R d_{\max}}{a f(h)} = \varrho(a) \frac{f(d_{\max})}{f(h)}$, where $\varrho(a)$ is defined in the text. Next, compute the derivative, re-group, simplify, and use the definition of $\eta(\cdot)$ and $\varrho(a, h)$

$$\frac{\partial F}{\partial a} = u'(c_0) \kappa f'(h) + \left(1 - \kappa a f'(h)\right)\left[-u''(c_0)\right] \kappa f(h) + (1 - \kappa R) \beta f'(h) u'(c_1)\left[1 - \frac{1}{\eta(c_1)} \frac{1}{1 - \varrho(a, h)}\right]. \tag{19}$$

Since the agent is constrained, we have that $u'(c_0) \geq \beta R u'(c_1)$. Using this inequality in the first term and ignoring the second term because it is always positive, obtain

$$\frac{\partial F}{\partial a} \geq \beta u'(c_1) f'(h)\left(R\kappa + (1 - \kappa R)\left[1 - \frac{1}{\eta(c_1)} \frac{1}{1 - \varrho(a, h)}\right]\right).$$

The RHS is positive when $\eta(c_1) > \frac{1 - \kappa R}{1 - \varrho(a, h)}$, which is stated as sufficient condition (a). Next, impose sufficient condition (b) on $\frac{dF}{da}$ in equation (19). Since $\beta R \leq 1$, we have $c_0 < c_1$ and $u'(c_0) \geq u'(c_1)$. Take $u'(c_1) f'(h) > 0$ as a common factor, and in the second term use the FOC implied equality $u'(c_1) = \frac{\left(1 - \kappa a f'(h)\right)}{\beta(1 - \kappa R) a f'(h)} u'(c_0)$. Also, divide and multiply by c_0 and simplify to obtain:

$$\begin{aligned}
\frac{\partial F}{\partial a} &= u'(c_1) f'(h)\left\{\kappa \frac{u'(c_0)}{u'(c_1)} + \frac{\kappa\left(1 - \kappa a f'(h)\right) f(h)}{f'(h)} \frac{-c_0 u''(c_0)}{c_0 \frac{(1 - \kappa a f'(h)) u'(c_0)}{\beta(1 - \kappa R) a f'(h)}} + \beta(1 - \kappa R)\left(1 - \frac{1}{\eta(c_1)\left[1 - \varrho(a, h)\right]}\right)\right\}, \\
&\geq u'(c_1) f'(h)\left\{\kappa + \frac{\kappa\beta}{\eta(c_0)} \frac{1}{1 - \varrho(a, h)} + \beta(1 - \kappa R)\left[1 - \frac{1}{\eta(c_1)\left[1 - \varrho(a, h)\right]}\right]\right\},
\end{aligned}$$

where the second line uses the definition of $\varrho(a, h)$, the fact that $c_1 \geq c_0$ and that $u'(c_0) \geq u'(c_1)$. Finally, since by assumption $\eta(\cdot)$ is increasing,

$$\frac{\partial F}{\partial a} \geq u'(c_1) f'(h)\left\{\kappa + \frac{\kappa\beta}{\eta(c_0)} \frac{1}{1 - \varrho(a, h)} + \beta(1 - \kappa R)\left[1 - \frac{1}{\eta(c_0)\left[1 - \varrho(a, h)\right]}\right]\right\}.$$

This inequality holds whenever the term inside brackets is positive, which holds if $\eta(c_0) \geq \frac{1}{1-\varrho(a,h)} \frac{1-\kappa(R+1)}{1+\kappa(\beta^{-1}-R)}$. Since $\varrho(a,h) \leq \varrho(a)$, the sufficiency of condition (b) follows.

Part (2): The FOC in this case is given by

$$F \equiv u'(c_0)\kappa a f'(h) + \beta u'(c_1)\left[a f'(h)(1-\kappa R) - R\right] = 0,$$

where $c_0 = w + \kappa a f(h)$ and $c_1 = a f(h)(1-\kappa R) - Rh$. To prove (i) notice that if the agent is constrained, then $u'(c_0) > \beta R u'(c_1)$ and $F = 0$ implies that $\kappa a f'(h) < \frac{R - a f'(h)(1-\kappa R)}{R}$. Re-arranging, we get $a f'(h) < R$, or equivalently $h > h^U(a)$ because of the strict concavity of $f(\cdot)$. To prove (ii), compute $\frac{dF}{dw} = \kappa a f'(h) u''(c_0) < 0$. The result follows from implicit differentiation ($\frac{\partial h^{G+L}(a,w)}{\partial w} = -\frac{\partial F/\partial w}{\partial F/\partial h}$) and the second order condition ($\partial F/\partial h < 0$). Similarly, for (iii) sign$\left\{\frac{\partial h^{G+L}(a,w)}{\partial a}\right\}$ = sign $\left\{\frac{\partial F}{\partial a}\right\}$. First, compute the derivative

$$\frac{\partial F}{\partial a} = u'(c_0)\kappa f'(h) + \kappa a f'(h) u''(c_0)\frac{\partial c_0}{\partial a} + \beta u'(c_1)\left[f'(h)(1-\kappa R)\right] + \beta\left[a f'(h)(1-\kappa R) - R\right] u''(c_1)\frac{\partial c_1}{\partial a}.$$

Notice that only the second term in this expression can be negative. Take $\frac{1}{a}$ as a common factor and then add and subtract $R\beta u'(c_1)$ to get:

$$\frac{\partial F}{\partial a} = \begin{array}{c} \frac{1}{a}\left\{u'(c_0)\kappa a f'(h) + \beta u'(c_1)\left[a f'(h)(1-\kappa R) - R\right]\right\} \\ + \frac{1}{a}\left\{R\beta u'(c_1) + \kappa a f'(h) u''(c_0)\frac{\partial c_0}{\partial a} a + \beta\left[a f'(h)(1-\kappa R) - R\right] u''(c_1)\frac{\partial c_1}{\partial a} a\right\}. \end{array}$$

The FOC implies that the first line equals zero. Take $R\beta u'(c_1)$ as common factor and multiply and divide the second term by $u'(c_0)$:

$$\frac{\partial F}{\partial a} = \frac{R\beta u'(c_1)}{a}\left\{1 + \frac{1}{R}\left(\frac{u'(c_0)\kappa a f'(h)}{\beta u'(c_1)}\right)\frac{u''(c_0)}{u'(c_0)}\frac{\partial c_0}{\partial a} a + \frac{1}{R}\left[a f'(h)(1-\kappa R) - R\right]\frac{u''(c_1)}{u'(c_1)}\frac{\partial c_1}{\partial a} a\right\}.$$

Because of the FOC, the expression inside parentheses in the second term equals $\left[R - a f'(h)(1-\kappa R)\right]$. After dividing and multiplying the last two terms by c_0 and c_1, respectively, and using the definition of the IES, $\eta(c_i) \equiv -\frac{c_i u''(c_i)}{u'(c_i)}$ and re-grouping:

$$\frac{\partial F}{\partial a} = \frac{R\beta u'(c_1)}{a}\left\{1 + \left[1 - \frac{a f'(h)}{R}(1-\kappa R)\right]\left[\frac{1}{\eta(c_1)}\left(\frac{\partial c_1}{\partial a}\frac{a}{c_1}\right) - \frac{1}{\eta(c_0)}\left(\frac{\partial c_0}{\partial a}\frac{a}{c_0}\right)\right]\right\}.$$

The term $1 - \frac{a f'(h)}{R}(1-\kappa R) > 0$, since there is over-investment, i.e. $a f'(h) < R$. Therefore, $\frac{\partial F}{\partial a}$ can only be negative if $\frac{1}{\eta(c_1)}\left(\frac{\partial c_1}{\partial a}\frac{a}{c_1}\right) - \frac{1}{\eta(c_0)}\left(\frac{\partial c_0}{\partial a}\frac{a}{c_0}\right)$. However, notice that since $\frac{\partial c_1}{\partial a}\frac{a}{c_1} = \frac{a f(h)(1-\kappa R)}{a f(h)(1-\kappa R) - Rh} > 1$ and $\frac{\partial c_0}{\partial a}\frac{a}{c_0} = \frac{\kappa a f(h)}{w + \kappa a f(h)} < 1$, this possibility is ruled out if $\eta(c_1) \leq \eta(c_0)$, which clearly holds if $\eta(\cdot)$ is constant. ■

C Proofs and Other Aspects of the Quantitative Model

Thresholds. Let $m^U(a) \equiv \Phi_{[P,R]} a \left[h^U(a)\right]^\alpha \frac{\Theta_{[S,P]}}{\Theta_{[S,T]}} + e^{r(P-S)}\frac{\Theta_{[P,T]}}{\Theta_{[S,T]}} h_I^U(a)$. Then, $d^U(a,w) = m^U(a) - e^{r(P-S)}\Theta_{[P,T]}/\Theta_{[S,T]} w$, and:

$$\begin{aligned} w^X(a) &\equiv e^{-r(P-S)}\frac{\Theta_{[S,T]}}{\Theta_{[P,T]}}\left[m^U(a) - d_{\max}\right], \\ w^L(a) &\equiv e^{-r(P-S)}\frac{\Theta_{[S,T]}}{\Theta_{[P,T]}}\left[m^U(a) - \kappa_L \Phi_{[P,R]} a\left[h^U(a)\right]^\alpha\right], \\ w^G(a) &\equiv e^{-r(P-S)}\frac{\Theta_{[S,T]}}{\Theta_{[P,T]}}\left[m^U(a) - \min\left\{e^{r(P-S)} h^U(a), d_{\max}\right\}\right]. \end{aligned}$$

For our baseline model:

$$w^{G+L}(a) \equiv e^{-r(P-S)}\frac{\Theta_{[S,T]}}{\Theta_{[P,T]}}\left[m^U(a) - \kappa_1 \Phi_{[P,R]} a\left[h^U(a)\right]^\alpha - \kappa_2 d_{\max}\right].$$

Derivation of the Credit Constraints. A non-defaulting individual retains access to formal credit markets, is able to optimally smooth consumption, and attains post-graduation lifetime utility:

$$V_P^R\left(a,h,d_g,d_p\right)=\Theta_{[P,T]}\frac{\left\{\left[\Phi_{[P,R]}ah^{\alpha}-d_g-d_p\right]/\Theta_{[P,T]}\right\}^{1-\sigma}}{1-\sigma}. \tag{20}$$

On the other hand, by defaulting on any private debt d_p, this individual would attain utility:

$$\begin{aligned}V_P^D(a,h,d_g,r(\cdot;d_g))=&\int_P^{P+\pi}e^{-\rho(t-P)}\frac{\left[(1-\gamma)\,ah^{\alpha}E\left(t-P\right)-r\left(t;d_g\right)\right]^{1-\sigma}}{1-\sigma}dt\\&+e^{-\rho\pi}\Theta_{[P+\pi,T]}\frac{\left\{\left[\Phi_{[P+\pi,R]}ah^{\alpha}-e^{r\pi}\left(d_g-R\left(P+\pi,d_g\right)\right)\right]/\Theta_{[P+\pi,T]}\right\}^{1-\sigma}}{1-\sigma},\end{aligned} \tag{21}$$

where $R\left(P+\pi,d_g\right)=\int_P^{P+\pi}e^{-r(t-P)}r\left(t;d_g\right)dt$ is the cumulative re-payments to GSL debt d_g from P to $P+\pi$. The first term is the discounted utility during the punishment period and the second the discounted utility post-punishment (when the individual has a fresh start and can fully smooth consumption).

Assume that for a period equal to or longer than the length of default punishment π, repayments to GSL loans are given by $r\left(t,d_g\right)=\delta ah^{\alpha}E\left(t-P\right)$, i.e. the individual must pay a constant fraction of his earnings.[58] Then, $R\left(P+\pi,d_g\right)=\delta\Phi_{[P,P+\pi]}ah^{\alpha}$ and the post-punishment balance of GSL debt is $e^{r\pi}\left(d_g-\delta\Phi_{[P,P+\pi]}ah^{\alpha}\right)$. Even under this restriction, we can investigate the interaction between the pace of repayments of GSL loans with repayment incentives and credit constraints of private debt. At one extreme is the "fastest" repayment case when $\delta=\delta_{fast}=d_g/\left(\Phi_{[P,P+\pi]}ah^{\alpha}\right)$ and all GSL debt must be repaid during the punishment period. This is the most disruptive case and is only relevant if earnings are high enough to cover the debt and leave positive consumption during the punishment period (i.e. $d_g/\Phi_{[P,P+\pi]}ah^{\alpha}<1-\gamma$). The attainable utility of a defaulting individual is

$$V_P^D(a,h,d_g,\delta_{fast})=\Delta\frac{\left[(1-\gamma)\,ah^{\alpha}-d_g/\Phi_{[P,P+\pi]}\right]^{1-\sigma}}{1-\sigma}+e^{-\rho\pi}\left(\Theta_{[P+\pi,T]}\right)^{\sigma}\frac{\left[\Phi_{[P+\pi,R]}ah^{\alpha}\right]^{1-\sigma}}{1-\sigma},$$

where $\Delta\equiv\int_P^{P+\pi}e^{-\rho(t-P)}E\left(t-P\right)^{1-\sigma}dt$. At the opposite extreme is the case of "slowest" repayment in which no repayment is made while the individual is being punished, i.e. $\delta=\delta_{slow}=0$. All GSL debt is rolled-over to the post-punishment period, leading to a balance of $e^{r\pi}d_g$ at time $P+\pi$. This case is relevant only if $\Phi_{[P+\pi,R]}ah^{\alpha}>e^{r\pi}d_g$. It leads to utility

$$V_P^D(a,h,d_g,\delta_{slow})=\Delta\frac{\left[(1-\gamma)\,ah^{\alpha}\right]^{1-\sigma}}{1-\sigma}+e^{-\rho\pi}\left(\Theta_{[P+\pi,T]}\right)^{\sigma}\frac{\left[\Phi_{[P+\pi,R]}ah^{\alpha}-e^{r\pi}d_g\right]^{1-\sigma}}{1-\sigma},$$

which, in general, is higher than $V_P^D(a,h,d_g,\delta_{fast})$, because repayments are scheduled in a way that minimizes the disruption of consumption smoothing.

In general, for intermediate values of δ, we can use (21) with the condition $V^R\geq V^D$ to obtain a closed form for the constraint on private credit:

$$d_p\leq\Phi_{[P,R]}ah^{\alpha}-d_g-\left[M_0\left(ah^{\alpha}\right)^{1-\sigma}+M_1\left(M_2ah^{\alpha}-e^{r\pi}d_g\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}},$$

with $M_0\equiv\Delta\left[1-\gamma-\delta\right]^{1-\sigma}/\left(\Theta_{[P,T]}\right)^{\sigma}$, $M_1\equiv e^{-\rho\pi}\left(\frac{\Theta_{[P+\pi,T]}}{\Theta_{[P,T]}}\right)^{\sigma}$, and $M_2\equiv\Phi_{[P+\pi,R]}+e^{r\pi}\delta\Phi_{[P,P+\pi]}$. Clearly, private debt limits are positively linked to post-school earnings $\Phi_{[P,R]}ah^{\alpha}$ and negatively linked to the amount of GSL debt d_g. However, as expected from its superior enforcement, GSL debt does not lead to a one-to-one reduction in the capacity to borrow from private lenders as captured by the fact that the CES term in the right-hand-side is negatively related to d_g. Thus, in general, an expansion of the GSL credit limit $d_{\max}$ leads to an overall expansion in available credit.

[58] Given our assumptions, the timing and structure of repayments does not matter if the agent does not default.

For our baseline case, we set $\delta = \delta^* \equiv (1-\gamma)\frac{e^{r\pi}\left(d_g - \delta\Phi_{[P,P+\pi]}ah^{\alpha}\right)}{\Phi_{[P+\pi,R]}ah^{\alpha}}$. In this case, the fraction $(1-\gamma-\delta^*)$ of income (net of garnishments) available to consume during the punishment period is equal to the fraction $\frac{\Phi_{[P+\pi,R]}ah^{\alpha} - e^{r\pi}\left(d_g - \delta\Phi_{[P,P+\pi]}ah^{\alpha}\right)}{\Phi_{[P+\pi,R]}ah^{\alpha}}$ of the present value of labor earnings (net of GSL debt payments) available for consumption during the post-punishment periods. Imposing this equality, we can write

$$V_P^D(a,h,d_g,\delta^*) = \Delta\frac{\left[(1-\gamma)\left(ah^{\alpha} - \frac{d_g}{\Phi_{[P,R]}-\gamma\Phi_{[P,P+\pi]}}\right)\right]^{1-\sigma}}{1-\sigma} + e^{-\rho\pi}\left(\Theta_{[P+\pi,T]}\right)^{\sigma}\frac{\left[\Phi_{[P+\pi,R]}\left(ah^{\alpha} - \frac{d_g}{\Phi_{[P,R]}-\gamma\Phi_{[P,P+\pi]}}\right)\right]^{1-\sigma}}{1-\sigma},$$

Define $\Theta_D \equiv \Delta(1-\gamma)^{1-\sigma} + e^{-\rho\pi}\left[\frac{\Theta_{[P+\pi,T]}}{\Phi_{[P+\pi,R]}}\right]^{\sigma}\Phi_{[P+\pi,R]}$ and factorize $\left[ah^{\alpha} - \frac{d_g}{\left[\Phi^y_{(P,R)}-\gamma\Phi^y_{(P,P+\pi)}\right]}\right]^{1-\sigma}$, then

$$V_P^D(a,h,d_g,\delta^*) = \Theta_D\frac{\left[ah^{\alpha} - \frac{d_g}{\left[\Phi_{[P,R]}-\gamma\Phi_{[P,P+\pi]}\right]}\right]^{1-\sigma}}{1-\sigma}.$$

This expression and the condition $V^R \geq V^D$ leads to the formula $d_p \leq \kappa_1\Phi_{(P,R)}ah^{\alpha} + \kappa_2 d_g$ in the text, where

$$\kappa_1 \equiv 1 - \frac{\Theta_{[P,T]}}{\Phi_{[P,R]}}\left(\frac{\Theta_D}{\Theta_{[P,T]}}\right)^{\frac{1}{1-\sigma}} \text{ and } \kappa_2 = \left(\frac{\Theta_D}{\Theta_{[P,T]}}\right)^{\frac{1}{1-\sigma}}\frac{\Theta_{[P,T]}}{\Phi_{[P,R]}-\gamma\Phi_{[P,P+\pi]}} - 1.$$

Direct inspection of these formulas verifies that $0 \leq \kappa_1 \leq 1$, $\kappa_2 > -1$, as well as the other properties stated in the text. Finally, when GSL loans must repaid within $[P,Q]$ for $\pi < Q < P$, the private credit constraint becomes

$$d_p \leq \Phi_{P,R}ah^{\alpha} - d_g - \left(\frac{\Delta}{\left[\Theta_{P,T}\right]^{\sigma}}\left[(1-\gamma)ah^{\alpha} - \frac{d_g}{\Phi_{P,Q}}\right]^{1-\sigma} + e^{-\rho\pi}\left[\frac{\Theta_{P+\pi,T}}{\Theta_{P,T}}\right]^{\sigma}\left[\Phi_{P+\pi,R}ah^{\alpha} - e^{r\pi}\left(1 - \frac{\Phi_{P,P+\pi}}{\Phi_{P,Q}}\right)d_g\right]^{1-\sigma}\right)^{\frac{1}{1-\sigma}}.$$

In the text, we consider $M \equiv Q - \pi = 15$.

Proof Proposition 4. All three items in Part 1 and items (i) and (ii) of Part 2 follow virtually the same lines as in Proposition 3 of the two-period case. We proceed to prove item (iii). Our case of interest is when $a > \bar{a}$, and $h > h^* \equiv h_0 + (1+s)d_{\max} = \left[\alpha(1+s)\bar{a}e^{-r(P-S)}\Phi_{[P,R]}\right]^{\frac{1}{1-\alpha}}$. In this case, if the individual is constrained, then $d_g = d_{\max}$ and $d_p = \kappa_1\Phi_{[P,R]}ah^{\alpha} + \kappa_2 d_{\max}$. As in the two-period model, for $a > \bar{a}$ and $h > h^*$, define $\varrho(a,h)$ as the fraction of life-time labor earnings an individual must pay to cover the maximum debt from the GSL, $\varrho(a,h) \equiv \frac{d_{\max}}{\Phi_{[P,R]}ah^{\alpha}} \leq \varrho(a) \leq \frac{\alpha e^{-r(P-S)}}{1+\frac{h_0}{(1+s)d_{\max}}} < \alpha < 1$. To keep notation manageable, define $C_0 = \left(w + e^{-r(P-S)}\left[\kappa_1\Phi_{[P,R]}ah^{\alpha} + (\kappa_2+1)d_{\max}\right] - h_I\right)/\Theta_{[S,P]}$ and $C_1 = \left(\Phi_{[P,R]}ah^{\alpha}(1-\kappa_1) - (\kappa_2+1)d_{\max}\right)/\Theta_{[P,T]}$. The problem for a constrained agent is entirely in terms of h_I :

$$\max_{\{h_I\}}\Theta_{[S,P]}\frac{[C_0]^{1-\sigma}}{1-\sigma} + e^{-\rho(P-S)}\Theta_{[P,T]}\frac{[C_1]^{1-\sigma}}{1-\sigma}.$$

Since $a > \bar{a}$, the solution is interior and given by the FOC

$$F \equiv C_0^{-\sigma}\left[\alpha\kappa_1\Phi_{[P,R]}ah^{\alpha-1}(1+s) - 1\right]e^{-r(P-S)} + e^{-\rho(P-S)}C_1^{-\sigma}\alpha\Phi_{[P,R]}ah^{\alpha-1}(1-\kappa_1)(1+s) = 0.$$

As before, the relationship between ability and investment is given by $\frac{dF}{da}$, which is:

$$\begin{aligned}\frac{\partial F}{\partial a} &= C_0^{-\sigma}\left[\alpha\kappa_1\Phi_{[P,R]}h^{\alpha-1}\right]e^{-r(P-S)}(1+s) - \sigma C_0^{-(\sigma+1)}\left[\alpha\kappa_1\Phi_{[P,R]}ah^{\alpha-1}(1+s) - 1\right]e^{-r(P-S)}\frac{\partial C_0}{\partial a} \\ &+ e^{-\rho(P-S)}C_1^{-\sigma}\alpha\Phi_{[P,R]}h^{\alpha-1}(1-\kappa_1)(1+s) - \sigma e^{-\rho(P-S)}C_1^{-\sigma-1}\alpha\Phi_{[P,R]}ah^{\alpha-1}(1-\kappa_1)(1+s)\frac{\partial C_1}{\partial a}.\end{aligned}$$

Using F=0 and factorizing $\Psi \equiv e^{-\rho(P-S)}C_1^{-\sigma}\alpha\Phi_{[P,R]}h^{\alpha-1}(1-\kappa_1)(1+s) > 0$,

$$\frac{\partial F}{\partial a} = \Psi\left\{\left[\frac{C_1}{C_0}\right]^{\sigma}e^{(\rho-r)(P-S)}\frac{\kappa_1}{1-\kappa_1} + \sigma\frac{\partial C_0}{\partial a}\frac{a}{C_1}\frac{C_1}{C_0} + 1 - \sigma\frac{\partial C_1}{\partial a}\frac{a}{C_1}\right\},$$

where the second term has been multiplied and divided by C_1. Since the individual is constrained, $C_1/C_0 > e^{\left[\frac{r-\rho}{\sigma}\right](P-S)}$ and $(C_1/C_0)^{\sigma} > e^{[r-\rho](P-S)}$. Using these inequalities in the second and first term, respectively, and then simplifying:

$$\frac{\partial F}{\partial a} > \Psi \left\{ \frac{1}{1-\kappa_1} + \sigma \left(e^{\left[\frac{r-\rho}{\sigma}\right](P-S)} \frac{\partial C_0}{\partial a} \frac{a}{C_1} - \frac{\partial C_1}{\partial a} \frac{a}{C_1} \right) \right\}.$$

Using the definitions of C_0 and C_1, compute $\frac{\partial C_0}{\partial a} \frac{a}{C_1} = \frac{\Theta_{[P,T]}}{\Theta_{[S,P]}} \frac{\kappa_1 e^{-r(P-S)}}{[1-\kappa_1-(\kappa_2+1)\varrho(a,h)]}$ and $\frac{\partial C_1}{\partial a} \frac{a}{C_1} = \frac{1-\kappa_1}{[1-\kappa_1-(\kappa_2+1)\varrho(a,h)]}$. Plug these expressions in and simplify to obtain

$$\frac{\partial F}{\partial a} > \Psi \left[\frac{1}{1-\kappa_1} + \frac{\sigma \left(\kappa_1 \frac{\Theta_{[S,T]}}{\Theta_{[S,P]}} - 1 \right)}{[1-\kappa_1-(\kappa_2+1)\,\varrho\,(a,h)]} \right].$$

The RHS is positive iff

$$\sigma \left(\kappa_1 \frac{\Theta_{[S,T]}}{\Theta_{[S,P]}} - 1 \right) \geq -\frac{[1-\kappa_1-(\kappa_2+1)\,\varrho\,(a,h)]}{1-\kappa_1}, \tag{22}$$

which holds if $\kappa_1 \geq \frac{\Theta_{[S,P]}}{\Theta_{[S,T]}}$, i.e. sufficient condition (a), because by construction $1-\kappa_1-(\kappa_2+1)\,\varrho\,(a,h) > 0$ when $a > \bar{a}$ and $h > h^*$. If $\kappa_1 \left(\frac{\Theta_{[S,T]}}{\Theta_{[S,P]}} \right) < 1$, dividing both sides of (22) by it leads to $\sigma \leq \left[1 - \frac{(\kappa_2+1)}{1-\kappa_1} \varrho\,(a,h) \right] \left[1 - \kappa_1 \frac{\Theta_{[S,T]}}{\Theta_{[S,P]}} \right]^{-1}$. The claim for sufficient condition (b) holds because $\varrho\,(a,h) \leq \varrho\,(a)$. ■

References

Aiyagari, R., J. Greenwood, and A. Seshadri (2002), "Efficient Investment in Children", *Journal of Economic Theory*, 102(2), 290–321.

Albuquerque, R., and H. Hopenhayn (2004), "Optimal Lending Contracts and Firm Dynamics", *Review of Economic Studies*, 71(2), 285–315.

Alvarez, F., and U. Jermann (2000), "Efficiency, Equilibrium, and Asset Pricing with Risk of Default", *Econometrica*, 68(4), 775–97.

Andolfatto, D. and M. Gervais, (2006). "Human Capital Investment and Debt Constraints," *Review of Economic Dynamics*, 9(1), 52-67.

Becker, G. (1975), *Human Capital*, 2nd Ed., New York, NY: Columbia University Press.

Brown, M., J.K. Scholz, and A. Seshadri, (2007) "A New Test of Borrowing Constraints for Education," Working paper.

Browning, M., L.P. Hansen, and J.J. Heckman (1999), "Microdata and General Equilibrium Models", in J.B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, vol. 1, Amsterdam: Elsevier Science.

Belley, P., and L. Lochner (2007), "The Changing Role of Family Income and Ability in Determining Educational Achievement", *Journal of Human Capital*, 1(1), 37-89.

Berkner, L. (2000), *Trends in Undergraduate Borrowing: Federal Student Loans in 1989-90, 1992-93, and 1995-96*, Washington DC: National Center for Education Statistics Report 2000-151.

Cameron, S., and J.J. Heckman (1999), "Can Tuition Policy Combat Rising Wage Inequality?" in M. Kosters (ed.), *Financing College Tuition: Government Policies and Educational Priorities*, Washington: American Enterprise Institute Press.

Cameron, S., and J.J. Heckman (1998), "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males", *Journal of Political Economy*, 106, 262-333.

Cameron, S., and J.J. Heckman (2001), "The Dynamics of Educational Attainment for Black, Hispanic and White Males", *The Journal of Political Economy*, 109, 455-499.

Cameron, S., and C. Taber (2004), "Estimation of Educational Borrowing Constraints Using Returns to Schooling", *The Journal of Political Economy*, 112, 132-182.

Card, D. (1995) "Earnings, Schooling, and Ability Revisited", in S. Polachek (ed.), *Research in Labor Economics*, vol. 14, Greenwich, CT: JAI Press.

Carneiro, P., and J.J. Heckman (2002), "The Evidence on Credit Constraints in Post-Secondary Schooling", *Economic Journal*, 112, 989-1018.

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

Cawley, J., J. Heckman, L. Lochner and E. Vytlacil (2000), "Understanding the Role of Cognitive Ability in Accounting for the Recent Rise in the Economic Return to Education," in K. Arrow, S. Bowles and S. Durlauf (eds.), *Meritocracy and Economic Inequality*, Princeton Univ. Press.

Chatterjee, S. and Ionescu, F. (2009) "Insuring College Failure Risk" Federal Reserve Bank of Philadelphia/Colgate University, working paper.

College Board (2005), *Trends in College Pricing 2005*.

Cordoba, J., and M. Ripoll (2009) "Schooling and Development: The Role of Credit Limits, Public Education, Fertility and Mortality" Mimeo, University of Pittsburgh.

Cunha, F., J. Heckman, L. Lochner, and D. Masterov (2007), "Interpreting the Evidence on Life Cycle Skill Formation", in E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, Amsterdam: Elsevier Science.

Ellwood, D., and T. Kane (2000), "Who Is Getting a College Education? Family Background and the Growing Gaps in Enrollment", in S. Danziger and J. Waldfogel (eds.), *Securing the Future: Investing in Children from Birth to College*, Russell Sage Foundation.

Fernandez-Villaverde, J. and D. Krueger (2004). "Consumption and Saving over the Life Cycle: How Important are Consumer Durables?" Working paper.

Gropp, R., J. K. Scholz, and M. White (1997), "Personal Bankruptcy and Credit Supply and Demand", *The Quarterly Journal of Economics*, 112(1), 217-51.

Hansen, W.L., and B. Weisbrod (2003), "The Distribution of Costs and Direct Benefits of Public Higher Education: The Case of California", *Journal of Human Resources*, 4(2), 176–91.

Hanushek, E., C. Leung, and K. Yilmaz (2003), "Redistribution through Education and Other Transfer Mechanisms", *Journal of Monetary Economics*, 50(8), 1719–50.

Heckman, J., L. Lochner and C. Taber (1998), "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents", *Review of Economic Dynamics*, 1, 1-58.

Heckman, J., L. Lochner, and P. Todd (2008), "Earnings Functions and Rates of Return", *Journal of Human Capital*, 2(1).

Ionescu, F. (2008), "Consolidation of Student Loan Repayments and Default Rates Across Income Groups", *The B.E. Journal of Macroeconomics (Topics)*, 8(1).

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

Kane, T. (2007), "Public Intervention in Postsecondary Education", in E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, Amsterdam: Elsevier Science.

Katz, L., and D. Autor (1999), "Changes in the Wage Structure and Earnings Inequality", in O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, vol. 3, New York: North-Holland.

Keane, M., and K. Wolpin (2001), "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment", *International Economic Review*, 42, 1051–1103.

Kehoe, T., and D. Levine (1993), "Debt-Constrained Asset Markets", *Review of Economic Studies*, 60(4), 865–888.

Kehoe, P. and F. Perri (2002). "International Business Cycles with Endogenous Incomplete Markets", *Econometrica*, 70(3), 907–928.

Kocherlakota, N. (1996), "Implications of Efficient Risk Sharing without Commitment", *Review of Economic Studies*, 63(4), 595–609.

Krueger, D. and F. Perri, (2006). "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies*, 73(1), 163–193.

Lang, K. (1993), "Ability Bias, Discount Rate Bias and the Return to Education", Unpublished Manuscript, Boston University.

Livshits, I., J. MacGee, and M. Tertilt (2007), "Consumer Bankruptcy: A Fresh Start", *American Economic Review*, 97(1), 402–418.

Lochner, L. and A. Monge-Naranjo (2002) "Human Capital Formation with Endogenous Credit Constraints." NBER working paper 8815.

McPherson, M., and M. Schapiro (1998), *The Student Aid Game: Meeting Need and Rewarding Talent in American Higher Education*, Princeton, NJ: Princeton University Press.

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

Monge-Naranjo, A. (2009) "Entrepreneurship and Firm Heterogeneity with Limited Enforcement" *The Annals of Finance*, 5(3-4), 465–494.

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", Working Paper.

Titus, M. (2002), *Supplemental Table Update for Trends in Undergraduate Borrowing: Federal Student Loans in 1989-90, 1992-93, and 1995-96*, http://nces.ed.gov/pubs2000/2000151update.pdf.

Wegmann, C., A. Cunningham, and J. Merisotis (2003), *Private Loans and Choice in Financing Higher Education*, Washington DC: Institute for Higher Education Policy.

Table D1: Total Schooling Costs for each year of college by AFQT Quartile (1999 Dollars)

Years of	Direct	Foregone Earnings:				Total Costs:			
College	Expenditures	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Quartile 1	Quartile 2	Quartile 3	Quartile 4
1	6,322	3,604	8,560	6,716	6,841	9,927	14,882	13,038	13,163
2	12,343	8,689	19,446	17,530	15,476	21,032	31,789	29,873	27,819
3	58,275	14,844	30,467	29,257	26,288	73,119	88,742	87,532	84,563
4	75,880	21,222	40,825	40,350	39,106	97,102	116,704	116,229	114,985
5	92,646	26,606	51,201	51,232	51,300	119,252	143,847	143,878	143,947
6	108,615	31,799	60,135	60,509	61,431	140,413	168,750	169,124	170,045
7	123,822	35,531	67,669	69,302	70,733	159,354	191,491	193,125	194,555
8	138,306	36,243	72,981	76,520	78,758	174,549	211,287	214,826	217,064

Notes:

1) Direct expenditures based on average expenditures per student in all colleges and universities. Expenditures for first two years of college are based on 2-yr school averages for school years 1980-81 to 1984-85. Expenditures for grades 15+ are based on 4-yr school averages for school years 1980-81 to 1989-90. Costs are discounted at a 5% annual interest rate back to grade 12. (Source: Digest of Education Statistics, Table 342, 1999.)

2) Foregone earnings reflect the PV of average earnings relative to someone with 12 years of completed schooling, taking into account earnings during college. See text for details.