The Nature of Credit Constraints and Human Capital*

Lance J. Lochner
University of Western Ontario and NBER

Alexander Monge-Naranjo
Pennsylvania State University

May 2009

Abstract

We develop a human capital model with borrowing constraints explicitly derived from government student loan programs and private lending under limited commitment. This model 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. 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

In this paper, we show that a human capital model with borrowing constraints explicitly derived from Government Student Loan (GSL) programs and private lending under limited commitment can explain a number of important empirical observations in the U.S. higher education sector since the early 1980s. Specifically, we identify four stylized facts about college attendance and student borrowing:

1. a strong positive correlation between ability and college attendance for youth from all economic backgrounds (both early 1980s and early 2000s)

2. a rising importance of family income as a determinant of college attendance (from the early 1980s to early 2000s),

3. an increase in the share of undergraduates borrowing the maximum from government student loan programs (since the early 1990s),

4. a dramatic rise in student borrowing from private lenders (since the mid-1990s).

Our model of limited public and private lending explains all of these major changes, qualitatively and quantitatively, as natural responses to the well-documented rising costs and returns to college in the U.S. over the 1980s and 1990s.[1] We argue that the rising real costs and returns to college have increased the demand for credit among students, but real federal student loan limits have not risen in response (Kane 1999). While GSL programs (combined with generous education subsidies) appear to have provided adequate credit for the vast majority of American youth in the early 1980s (Cameron and Heckman 1998, 2001, Keane and Wolpin 2001, Carneiro and Heckman 2002, Cameron and Taber 2004), this no longer appears to be true. Many more students borrowed the maximum allowable amount from their federal Stafford loans in 1999-2000 compared with a decade earlier (Berkner 2000 and Titus 2002), and family income has become an important determinant of college-going (Belley and Lochner 2007). It is only natural that private lenders would want to step in to help fill the growing gap between student credit demands and government loan limits. The increase in earnings associated with college attendance has enabled students to credibly commit to repay more debt than in the past, which enables private lenders to extend more credit. Yet, the fact that family income has become an important determinant of college attendance implies that the expansion of private credit has not fully met all student demands.

Since Becker (1975), economists have long appealed to credit constraints as an explanation for college attendance gaps by family income (e.g. see Manski and Wise 1983, Elwood and Kane 2000). Yet, the human capital literature has paid little attention to the nature of borrowing constraints. Existing models typically assume that either interest rates increase with the amount borrowed or that there is a fixed maximum amount that individuals can borrow.[2] Both approaches neglect the link between borrowing opportunities and investment decisions that plays a key role in both

[1] See College Board (2005) for evidence on rising costs and Katz and Autor (1999) Heckman, Lochner and Todd (2008) for evidence on rising returns.

[2] Studies assuming variable interest rates (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), Hanushek, Leung, and Yilmaz (2003), and Keane and Wolpin (2001).

GSL programs and private lending as we describe below. We show that without this link, the canonical model of exogenous borrowing constraints predicts a *negative* relationship between ability and human capital investment among constrained borrowers when the consumption intertemporal elasticity of substitution (IES) is less than one. This is troubling, since most empirical estimates of the IES are below one (Browning, Hansen, and Heckman 1999) and a strong positive ability – college attendance relationship exists for all family income and wealth levels in both the NLSY79 and NLSY97 (as we show in Section 3). Additionally, models with exogenous borrowing constraints offer no insights regarding the recent rise in private student lending, the interaction between private and public lending, or how lending opportunities respond to important economic and policy changes. Our framework offers insights on all of these issues.

GSL programs directly tie student credit to the level of investment — students can borrow to help finance college-related expenses only if they are enrolled in school. We show that private lenders, facing limited repayment incentives from borrowers, also link credit limits to the level of investment, as well as observable individual characteristics that affect the returns to investment. These features of actual public and private credit limits help generate a positive relationship between ability and investment (even when the IES is less than one) while still predicting a positive relationship between family resources and investment among constrained youth.

GSL programs have two distinct forms of limits: (i) a pre-specified maximum loan limit, and (ii) an endogenous limit that restricts students from borrowing more than they spend on their education. Youth constrained only by the second limit will invest the same amount in their human capital as if they were completely unconstrained, since they do not face a tradeoff between additional investment and consumption while in school. Consumption decisions may be severely distorted even when schooling and investment decisions are not. Standard empirical tests for borrowing constraints that are based on differences in educational attainment or marginal rates of return to investment by family income (or other categories used to differentiate the 'constrained' from 'unconstrained') will fail to detect this constraint.[3] By introducing the restriction that borrowing cannot exceed investment, GSL programs effectively increase the population of students who invest the unconstrained optimal amount.

The rise in private student lending since the mid-1990s (and drop in the current credit crisis) highlights the importance of understanding how private lenders determine student credit levels. Even if human capital cannot be directly repossessed by lenders, creditors can punish defaulting borrowers in a number of ways (e.g. lowering credit scores, seizing assets, garnisheeing a fraction of labor earnings), which tend to have a greater impact on debtors with higher post-school earnings. Higher ability students who invest more through education will be offered more credit by private lenders, because they can credibly commit to re-pay more given the punishments they face upon default.[4] These mechanisms effectively link private borrowing limits for students to their abilities and human capital investments. This dependence of credit limits on investment and ability generates a positive ability – investment relationship for *all constrained borrowers* under standard parameterizations of preferences.

[3] Thus, evidence that family resources do not affect educational attainment or financial returns does not necessarily imply that credit constraints are non-binding. These common tests will tend to under-estimate the fraction of the population that is constrained as well as the adverse impacts of constraints on welfare.

[4] In the following section, we discuss evidence that private lenders do tend to extend more credit to youth with characteristics and investments that lead to higher expected post-school earnings.

Our model of private lending is related to the literature on endogenous credit constraints, which has generally focused on implications for 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 2008). Our assumed punishments for default are similar to those employed by Livshits, MacGee, and Tertilt (2007) in their analysis of bankruptcy over the lifecycle. Andolfatto and Gervais (2006) study human capital accumulation with limited commitment, but they focus on the optimal set of intergenerational transfers and not on cross-sectional implications for investment.

An economy with both GSL programs and private lending under limited commitment will be characterized by a positive ability – investment relationship under plausible assumptions about punishments for private loan default. Furthermore, investment will be decreasing in family resources for youth constrained by both upper GSL loan limits and private lending limits. Thus, this framework is consistent with key education patterns in the NLSY79 and NLSY97 data. The standard model with exogenous constraints cannot simultaneously explain the positive ability – college attendance and income – attendance patterns in the NLSY97.

The rest of the paper proceeds as follows. In the next section, we describe borrowing opportunities from U.S. GSL programs and the recent emergence of private lenders. In the third section, we discuss evidence on the relationship between ability, family income and college attendance in the U.S. and briefly survey the literature on the prevalence of credit constraints. In Section 4, we develop a simple two-period human capital investment model to analytically compare the cross-sectional implications for borrowing and investment under alternative assumptions about credit markets. Section 5 extends our framework with public and private lending to a lifecycle model that incorporates government subsidies for education. We calibrate this model using U.S. data on schooling, ability, government subsidies, and post-school earnings in the early 1980s. Simulating the recent rise in costs and returns to schooling, we show that our model with both public and private lending does a good job reproducing observed cross-sectional human capital investment and private borrowing patterns for the early 1980s and 2000s. The model with exogenous constraints does not. Section 6 concludes.

2 Available Sources of Credit

In this section, we briefly review the main sources of credit in the U.S. for college education. We begin with a description of key institutional features of government student loan (GSL) programs. We then discuss private student loan programs, their recent expansion, and highlight the similarities and differences between GSL and private 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. 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.[5]

Historically, private lenders have provided the capital to student borrowers (and their parents) under the Stafford and PLUS programs, while the government guarantees those loans with a promise to cover any unpaid amounts. However, since the 1994-95 academic year, the federal government has directly provided loans to some students under the same rules and terms.[6] While Stafford loans are disbursed to students, PLUS loans can be taken out by parents to help cover the costs of their children's schooling. The Perkins Loan Program provides an additional source of government funds to students most in need; however, its loan offerings depend on the level of program funding at the post-secondary institution attended by a student.

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, students cannot borrow from GSL programs to finance non-schooling related consumption goods or activities. Second, student loan programs set cumulative and annual loan upper limits on the total amount of credit available for each student. Thus, students face exogenously pre-specified limits for their borrowing from U.S. federal loan programs.[7] Third, loans covered by GSL programs typically have extended enforcement rules compared to unsecured private loans. For instance, students cannot expunge payment of student loans via bankruptcy.

Table 1 reports loan limits (based on the dependency status and class level of each student) for Stafford and Perkins student loan programs for the period 1993-2007.[8] In recent years, 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 would 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.[9] Student borrowers can defer loan re-payments until six (Stafford) to nine (Perkins) months after leaving school.

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 NLSY97 and NLSY79 respondents we study below 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.[10] The relative stability of real GSL

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

[6] 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. Most students under age 24 are considered dependent, so their determined need is a decreasing function of their parents' income. 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.

[7] 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 less other financial aid received by the student.

[8] Stafford loan limits for freshman, sophomores, and graduate students increased slightly in July, 2007.

[9] 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.

[10] From 1982-83 to 2002-03, Stafford borrowing limits for undergraduates declined by 44% for first-year students

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

	Stafford Loans Dependent Students	Stafford Loans 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.
** Subsidized Stafford loan amounts can be no greater than the borrowing limits for dependent students; independent students can also borrow unsubsidized Stafford loans provided that their total (subsidized and subsidized)loan amount is not greater than the independent student limits.
*** Cumulative graduate loan limits include loans from undergraduate loans.

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 went up 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).

As mentioned above, 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 generally be expunged through bankruptcy. If a suitable re-payment plan is not agreed upon with the lender once a 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.[11] 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.[12]

2.2 Private Lending

Historically, private financing of higher education has been relatively minor and mostly restricted to professional school students (e.g. law, business administration and medicine), whose post-graduation earnings (and the ability to repay loans) 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). 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.[13]

The rise in borrowing from private lenders outside the Stafford and Perkins Loan Programs suggests that the GSL limits are no longer enough to satisfy many students' demands for credit.[14] 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 design of private lending programs is broadly consistent with the problem of lending under limited repayment incentives. Private lenders directly and indirectly link credit to educational investment expenditures and (sometimes) to student earnings potential. Most directly, 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,

and 25% for second-year students, while they increased by about 20% for college students enrolled in years three through five. Throughout 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 our NLSY respondents would have began attending graduate school.

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

[12] Since the early 1990s, the government has also begun to punish educational institutions with high student default rates by making their students ineligible to borrow from federal lending programs.

[13] These figures do not include student borrowing on credit cards, which has also increased considerably over this period. See College Board (2005).

[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.

they vary substantially in their terms and eligibility requirements.[15] 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.[16] 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) recently reported that in response to the current credit crisis in the U.S., a number of private lenders have discontinued lending to students at community colleges and lower quality four-year institutions, while they have continued to lend to students at higher quality schools where graduates are expected to earn more after school.

Enforcement of private student loans is regulated by U.S. bankruptcy code. In filing for bankruptcy under Chapter 7, former students could (until very recently) discharge all private student loan obligations after leaving school.[17] 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, which affects future access to credit and may spill over into other consumer domains (e.g. landlords often request credit reports before renting to potential tenants). 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.

3 The Role of Ability and Family Income on College Attendance

In this section, we discuss the empirical relationship between family income, cognitive ability and college attendance in the U.S. We briefly review the recent literature and data from the NLSY79 and NLSY97 and offer evidence documenting three stylized facts on college attendance. First, in the early 1980s, there was a weak link between family income and college attendance. Second, for recent student cohorts, there is a much stronger relationship between family income (or wealth) and college attendance. Third, in both the early 1980s and the early 2000s, there has been 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.[18]

[15] Among those that limit borrowing to the cost of schooling less financial aid, most use a broader concept of schooling costs than do GSL programs. Specified maximum loan limits are also generally quite high, especially for students in professional schools.

[16] For example, the relatively new private lending institution, 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." Lenders also generally offer more credit (and sometimes better terms) to student borrowers with a creditworthy co-signer.

[17] 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.

[18] 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

A number of empirical studies using the NLSY79 data have shown that family income played little role in college attendance decisions in the U.S. during the early 1980s. 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 small differences in college enrollment rates and other college-going outcomes by family income after accounting for differences in family background and AFQT.[19]

More recently, Belley and Lochner (2007) show that family income has become a much more important determinant of college attendance in the early 2000s.[20] Youth from high income families in the NLSY97 are sixteen 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 nearly twice the effect observed in the NLSY79. The combined effects of family income and wealth are even more dramatic in the NLSY97. 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 important changes in the relationship between family resources and college attendance, the relationship between ability and schooling has remained relatively stable over time. Figure 1 shows college attendance rates by AFQT quartiles and either family income or family wealth quartiles in the NLSY79 and NLSY97.[21] For all family income or wealth categories 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.

That ability is positively related with schooling is usually taken for granted by economists. However, as we discuss below, the standard exogenous borrowing constraints model predicts a *negative* relationship between ability and educational attainment for constrained youth (under empirically relevant assumptions about preferences). Therefore, a close examination of the ability – college attendance relationship by family income and wealth is warranted. Figure 1 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 for youth from both the lowest family income and wealth quartiles is 47%, which is actually larger than the 37% gap among those from both the highest family income and wealth quartiles. These patterns are inconsistent with a negative ability – schooling relationship among constrained youth as long as the fraction of youth constrained is decreasing in family income and wealth.[22]

additional details.

[19] A few other studies explore different features of the NLSY79 data and argue that credit constraints had little effect on educational outcomes in the early 1980s (e.g. Cameron and Taber 2004, Keane and Wolpin 2001).

[20] 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.

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

[22] We observe similar patterns in the NLSY97 for age 20 enrollment in four-year colleges/universities conditional

Of course, AFQT scores may be correlated with other family background variables that influence college attendance decisions conditional on family resources. We, therefore, control for a host of other family background measures in addition to AFQT quartiles using ordinary least squares. Table 2 reports the estimated effects of AFQT (these estimates reflect the difference in attendance rates between the reported AFQT quartile and AFQT quartile 1) on college attendance after controlling for family background characteristics.[23] Results are reported for separate regressions by family income or wealth quartile. The estimates confirm the general patterns observed in Figure 1: ability has strong positive effects on college attendance for all family income and wealth quartiles in both NLSY samples.

4 Modelling Student Credit

We first use a two-period model to examine how the nature of public and private credit markets shapes the behavior of human capital investment. We show that contrary to a canonical exogenous borrowing constraints model, incorporating the key features of public and private student lending sources produces investment – wealth and investment – ability relationships that are qualitatively consistent with the empirical patterns discussed above.

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 $u(\cdot)$ satisfies:

Assumption 1. $u : \mathbb{R}_+ \to \mathbb{R}$ *is strictly increasing, strictly concave, twice continuously differentiable and* $\lim_{c \searrow 0} u'(c) = +\infty$.

Each individual is endowed with financial assets $w \geq 0$ and ability $a > 0$. Initial assets capture all familial transfers while ability represents all 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.[24]

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 difference is 41%, while it is 52% for the highest income quartile.

[23] We control for the following: gender, race/ethnicity, mother's education, 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.

[24] In an online appendix, we derive investment – wealth and investment – ability relationships under 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 'altruistic' model, all of our key results hold with the sole reinterpretation of initial wealth w as the parent's wealth rather than the child's wealth (consistent with our empirical evidence). In the 'warm-glow' model, parental transfers are entirely determined by parental wealth and preferences for giving. The lending environment has no impact on transfers, and all of our results go through without qualification (interpreting w as either parent's wealth or total transfers to children). Finally,

Labor earnings at $t = 1$ are $y = af(h)$, where h is schooling investment and $f(\cdot)$ satisfies:

Assumption 2. $f : \mathbb{R}_+ \to \mathbb{R}_+$ *is strictly increasing, concave, twice continuously differentiable,* $\lim_{h \searrow 0} f'(h) = +\infty$ *and* $\lim_{h \nearrow \infty} f'(h) = 0$.

Note that both a and h enhance earnings and that they are complementary with each other.[25] Assumptions 1 and 2 are standard, and we make use of them without further reference. They imply that optimal solutions are interior (positive and finite) and uniquely determined by first order conditions.

Human capital investment, h, is in units of the consumption good.[26] 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 h to maximize the present value of lifetime net resources, $w - h + af(h)/R$:

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

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, borrowing an amount $d^U(a, w)$ to satisfy the Euler equation:

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

Therefore, unconstrained borrowing is strictly decreasing in wealth and increasing in ability. For a fixed w, optimal debt is strictly increasing in ability because of two reasons. First, more able individuals wish to finance a larger investment. Second, for any given level of investment h, more able individuals earn higher net lifetime income and would want to consume more early on.The latter force implies that borrowing increases more quickly in ability than does human capital investment. The following lemma states these findings:

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* $-1 < \frac{\partial d^U(a,w)}{\partial w} < 0$.

under 'paternalistic' preferences, results for exogenous constraints and private lending under limited commitment hold subject to a few additional minor conditions. Results for the GSL hold under a slight modification of the tied-to-investment constraint.

[25] Our assumptions 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 standard theoretical models of human capital (e.g. the widely used Ben-Porath (1967) model). Below, we briefly discuss the implications of relaxing the assumption of a unitary elasticity of substitution.

[26] While the model formally abstracts from foregone earnings, it is isomorphic to one in which foregone earnings for any given investment amount, h, are independent of ability.

Proofs for all results and other analytical details of this section are provided in Appendix ??.

We make repeated use of Lemma 1 to characterize the behavior of investment with borrowing constraints. Since whenever unconstrained, individuals invest $h^U(a)$ and borrow or lend $d^U(a, w)$, our discussion will be focused on the behavior of constrained individuals.

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.[27] 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 ($w \geq w^X_{\min}(a)$). It turns out that this threshold is increasing in ability. Then, constrained persons are those of high ability or low wealth. Appendix ?? explains further the function $w^X_{\min}(a)$ as well as the threshold defined by the other constraints considered in this section.

Individuals constrained by (EXC) have exhausted their ability to bring resources to the early (investment) period. In setting their human capital investment they must strike a balance between increasing lifetime earnings and smoothing consumption over time. For them, optimal investment $h^X(a, w)$ is uniquely determined by

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

the equality between the present marginal cost of investing and its future marginal returns.

We highlight four empirically relevant results of this model. First, constrained investment never exceeds unconstrained investment. Second, constrained investment is strictly increasing in wealth. Third, the return on human capital investment exceeds the return on borrowing/saving for constrained individuals. Fourth, the relationship between constrained investment and ability depends on the intertemporal elasticity of substitution (IES), $-u'(c) / [cu''(c)]$. More formally:

Proposition 2 *Consider individuals whose wealth w falls below the threshold $w^X_{\min}(a)$ 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), (ii) and (iii) are well-known (see 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. Lang

[27] See, for example, Aiyagari, Greenwood, and Seshadri (2002), Belley and Lochner (2007), Caucutt and Kumar (2003), 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.

(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.

Our interest is in part (iv), i.e. that the model predicts a *negative* relationship between ability and investment for an IES below one.[28] This is a serious shortcoming of the model, because most estimates of the IES fall below one (see Browning, Hansen, Heckman 1999) and, as we showed earlier, a positive relationship between ability and investment is a robust empirical regularity. The intuition behind this counterfactual implication of the model is simple. For constrained individuals, the relationship between ability and investment is driven by two opposing forces. The first is an intertemporal substitution effect: A higher ability encourages investment by raising the returns of human capital investment. The second is a life-time income effect: A higher ability reduces investment since with a higher lifetime income the individual wants to increase consumption and, if constrained, that can only be done by reducing investment. With strong preferences for intertemporal consumption smoothing (e.g. IES$\leq$1) the second effect dominates.

Aside from the counter-factual implication for cross-section behavior of human capital investments, this result can also impair the ability of the model to replicate the aggregate behavior of investment. Indeed, this canonical exogenous constraints model, predicts a reduction in the human capital investments of constrained individuals as a response to an economy-wide increased rate of return to college education. We now show that incorporating some crucial features of the sources of financing for human capital is sufficient to overturn this negative result.

4.4 Government Student Loan Programs

Now, consider an economy in which student credit is governed by the key features of GSL programs described in Section 2. First, lending is *tied to investment* and cannot be used to finance non-schooling related consumption goods or activities:

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

In the absence of other sources of credit, (TIC) is equivalent to $c_0 \geq w$, the requirement that early consumption must be self-financed by the individual or his family. Second, borrowing is constrained by an upper limit $0 < d_{\max} < \infty$ for the total credit to each student:

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

This second constraint is effectively the same as the exogenous constraint above. The overall credit limits induced by GSL programs are:

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

Third, the government has enhanced disciplining devices to enforcement repayment of GSL credit. To capture this feature we assume that GSL is fully enforceable. We now characterize the behavior of human capital investment in this environment, using the superscript G to designate them.

To isolate the role of (TIC), first assume that it is the only constraint individuals face.[29] In this case, individuals are unconstrained as long as desired investment exceeds desired borrowing.

[28] An $IES \leq 1$ is only a sufficient condition for a negative ability – investment relationship. More generally, the model may predict a negative relationship for IES values greater than one. Result (iv) holds more generally in a model with foregone earnings as long as youth wage rates are not strictly decreasing in ability.

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

Because desired investment is increasing in ability, the (TIC) constraint is less stringent than (EXC) for higher ability individuals and more stringent for those with lower ability.

When $d = h$, borrowing finance investment (and no more) and early consumption equals initial wealth. The problem boils down to choosing h so as to

$$\max_h \{u(w) + \beta u[af(h) - Rh]\},$$

which is equivalent to maximizing $af(h) - Rh$, and investment coincides with amount $h^U(a)$. By itself, (TIC) does not lead to conflict between smoothing consumption and maximizing net lifetime resources because credit cannot be use for anything other than investment. Despite potentially large distortions in the consumption overtime, 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 always fail to capture the consumption distortions and conclude that borrowing constraints are non-binding or inexistent. *A simple but robust message of this analysis is that empirical tests on the impact of credit constraints must also incorporate measures of consumption over time.*

Now consider the full GSL constraint (GSLC), and, to facilitate the comparison with the X model, assume that $d_{\max} = d_0$. Appendix **??** shows that the threshold $w^G_{\min}(a)$ is increasing in ability and that $w^G_{\min}(a) \geq w^X_{\min}(a)$ when $d_{\max} = d_0$, because the GSL imposes more constraints that the X model.

Unconstrained individuals, i.e. $w \geq w^G_{\min}(a)$, are those who possess relatively high assets given their ability, and will include very rich individuals or very low ability youth in the population. The remaining population of constrained individuals will fall in one of three categories. The first one is composed of lower ability persons who are constrained by (TIC) only. They invest the unrestricted level $h^U(a)$ but would like to borrow to increase consumption while in school. The other two groups are more able individuals who would like to invest more than $d_{\max}$. From these, one group is constrained only by the upper limit on borrowing (6). They borrow up to the maximum $d_{\max}$ and invest beyond that using some of their initial assets For them, investment coincides with $h^X(a, w)$, because (TIC) is slack. Finally, if $h^X(a, w)$ is decreasing in a, there may be a third group which is composed of very poor but very high ability youth, which find themselves constrained by both (6) and (TIC). Individuals in this group borrow and invest the maximum amount, $d_{\max}$, consuming their initial wealth while in school. They invest below the unrestricted amount, but if they get an additional dollar they would consume it and not invest it since their consumption while in school is much lower than their future earnings.

The following proposition summarizes the previous discussion:

Proposition 3 *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 $\hat{w} : [\bar{a}, \infty) \to \mathbb{R}_+$ 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:*

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

Regardless of the IES, there is always a region in which $h^G(a, w)$ is increasing in ability and independent of initial wealth; there may also be a region in which it is constant and equal to $d_{\max}$.[30] If utility has an IES ≤ 1, there is a region (of middle-high abilities) in which investment decreases with ability as in the exogenous constraint model, but the additional constraint (TIC) shrinks this region relative to the exogenous constraint case alone.

Figures **??** and **??** 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 **??** displays investment and borrowing behavior for two low levels of wealth, $\bar{w}$ and $w_L < \bar{w}$, while Figure **??** illustrates investment behavior for a higher level of wealth $w_H > \bar{w}$.[31]

We highlight three important points about investment under the GSL as exemplified in these figures. First, investment under the GSL equals the unconstrained level for a larger range of middle ability and low/middle wealth individuals than under exogenous constraints (e.g. individuals with wealth w_L and ability $a \in (a_2, \bar{a}]$ in Figure **??**). The constraint (TIC) encourages investment for those who would like to borrow more than they spend on schooling. This implies a positive relationship between investment and ability and no relationship between investment and wealth for a broader range of ability and wealth levels. Second, among higher ability and middle/high wealth individuals, the (TIC) restriction ensures that investment never falls below $d_{\max}$. With an IES less than one, this shrinks the range of abilities for which investment is negatively related to ability (e.g. individuals with ability $a > a_4$ in Figure **??**). 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 **??**). Altogether, the implied investment – ability and investment – wealth relationships in the GSL model are more closely aligned with the empirical findings discussed earlier due to the additional constraint that borrowing cannot exceed investment. In particular, the set of individuals whose investment declines with ability is smaller than in the traditional exogenous constraint model.

While the extra (TIC) restriction imposed by GSL programs adversely affects utility and early consumption levels, it encourages investment in human capital. The following lemma compares allocations and utility under the GSL, exogenous constraints, and unconstrained models.

Lemma 4 *Impose $d_0 = d_{\max}$ and let $\{h^m, c_0^m, c_1^m, U^m\}$, denote the optimal allocations and attained utilities in models $m = U, X, G$ for arbitrarily fixed $(a, w) \in \mathbb{R}_+^2$. Then:*

$$h^U \geq h^G \geq h^X, \qquad c_0^U \geq c_0^X \geq c_0^G, \qquad c_1^G \geq c_1^U \geq c_1^X, \qquad U^U \geq U^X \geq U^G,$$

and any of the inequalities is strict if the extra constraint between a pair of models is binding.

[30] If $h^X(a, w)$ is always increasing in a (e.g. for an IES sufficiently greater than one), then $h^G(a, w)$ is globally increasing in both arguments. The characterization is as follows: $h^G(a, w) = h^U(a)$, for $a \leq \bar{a}$ or $w \geq w_{\min}^X(a)$; $h^G(a, w) = d_{\max}$ for $a > \bar{a}$ and $w < \hat{w}(a)$ and $h^G(a, w) = h^X(a, w)$ otherwise. The flat region where investment equals $d_{\max}$ may not exist.

[31] Note that $\bar{w} \equiv w_{\min}^G(\bar{a})$ 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. $h^U(\bar{a}) = d_{\max}$).

Figure 3: d^U, h^U, h^X, and h^G for low wealth individuals ($w \leq \bar{w}$)

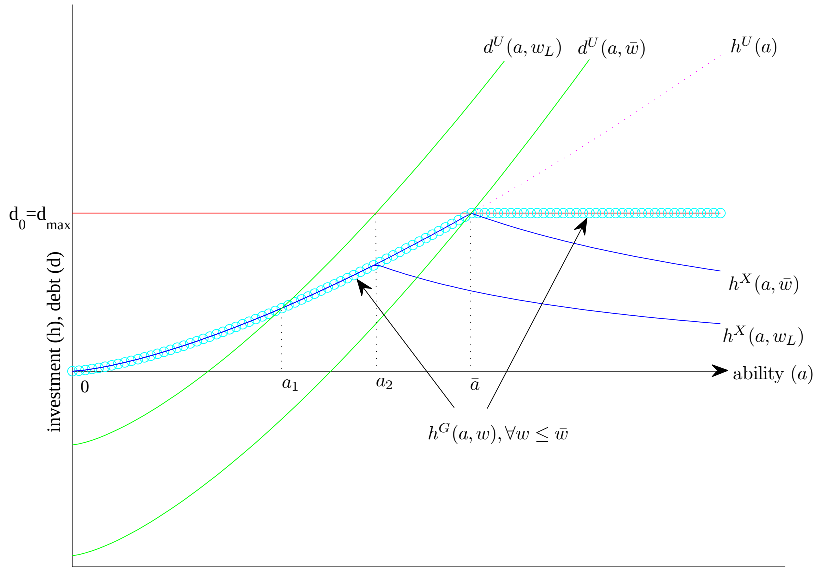
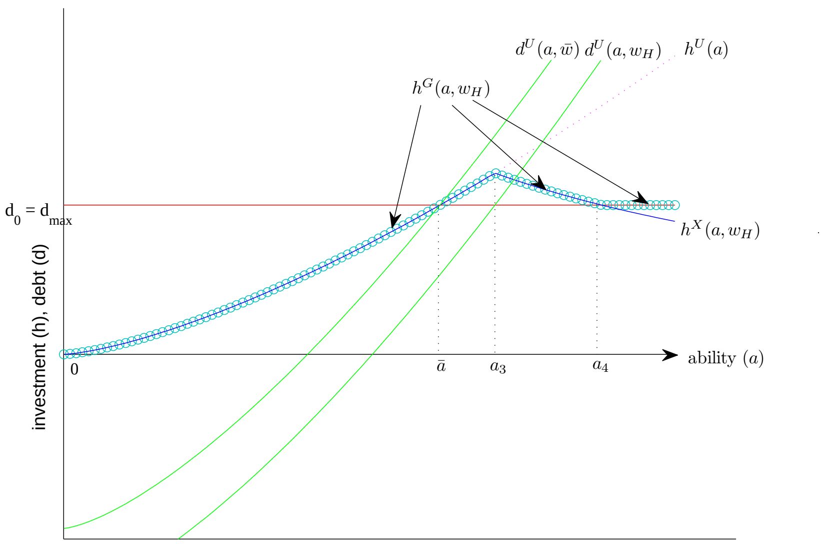


Figure 4: d^U, h^U, h^X, and h^G for high wealth individuals ($w > \bar{w}$)



4.5 GSL Programs and Private Lenders

Private lending has recently emerged as an important source of funding for many college students in the U.S. Therefore, our baseline model allows private credit markets to co-exist with GSL credit. We assume that private lenders are competitive but face limited repayment incentives. These limited repayment incentives are due to the inalienability of human capital and the lack of other forms of collateral by many students. We preserve our assumption that the repayment of GSL loans is fully enforced to capture the superior repayment enforcement of GSL loans vis-à-vis private loans described in Section 2.

A rational borrower repays private loans if and only if the cost of repaying is lower than the cost of defaulting. In turn, these limited incentives can be foreseen by rational lenders who, in response, limit their supply of credit.[32] 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 — the credit offered to an individual would be directly related to his perceived future earnings. But since expected earnings are determined by ability and investment, private credit limits and investments will be co-determined in equilibrium.

In the richer life-cycle model of next section we derive a simple expression for the maximum amount of credit based on empirically based assumptions of exclusions of formal credit markets. In the two-period model we capture a similar form of credit limit by assuming that private lenders punish garnishee a fraction $\tilde{\kappa} \in (0, 1)$ of the labor earnings defaulting borrowers.[33] Implied repayment decisions are simple: borrowers repay (principal plus interest on a private debt d_p) if the payment Rd_p is less than the punishment cost $\tilde{\kappa} a f(h)$. Foreseeing this, borrowing from private lenders is limited to:

$$d_p \leq \kappa a f\left(h\right), \tag{7}$$

where $\kappa \equiv R^{-1}\tilde{\kappa} < R^{-1}$.

It is convenient to start by considering the allocations with only private lenders, i.e. $d_{\max} = 0$. The wealth threshold $w_{\min}^{L}(a)$ above which individuals are unconstrained can be decreasing in a if punishments are severe enough because more able individuals would be interested in repaying higher debts. It is possible that more able persons may be unconstrained while the least able are constrained. It is also possible that individuals at the top end of the ability distribution may be unconstrained regardless of their of their financial wealth.

Now, consider both public and private lending together. Borrowing from the GSL, denoted by d_g, must satisfy (GSLC). If a person must re-pay GSL loans regardless of whether he defaults on private loans, then the constraint on private borrowing is still given by (7). Therefore, the constraint on total borrowing constraint is given by

$$d_g + d_p \leq \min\left\{h, d_{\max}\right\} + \kappa a f\left(h\right). \tag{8}$$

We refer to this case with the superscript $G + L$.

The threshold $w_{\min}^{G+L}(a)$ of assets below which individuals are constrained despite the two sources of credit is below both $w_{\min}^{G}(a)$ (GSL only) and $w_{\min}^{L}(a)$ (private lenders only). The threshold can be decreasing in a so the least able may be constrained and the most able unconstrained. The threshold

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

[33] Penalty avoidance actions like re-locating, working in the informal economy, borrowing from loan sharks, or renting instead of buying a home are all costly to those who default and would contribute to $\tilde{\kappa}$.

is decreasing in $d_{\max}$ and κ because higher values of those parameters increase the amount of credit available.

Next, we provide sufficient conditions on punishments and preferences which ensure that optimal investment $h^{G+L}(a,w)$ is exhibits a positive relationship with ability.

Assumption 3. *Either of the following holds: (i) The IES is uniformly bounded below by* $(1-\kappa R)$*, or (ii) the IES is non-decreasing in consumption,* $\beta R \geq 1$*, and,* $IES\,(c_0) \geq 1-(1+R)\,\kappa$*.*

Proposition 5 *Consider individuals whose wealth* w *is below the threshold* $w_{\min}^{G+L}(a)$*,so constraint (8) binds. First, if* $a \leq \bar{a}$*, then* $h^{G+L}(a,w) \geq h^{U}(a)$*, and for* (w,a) *low enough the inequality is strict. Second, 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*; and (iii) if Assumption 3 holds* $h^{G+L}(a,w)$ *is strictly increasing in* a*.*

The first part of the proposition shows that for low ability individuals, available sources of credit suffice to attain unrestricted investment levels and $h^{U}(a)$ is a lower bound to their human capital investment. There is an important difference with respect to an economy with GSL credit only. The interaction of between GSL and private credit can lead a person with low wealth and low ability to invest above the unconstrained level $h^{U}(a)$. We call this situation *"over-investment"* since the marginal return of human capital $af\prime(h)$ is below the returns on financial savings R. Over-investment occurs only for individuals with little financial resources that borrow from the GSL to invest in human capital beyond $h^{U}(a)$ which in turn increases their early consumption by enhanced ability to borrow $\kappa af(h)$ from private lenders. Appendix **??** contains the details and provides an illustration.

The second part of the proposition shows that higher ability indviduals, for whom the GSL credit is not enough to attain $h^{U}(a)$, the existence of private lenders allows for higher investment. The responsiveness of private credit limits to ability and investment not only creates a tendency for more able persons to be unconstrained, but it can also generate a positive ability – investment relationship. Moreover, under the conditions of Assumption 3, the investment of those agents would be a strictly increasing function of their ability a and their family resources w. The combination of GSL programs with private lending with limited commitment leads to a model that is consistent with patterns observed in the data. We show in Section 5 that this general conclusion holds true in the context of a richer life-cycle economy calibrated with *any* empirically reasonable values of the IES.

The interaction between public and private lending also enhances the investments of constrained persons with high ability:

Proposition 6 *Let* $h^{G+L}(a,w;d_{\max},\kappa)$ *denote optimal investment with both the GSL (characterized by upper loan limit* $d_{\max} > 0$*) and private markets with limited commitment (characterized by enforcement level* $\kappa > 0$*). Then: (i)* $w_{\min}^{G+L}(a) < \min\left\{w_{\min}^{G}(a), w_{\min}^{L}(a)\right\}$*; and (ii)* $\frac{\partial h^{G+L}(a,w;d_{\max},\kappa)}{\partial d_{\max}} \geq 0$ *and* $\frac{\partial h^{G+L}(a,w;d_{\max},\kappa)}{\partial \kappa} \geq 0$ *with strict inequality for both if* $a > \bar{a}$ *and* $w < w_{\min}^{G+L}(a)$*.*

Private and GSL credit both help reducing the number of persons whose investments are constrained. However, the the expansion of either has differential impacts on individuals depending on their ability levels. For example, for high ability individuals an expansion of the GSL limit

$d_{\max}$ might have a more than proportional impact on their investment, because in addition to the one-to-one impact of a larger GSL loan it would lead private lenders to increase their offer of credit for both consumption and further investments. However, a higher $d_{\max}$ would have no impact on the consumption or investments of lower ability individuals. An increase in private credit (i.e. an increase in κ) would allow for greater consumption for all individuals during the investment period, but it would only affect the investments of the most able.The presence of private lenders also enhances the responses in investment (among constrained individuals) to changes in GSL lending.

5 Quantitative Analysis

In this section we explore the quantitative implications of our model with public and private lending for schooling in the U.S. To this end we extend our two-period model to a multi-period setting which we calibrate using data on schooling costs, earnings, and other features of the U.S. economy in the early 1980s. Then, we compare the performance of the model in replicating the empirical patterns in Section 3.

5.1 A Multiperiod Model

The lifespan of all individuals is given by an interval $[S, T]$ of time. Lives are divided into three stages: "Youth", $t \in [S, P]$, when individuals invest $x(t)$ in school; "maturity," $t \in [P, R]$, when they work full-time earning $y(t)$; and "retirement," $t \in [R, T]$, when they consume from accumulated savings. Concretely, S stands for the start date of schooling, P for the entry date of participation in labor markets and R for the date of retirement.

Preferences are standard. As of any $t_0 \in [0, T]$, the utility of an individual is

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

where $c(t)$ is consumption at t, $\sigma > 0$ is the inverse of the IES, and $\rho > 0$ is a subjective discount rate.

We assume competitive financial markets. The market interest rate is $r \le \rho$. Consumption and investment decisions are restricted by the lifetime budget constraint

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

Here $w \ge 0$ indicates the individual's own financial wealth as of $t = S$.[34] Individuals are also endowed with an ability $a \ge 0$, which may reflect genetic traits, early family investments, and other characteristics that affect the returns on investment.

Regardless of their investments, all individuals have a minimum human capital $h_0 \ge 0$ at the time they enter the labor market at $t = P$. This minimum could be seen as the average human

[34] 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.

capital of individuals who do not attend college, which, and among other things, embeds free of charge government investments accessible by all youth. In our baseline specification we assume that this minimum human capital is uniform across abilities. However, we will also explore the quantitative results when $h_0(a)$ is a function of the ability of individuals.

In addition to h_0, individuals can invest to augment the stock of human capital with which they will enter the labor market. An investment flow $x(t)$ 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. With the purpose of taking the model to the data we explicitly assume that the government matches every unit of privately financed investment with a subsidy rate of $s \geq 0$. Altogether, total investment in human capital at the time of labor market entry is

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

In what follows, we ignore the timing of investment flows $x(t)$ and focus on accumulated private (h_I) and total (h) human capital investment.[35] As a normalization, the units of h, h_I,and $h_0(a)$ are in present value terms of $t = S$.

Labor earnings at date t, $y(t)$, depend positively on individual ability a, schooling investments, and the experience $E(t-P)$ accumulated since entry into labor markets at date P. Specifically, we assume

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

where $0 < \alpha < 1$ and $E(t-P) = \exp\left(\int_P^t g_s ds\right)$ where g_s is the rate of growth of earnings at date s.[36] 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 cumulation of the present value factors, as of date t_0, of the labor income that accrues between t_0 and t_1. With this definition, the present value of lifetime labor income, as of date $t = S$ is $e^{-r(P-S)} \Phi_{[P,R]} a h^{a}$.

5.2 Unrestricted Allocations

The presence of frictionless financial markets allows individuals to fully smooth consumption and maximize the present value of lifetime labor earnings net of investments costs

$$PVLI = \left\{ e^{-r(P-S)} \Phi_{[P,R]} a \left[h_0 + (1+s)\, h_I\right]^{\alpha} - h_I \right\}.$$

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 the publicly provided amount, therefore their human capital is equal to the lump-sum h_0 provided by the government. Those with $a > a_0$ will equate the marginal return on human capital investment with its private marginal cost of investing. The resulting unconstrained level of human capital investment is, therefore, given by

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

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

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

[36] Our main theoretical results readily extend to the case where $f(a,h)$ is a general function that satisfies Assumption 2 above and g_s is increasing in a (i.e. more able individuals have steeper wage profiles).

The optimal consumption path grows at the constant rate $\frac{r-\rho}{\sigma}$ over the life-cycle. For an individual with ability a and initial wealth w, expenditures for consumption and schooling investments during $[S,P]$ imply a level of debt (if positive; savings if negative) at labor market entry ($t=P$) equal to

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

where the function $\Theta_{[t_0,t_1]}\equiv\int_{t_0}^{t_1}e^{\left[\frac{r-\rho}{\sigma}-\rho\right](t-t_0)}dt$ is the cumulation of present value factors along the unconstrained optimal path for consumption between dates t_0 and t_1.

This function $d^U(a,w)$ shares the same essential properties of its equivalent in the two-period model, i.e. $-1<\frac{\partial d^U(a,w)}{\partial w}<0$ and $\frac{\partial d^U(a,w)}{\partial a}>\frac{\partial h^U(a)}{\partial a}\geq 0$.

5.3 Borrowing Constraints

Before introducing restrictions to the amount of borrowing available during school, it is convenient to describe the utility of the individual at the time of entry into labor markets. Consider an individual with ability a who, after his consumption and investment decisions during his school age, enters the labor market with human capital investment h and financial liabilities d. If he fully repays his debts, the net present value (as of P) of his lifetime resources available for consumption is equal to $\Phi_{[P,R]}ah^\alpha-d$. Optimal consumption smoothing, implies that discounted utility as of date $t=P$ is

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

where we have used the fact that $\Theta_{[t_0,t_1]}=\int_{t_0}^{t_1}e^{\left[\left(\frac{r-\rho}{\sigma}\right)(1-\sigma)-\rho\right](t-t_0)}dt$ for any two dates $t_0<t_1$. Lastly, the superscript R in $V_P^R(a,h,d)$ indicates that is the maximum utility attainable while fully repayment of all debts; the subscript P indicates that it is the discounted utility as of date $t=P$.

5.3.1 Government Student Loan Programs and Private Lending

We consider now the behavior of human capital with the credit limits defined by GSL programs (d_g) and private lending (d_p). Total available credit d is defined by the co-existence of these and given by $d=d_g+d_p$. Credit from GSL programs follows the same guidelines described above, while we continue to assume that private lending is restricted by limited enforcement problems.

As discussed earlier, GSL credit is tied to schooling-related expenses, generally imposes a maximum amount of possible credit, and have superior enforcement. To capture these features we assume that GSL credit limits are given by

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

for some $0<d_{\max}<\infty$. Note that credit is linked lending to to out-of-pocket investment expenses h_I and not to the resulting total investment h. With respect to repayment of debts d_g, we assume that GSL loans have a fixed repayment schedule $r(t;d_g)$ over individual's participation in labor markets, i.e. $\int_P^R e^{-r(t-P)}r(t;d_g)\,dt=d_g$. We assume that GSL loan repayments are fully enforced, and individuals pay it regardless of whether they default or not on private lending.

Private lenders face limited repayment incentives that stem from their own limited ability to punish default. Being unable to collect back the investments h_I, we assume they dispose of two punishments commonly assumed in the literature on consumer bankruptcy (e.g. see Athrea XXX, 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 operates via consumption smoothing and could be quite costly if the IES is low and labor earnings grow quickly with age and experience. The second punishment is that a defaulting borrower forfeits a fraction $\gamma \in [0, 1)$ of his labor earnings for a period following default. The fraction γ encompasses direct garnishments from lenders as well as the costs of actions taken by the borrower to avoid direct penalties (e.g. working in the informal sector, renting instead of owning a house, etc.). We assume that both penalties are active for an interval of length $\pi \in (0, R-P)$ that starts the moment default takes place. These two penalties reinforce each other as a temporary reduction in earnings is particularly problematic for an individual if he is also unable to borrow against future earnings.

To exclusively focus on limitations to finance schooling, we make three additional simplifying assumptions: (1) individuals can only default on private loans at the time of labor market entry; (2) individuals that default on private loans can access frictionless and fully enforceable credit markets after the punishment period; and (3) individuals that choose to repay their private student loans have access to perfect financial markets upon entry into the labor market. In short, we abstract from issues related to the optimal timing of default and the enforcement of post-school loans.[37] These assumptions greatly simplify the analysis and allow us to focus on the interaction GSL loans with the repayment incentives of private loans acquired during school.

As with the two period model, private creditors will restrict credit to amounts that will be repaid. Individuals compare the lifetime utility associated with full loan repayment, given in equation 15), against the lifetime utility associated with default and its associated punishments. For an individual with ability and human capital a and h and with financial liabilities d_g and d_p from GSL and private lenders, respectively, the value of default is

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

The first term reflects the discounted utility acquired during the punishment phase from P to $P + \pi$. During that time the agent cannot borrow or lend, and consumption equals earnings net of garnishments from private lenders and net of GSL loan repayment $r\,(t, d_g)$. The second term reflects the discounted utility acquired after the punishment phase. When entering this phase, the individual may carry some debt from private lenders, $d_g^{P+\pi} = e^{r\pi}\left[d_g - \int_P^{P+\pi} e^{-r(t-P)} r(t, d_g) dt\right]$, but is cleared of all debt from private creditors. During this latter phase, the individual can freely access credit markets again and obtains the utility $V_{P+\pi}^R\left(a, h, d_g^{P+\pi}\right)$ as defined in equation (15). While the value of repayment depends on the total amount of GSL and private debt (but not the timing of GSL repayment), the value of default depends on the actual timing of GSL repayment, $r\,(\cdot; d_g)$, but not on private debt.[38]

[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] 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. Private debt is irrelevant in the case of default, since borrowers are cleared of all debts.

An individual will choose to repay his private debts at the end of school if and only if $V_P^R(a,h,d) \geq V_P^D(a,h,d_g,r(\cdot;d_g))$. This condition implicitly defines the maximum private credit d_p that a borrower would repay (and a lender would grant) as a function of a, h, and d_g, given the GSL repayment schedule $r(\cdot;d_g)$ and punishments π and γ. In this life-cycle model, the repayment of GSL loans can follow many different paths $r(\cdot,d_g)$, which will affect the repayment incentives on private debt d_p and, therefore, limits on private credit. As we show below, the faster the individual has to repay GSL loans, the more costly it is to default on private loans, since higher payments must be made during the punishment period when income is low and the agent is unable to smooth consumption. This implies that shorter GSL repayment schedules will cause private lenders to extend more credit – an interesting result given recent moves in the U.S. to allow GSL recipients to greatly extend their repayment periods through loan consolidation and other options.

Analytically, it is convenient to impose additional structure on allowable repayment $r(t,d_g)$ schedules to further characterize credit limits. In what follows, 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.[39] (The structure of repayments after the punishment period is irrelevant for individuals, since we assume they can borrow and lend freely during that period.) Under this assumption, consumption during the punishment period equals $ah^\alpha E(t-P)(1-\gamma-\delta)$ and the value of GSL debt remaining at the end of the punishment period would be $e^{r\pi}\left(d_g - \delta\Phi_{[P,P+\pi]}ah^\alpha\right)$ in present value terms as of $t = P + \pi$.

If the GSL repayment rate δ is set such that individuals must repay a constant fraction of their net disposable income (i.e. net of any punishments in the case of default) over their full working careers, then the repayment rate $\delta^* = \frac{(1-\gamma)d_g}{\left[\Phi_{[P,R]}-\gamma\Phi_{[P,P+\pi]}\right]ah^\alpha}$ ensures full GSL repayment. Note that this repayment schedule effectively reduces the required GSL repayment rate during the punishment phase following a private student loan default. It also assumes individuals can extend loan repayment for their entire working lives. As such, a repayment rate of δ^* is somewhat 'generous' for GSL borrowers relative to actual repayment plans in the U.S.[40]

To simplify our analysis we assume the GSL repayment rate is δ^*, which, as we detail in Appendix XXX implies the simple private debt constraint of:

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

where $\kappa_1 \in (0,1)$ and $\kappa_2 < 1$ are both constants that depend on preferences (σ,ρ), the interest rate r, the GSL repayment requirement rate δ, and enforcement parameters (γ,π). It is important to note that κ_1 and κ_2 incorporate the effects of punishment for default but not on the values of government subsidies s and the minimum human capital levels h_0.

Private lending limits are directly linked to lifetime earnings just as in the two-period model. However, even if wage garnishment are not allowed ($\gamma = 0$), private lending could be sustained ($\kappa_1 > 0$) as long as for some period ($\pi > 0$) defaulting individuals are excluded from credit markets

[39] We implicitly assume that $\delta \leq \frac{d_g}{\Phi_{(P,P+\pi)}ah^\alpha}$ to ensure that total repayments during the punishment period do not exceed total GSL debt.

[40] Two features of GSL repayment schedules in the U.S. are worth noting here. First, GSL debtors can apply for forbearances, deferrements, and other reductions in GSL payments during periods of 'economic hardship'. Thus, GSL debtors may qualify for reduced payments during the punishment period when earnings are relatively low due to lifecycle forces and wage garnishments reduce available resources. Second, students can typically extend their loan repayment period up to 30 years (depending on the amount borrowed) through loan consolidation and other GSL repayment plans.

and unable to smooth consumption. Indeed, regardless of the value of γ, $\kappa_1 = 0$ only if $\pi = 0$, in which case the model converges to an exogenous constraint model $d_0 = 0$. In general, the amount of sustainable borrowing (the value of κ_1) 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) higher desire to smooth consumption (lower IES, i.e. higher σ), and (iv) less stable are earnings (i.e. the higher their post-graduation growth rate g_s).

Another very important feature of these private lending constraints is that $\kappa_2 < 1$. This implies that private credit does not decrease one-for-one with expansions of government credit and that total student credit is increasing in government loans. In contrast with the two-period model; however, there may be a partial 'crowding out' of private credit with expansions in GSL programs.

5.3.2 The Behavior of Human Capital Investment

Optimal schooling investment decisions maximize the initial discounted utility $U(a, h)$ at $t = S$

$$U(a,h) = \int_S^T e^{-\rho(t-S)} \left[\frac{c(t)^{1-\sigma}}{1-\sigma}\right] dt + e^{-\rho(P-S)} V_P^R(a, h, d_g + d_p)$$

s.t. $h = h_0 + (1+s)\, h_I$, 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 (17) and (16).

We have adopted the constraints defined by equation (17) as our baseline case. We compare our baseline model with (1) a model in which the only source of credit is the GSL program ($d_{\max} > 0$ and $\pi = 0$); (2) a model in which there is no GSL but there are private lenders ($d_{\max} = 0$ and $\pi > 0$), and, (3) a canonical exogenous constraints where total credit is constrained by $d \leq d_{\max}$. We refer to our baseline model with the superscript $G+L$, and to the other three models by the superscripts G, L and X, respectively. We defer to Appendix XX some analytical aspects of the models, including the thresholds of financial wealth, $w_{\min}^{G+L}(a)$, $w_{\min}^{G}(a)$, $w_{\min}^{L}(a)$, and $w_{\min}^{X}(a)$, that define the regions where the respective constraints bind.

Before turning to the quantitative implications of our model, we close this section with a simple yet important result on conditions under which our baseline model leads to a positive relationship between ability and investment for constrained individuals. Define,

$$\underline{\kappa} \equiv \left(\frac{\sigma - 1}{\sigma}\right)\left(\frac{1}{e^{\left(\frac{r-\rho}{\sigma} - r\right)(P-S)}\Theta_{[P,T]} + 1}\right),$$

which the following proposition uses as a lower bound for $\kappa_1$as a function of the parameters of the economy.[41]

Proposition 7 *Consider individuals whose ability a is above the minimum ability which it is privately optimal to invest beyond h_0, i.e. $h_I^U(a) > 0$ and whose wealth w is below the threshold $w_{\min}^{G+L}(a)$,so constraint (8) binds. First, if $a \leq \bar{a}$, then $h^{G+L}(a, w) \geq h^U(a)$, and for (w, a) low enough the inequality is strict. Second, if $a > \bar{a}$, then: (i) $h^{L+G}(a, w) < h^U(a)$; (ii) $h^{G+L}(a, w)$ is strictly increasing in w; and (iii) if $\kappa_1 > \underline{\kappa}$ then $h^{G+L}(a, w)$ is strictly increasing in a.*

[41] This lower bound is also a function of ρ and r; however, we focus on the role of the IES and σ based on our discussion above.

Aside from the possibility of zero private invesment, i.e. $h_I = 0$, the first part of this proposition is identical to the first part of Proposition 5, its two-period counterpart. Likewise, the second part is virtually identical to the second part of Proposition 5, except that the sufficient condition is $\kappa_1 > \underline{\kappa}$. Here, we highlight the fact that such condition enables us to examine whether our model with empirically relevant parameter values can overturn the negative results of an exogenous constraints model and deliver both, a positive relationship between ability and investment and a positive response to aggregate investment to increases in the return to human capital investment.

An important parameter in this regard is the IES. Recall from the discussion in Section 4.3 that strong preferences for smooth consumption (i.e. a high σ, low IES) generate a negative ability – investment relationship when credit constraints are exogenously fixed. This tendency also exists when constraints are endogenous since $\partial\underline{\kappa}/\partial\sigma \geq 0$, which implies that a stronger link between investment and credit limits (i.e. a larger κ_1) is needed to ensure a positive ability – investment relationship as preferences for smooth consumption become stronger. However, a greater preference for smooth consumption profiles also makes the default punishment of exclusion from credit markets more painful. Private lenders will be willing to offer more credit if the cost of defaulting is higher. Thus, κ_1 might be also increasing in σ under limited commitment. In fact, it may be possible that higher values of σ (or lower IES) makes it more rather than less likely that $\kappa_1 > \underline{\kappa}$ holds.

5.4 Parameter Values

We now discuss the parameter values we use to explore the quantitative implications of the model. We set a time interval of unitary length to represent a calendar year. All dollar amounts are denominated in 1999 dollars using the Consumer Price Index (CPI-U). We set some of the parameters values to match the US economy and the others are estimated using data on earnings and educational attainment from the random sample of males in the NLSY79. To measure ability, we use indicator variables for the quartile of the AFQT score of each individuals in our sample. Table **??** reports the values of all parameters used in our baseline model.

Calibrated Parameters			GMM estimates		
Parameter	Value	To match:	Parameter	Value	Coefficient on:
S	19	US Demographics	g_0	0.03	Experience
P	26				
R	65		α	0.70	Schooling investment
T	80		h_0	160,322	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
$d_{\max}$	46,000	GSL Loan Limits	a_2	1.55	2
γ	0.1	Garnishments & other costs	a_3	1.60	3
s	0.526	Subsidy school grades 10+	a_4	1.72	4

With our focus on college education, we assume that youth (investment period) begins at age $S = 19$ and ends at age 25. Maturity (labor market participation period) runs from age $P = 26$ until age 64. Retirement runs from age $R = 65$ until death at age $T = 80$. With these numbers, our model roughly captures the demographics and the timing of college education and labor market decisions in the US and many other countries.

We set the annual interest rate r to be 5% as a compromise between the US historical averages of the risk-less rate (around 4.1%) and of return to capital before taxes (around 5.5%).[42] We also set ρ to be 5%. Local variations in ρ and r, including differences between those rates, only lead to insignificant changes in the results. We complete the calibration of preferences setting $\sigma = 2$ to match an IES of 0.5, an intermediate value in the estimates 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 the U.S. Bankruptcy Code, after filing for Chapter 7, individuals must wait a period of at least 7 years to qualify to file again. On the other hand, default records remain in the credit history of an individual for a period of 10 years. Thus, the relevant values for π are between 7 and 10. In our baseline we set $\pi = 10$, but as we show below, credit constraints are similar for $\pi = 7$. Also, we set $\gamma = 0.1$ for the fraction of lost earnings for individuals who default. Under the GSL program guidelines, defaulting borrowers face an explicit 15% wage garnishment. For private unsecured loans, an explicit garnishment rule does not exist. However, 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.) As we report below, the fractions κ_1 and κ_2 vary little with variations of γ within empirically plausible values. Finally, we calibrate $d_{\max}$ to be US$46,000, which is roughly consistent with the maximum available from GSL programs for independent students.

To set the values that govern the relationship between investment, human capital and earnings we employ a two step procedure which we explain in detail in Appendix D. The first step is to compute the total cost expenditures for the N_i years of investment of each individual i controlling for the ability of the individual. We split our NLSY79 sample according to AFQT ability quartiles indicated by dummy variables D_i^q, for $q = 1, 2, 3, 4$. Total cost of investment includes direct costs (tuition and other costs) and foregone earnings. Direct costs are computed based on current-fund expenditures per full-time equivalent student in all institutions of higher education as reported in the 1999 Digest of Education Statistics (Table 342). Direct expenditures for the first two years of college are based on 2-yr school averages for academic years 1980-91, while direct expenditures for 3+ years of college are based on 4-yr school averages for academic years 1980-1989-90.

Foregone earnings for each year of schooling an individual i attains are computed using the average earnings of working individuals in our NLSY79 sample that have the same AFQT quartile, age and years of schooling. Total schooling expenditures are the sum of direct schooling costs and foregone earnings. In our baseline model, we assume students can work, at least part-time, so we substract from the cost the average earnings reported by students while in school. From earnings we subtract off taxes using an average tax rate of 18%, the average federal plus state taxes from TAXSIM as reported by Feinberg on http://www.nber.org/taxsim/ally/ally.html. All net-of taxes expenditures are discounted to the end of the high school using the same interest rate $r = 5\%$ used to calibrate the model. Total investments $h\left(D_i^q, N_i\right)$ consists of the sum of those discounted expenditures, which, by construction, are in equivalent terms as the model's present-value investments as of $t = S$.

[42] This is the weighted average of the annual returns on equity and corporate bonds in the U.S. over the period 1950-2005, calculated by Lustig, Van Nieuwerburgh, and Syverson (2007).

The second step is to estimate the parameters that link the investments $h\left(D_i^q, N_i\right)$ with observed earnings. We assume that earnings profiles are log-linear in experience, i.e. $E\left(t-P\right) = \exp\left[g_0\left(t-P\right)\right]$. This rather standard assumption leads to an expression for log-earnings $\log y_{i,t-P}$ of the form

$$E\left[\log y_{i,t-P}|Z_i\right] = \sum_{q=1}^{4} \aleph_q D_i^q + \alpha \log\left[h_0\left(D_i^q\right) + h\left(D_i^q, N_i\right)\right] + g_0\left(t-P\right)_i, \qquad (18)$$

where the information set Z_i includes the ability D_i^q, investments $h\left(D_i^q, N_i\right)$ and experience $\left(t-P\right)_i$ of individual i. The parameters $\aleph_q$ provides the basis for a measure (see below) the ability levels of individuals in AFQT scores in quartile q.

A second moment condition can be obtained using the behavior of investments. In particular, under the results of Cameron and Heckman (1998) and Carneiro and Heckman (2002), NLSY79 respondents were not constrained. If so, the behavior of investments would depend solely on ability D_i^q and obey the following condition in our model,

$$E\left[h\left(D_i^q, N_i\right)|D_i^q\right] = h^U\left(a_q\right) - h_0\left(D_i^q\right), \qquad (19)$$

where $h^U\left(a_q\right)$ indicates the unrestricted investments for the estimated quartile ability levels a_q (see below) as computed using the expression (13) with the values calibrated for r, S and P and the value (here) estimated for g_0. To compute $h^U\left(a_q\right)$ we also need the value for the subsidy rate s, which we calibrate as follows. We first compute marginal subsidy rates for each year of college (1-8 years) by AFQT quartile. The marginal subsidy is computed as 0.77 times the ratio of direct-to-total expenditures, where 0.77 reflects the ratio of current-fund revenue that does not come from tuition and fees average over academic years 1980-81 to 1989-90 as reported by the Digest of Education Statistics, 2003 (Table 333). We use the average of these rates using the distribution of completed schooling in our NLSY79 sample. The resulting government subsidy matching rate of $s = 0.516$ is our baseline value.

We use GMM based on the moment conditions (18) and (19) to estimate the parameters $\{\aleph_q, \alpha, g_0, h_0\left(D_i^q\right)\}$.[43] With those estimates, our ability levels for each AFQT are $a_q = \exp\left(\aleph_q\right) E\left(e^{\nu_{i\tau}}|Z_i\right)$, $\nu_{i,t-P}$ being the residuals in (18). The estimates for α, g_0, and $h_0\left(D_i^q\right)$ are directly used as the parameter values in the model. In our baseline specification we impose that the minimum human capital level is the same for all ability levels, i.e. $h_0\left(D_i^q\right) = h_0$, but to check for robustness we also report the case in which $h_0\left(D_i^q\right)$ varies by ability. As a final robustness check, we also report the behavior of the model when earnings profiles, i.e. $g_0\left(D_i^q\right)$, also varies by ability.

Obviously, despite having set the parameter values $\{\aleph_q, \alpha, g_0, h_0\left(D_i^q\right)\}$ so as to match unrestricted investment, this does not imply that simulations of our baseline $G+L$ model would necessarily lead to unrestricted investment. We only have set the parameters of the earnings function and have not used any data regarding the individual's (or his family) wealth or income. Depending on the constraints κ_1 and κ_2, the model's implied investments can be quite different from the unrestricted model. Thus, the metric to test our calibrated model is whether it reproduces the

[43] The moments associated with the first equation could suffice to estimate the parameters of the model. However, those estimates led us to highly unstable behavior of investments. With the set of moments associated to the investments we obtained much more precise estimates and estable investments without sacrificing much in terms of fit with respect to earnings. The mean squared error (MSE) on earnings using the earnings equation alone is 0.593, and only goes to 0.604 when the second set of moments are also used in the estimation.

behavior of investments of NLSY79 participants for an empirically plausible distribution of wealth w available to youth at the time of their college investment decisions.

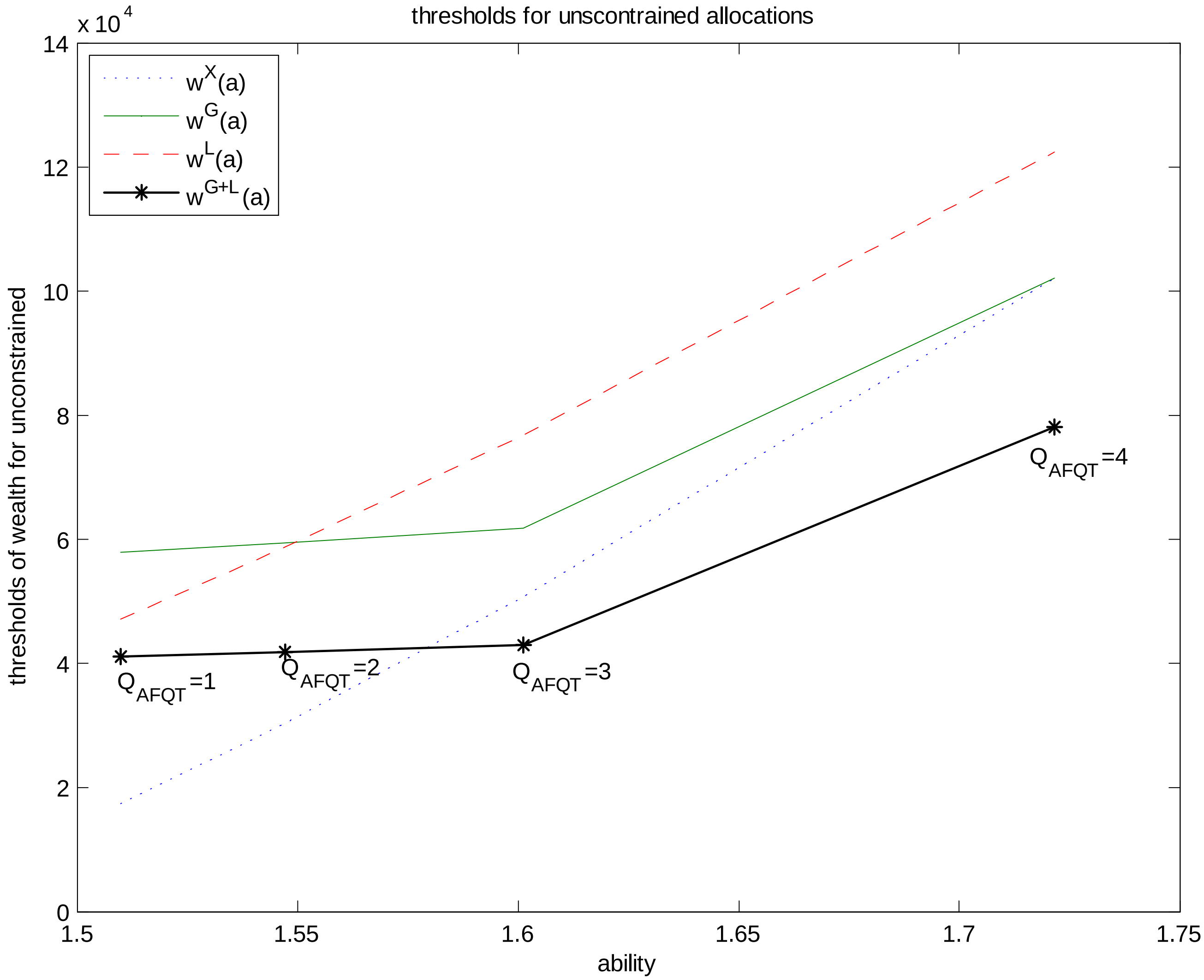


Figure 1: Thresholds of wealth for unconstrained allocations, several models.

5.5 Baseline Simulations

To be written

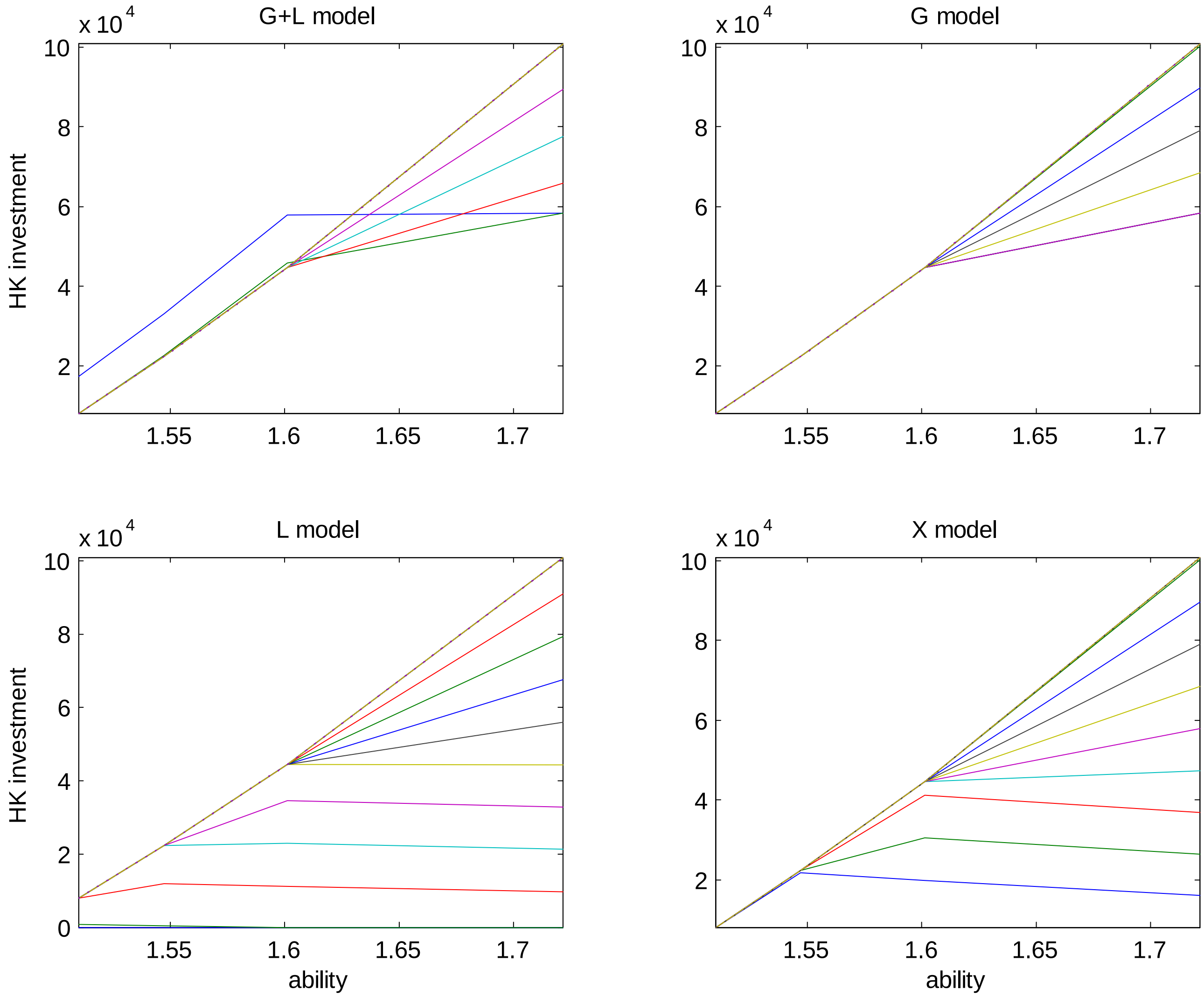


Figure 2: Relationship between Ability and Investment in the various models

6

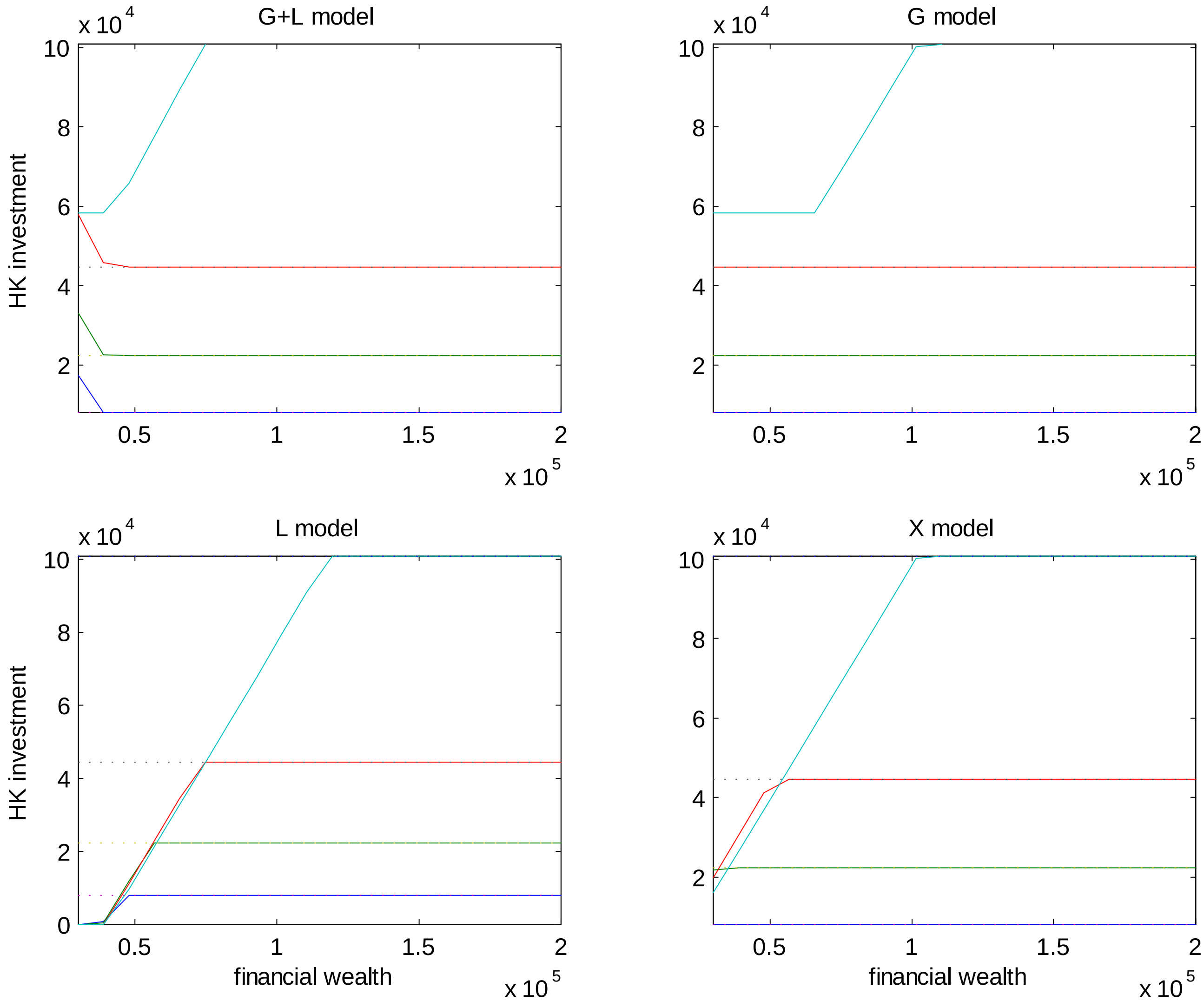


Figure 3: Relationship between initial Wealth and Investment, in the various models

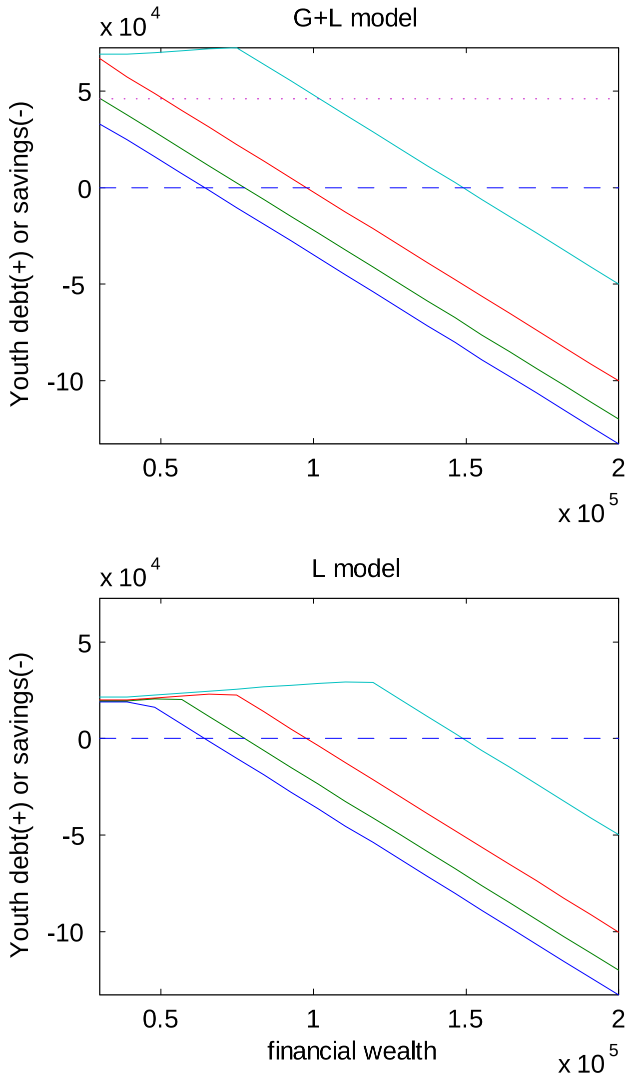
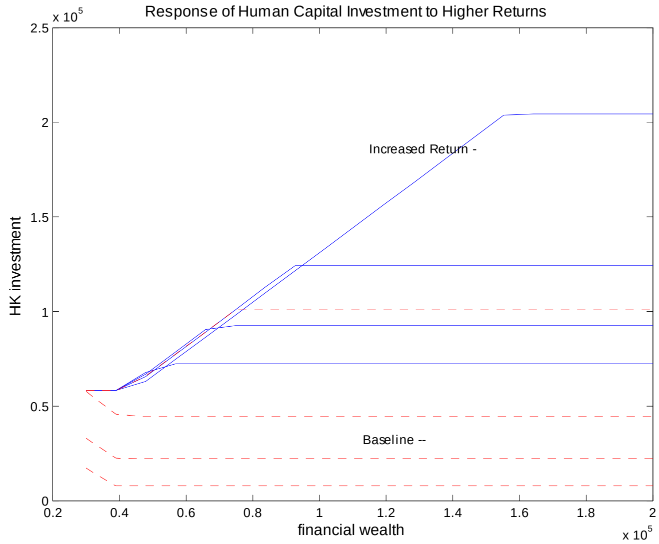
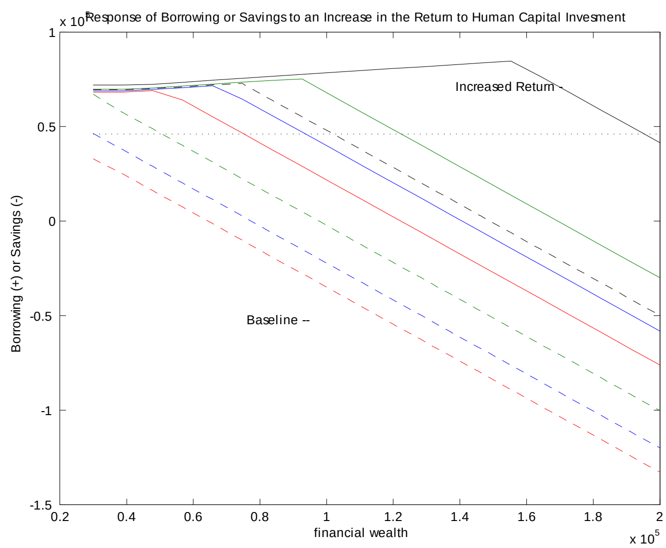


Figure 4: Relationship between borrowing, ability and initial assets: various models.





Counterfactual Exercises

6.1 Counterfactual 1: Increase in the Return to Investment

6.2 Counterfactual 2: Policy Experiments

To be added.

7 Conclusions

This paper develops a life-cycle human capital investment model that incorporates the borrowing opportunities of GSL programs and private lenders who face limited commitment by borrowers. Both types of lenders directly link credit to investment behavior, with private lenders further linking credit to observable borrower characteristics that determine investment productivity. These links are absent in previous models and we show that they play a central role in determining human capital investment behavior.

We draw three broad lessons. First, our model with endogenous borrowing constraints is consistent with empirical studies in that it implies positive effects of both family income and ability on schooling attainment among constrained borrowers. In contrast, under empirically plausible assumptions, a standard exogenous constraint model predicts a negative ability – schooling relationship for constrained borrowers. Second, the direct link between credit and investment inherent in GSL programs breaks the tradeoff between income maximization and consumption smoothing for some constrained borrowers. As a result, students constrained by GSL limits from borrowing more then they invest will choose to invest the unconstrained optimal amount. Previous empirical tests based only on educational attainment (or the marginal returns on investment) cannot detect this constraint. Third, our model is able to reproduce the increased effect of family income on college attendance, the increased fraction of students borrowing the maximum amount from GSL programs, and the increased student borrowing from private lenders over the last few decades as an equilibrium response to rising college costs and returns.

It is important to consider both GSL programs and private lenders when modelling human capital investment decisions. The features of GSL programs allow for the possibility that some student borrowers invest the optimal unconstrained amount even if they are constrained. For them, the existence of a private loan market allows for better smoothing of consumption over time. The presence of private lending generates a positive relationship between ability and investment for individuals from all income backgrounds – a robust empirical pattern. The co-existence of private and public sources of credit yields some important interactions. First and foremost, investment is higher when both sources are available than when only one or the other exists. Our quantitative analysis suggests that many more persons would be constrained in the absence of either GSL programs or a private student loan market.

We use an analytically tractable model to show that the main features of GSL programs and private lending under limited commitment are important for explaining observed investment patterns. An obvious next step is to introduce uncertainty about the returns to investment. While we do not expect such an extension to alter our predictions about the relationship between investment, ability and family resources in any important way, incorporating uncertainty opens new and interesting areas of inquiry. With uncertainty about labor market success, the option of default provides insurance against adverse outcomes. Private lenders and governments must strike a balance between providing this insurance to borrowers and enforcing repayment. This defines an interesting optimal lending and enforcement policy, which may be complicated by the fact that students possess private information about their own abilities or willingness to study. Additionally, the existence of labor market uncertainty generally implies default by some agents in equilibrium. This makes it possible to study which agents are most likely to default and how economic changes and public policies affect default behavior. We view our framework as a natural starting point for these types of analysis.

We also suggest that future empirical efforts to estimate school-choice models consider the types of endogenous constraints and punishments we emphasize here. With reliable data on schooling, borrowing, earnings, and loan repayment (an admittedly tall order), structural estimation may be able to identify more general punishment strategies than we have assumed in this paper. Such an analysis would provide important new insights about the role of borrowing constraints, who is likely to be constrained, and how higher education policies and economic changes affect 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.[44] 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.[45] 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.[46] 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 wealth (in NLSY97), and AFQT score quartiles.[47]

Our multivariate analysis controls for a host of family background variables. For both cohorts, we control for maternal education by categorizing mothers as high school dropouts, those who completed high school or more, and those who completed at least one year of college. We account for family structure in the NLSY79 by controlling for the number of siblings the youth reported in 1979. For the NLSY97, we control for the number of household members under the age of 18 as of the 1997 survey date. Additional family structure information is provided by an indicator variable for whether both parents are present in the home at age 14 in the NLSY79 and in 1997 (i.e. ages 13-17) in the NLSY97. Family residence in an urban (metropolitan) area at age 14 (age 12) is accounted for with the 1979 (1997) cohort. We control for the mother's age at birth as well as gender and race (blacks, Hispanics and whites for the NLSY79; blacks, Hispanics, other non-whites, and whites for the NLSY97 data). Finally, we allow for differences by year of birth.

B Proofs and Other Aspects of the Two-Period Model

B.1 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 $w_{\min}^X(a)$ is defined by $d^U\left(a, w_{\min}^X(a)\right) = d_0$, and therefore it is increasing in a. Consumption smoothing implies that $w_{\min}^X(a) \geq h^U(a) - d_0$ (the minimum wealth needed to finance $h^U(a)$ given maximum borrowing) and that $w_{\min}^X(a)$ is steeper than $h^U(a)$ as a function of a. To see this, implicit differentiation leads to $\frac{dw_{\min}^X(a)}{da} = \frac{\partial d^U\left(a, w_{\min}^X\right)}{\partial a} / \frac{\partial d^U\left(a, w_{\min}^X\right)}{\partial w} > \frac{\partial d^U\left(a, w_{\min}^X\right)}{\partial a} > \frac{dh^U(a)}{da} > 0$

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

[45] 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).

[46] 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.

[47] 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.

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

B.2 The Possibility Over-investment in the G+L model

We now show that the interaction between private lending and GSL programs can lead to over-investment in human capital, a situation where the marginal return on human capital investments is below is financial cost (i.e. $af'\left[h^{G+L}(w,a)\right] < R$) More interestingly, we now show that that inequality can be strict. For simplicity, assume that $a < \check{a} \equiv Rd_{\max}/f(d_{\max})$, so the constraint $d_g \leq d_{\max}$never binds. Assume also that $w = 0$, so that both constraints $d_g \leq h$and $d_g \leq h$bind. Under these conditions private lending will be financing consumption and GSL credit will be used exclusively for investment. The person solves

$$\max_h u\left[\kappa af(h)\right] + \beta u\left[af(h) - R\kappa af(h) - Rd\right],$$

where $R\kappa af(h)$ and Rh are the repayments to private and GSL debt respectively. The FOC is

$$\kappa af'(h)\, u'\left[\kappa af(h)\right] = \beta\left[R - af'(h)(1 - R\kappa)\right] u'\left[af(h)(1 - R\kappa) - Rh\right].$$

The left-hand side is positive and strictly decreasing in h; the right-hand side might be initially negative and non-monotone but it is eventually positive and increasing. For an optimum, the right-hand-side must cross the left-hand side from below. So, we would obtain over-investment if at $h = h^U(a)$, the left-hand-side is above the right-hand-side. Since $af'\left[h^U(a)\right] = R$, the condition for over-investment is

$$u'\left[\kappa af\left(h^U(a)\right)\right] > \beta Ru'\left[af\left(h^U(a)\right)(1 - R\kappa) - Rh^U(a)\right]. \tag{20}$$

which can hold under many parameter configurations. By continuity, over-investment can hold for positive values of $w$in a neighborhood around 0.

As an example, assume $\beta R = 1$. With condition (20) is simply that early consumption $\kappa af\left(h^U(a)\right)$ be strictly lower than late consumption $(1 - R\kappa)\, af\left[h^U(a)\right] - Rh^U(a)$. Upon rearranging, this is $Rh^U(a) < af\left(h^U(a)\right)\left[1 - \kappa(1 + R)\right]$.Finally, when $f(h) = h^\alpha$, then $h^U(a) = [\alpha a/R]^{\frac{1}{1-\alpha}}$and we would have over-invesment for low (w, a) individuals if $\alpha < 1 - \kappa(1 + R)$.

B.3 Proofs

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

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

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,

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

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

Proof of Proposition 2. From the FOC define

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

The second order condition imlies $\partial F/\partial h < 0$, which, combinaed with implicit differentiation leads to $sign\left\{\frac{\partial h}{\partial w}\right\} = sign\left\{\frac{\partial F}{\partial w}\right\}$ 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,

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

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

Proof of Proposition 3. Using the FOC of the exogenous constraint model,

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

which in principle could be $+\infty$. If $u(c) = c^{1-\sigma}/(1-\sigma)$, then a finite $\hat{a}(w)$ would be given by

$$\hat{a} : w\left(\beta f'[d_{\max}]\right)^{\frac{1}{\sigma}} = (\hat{a})^{\frac{\sigma-1}{\sigma}} f(d_{\max}) - Rd_{\max}(\hat{a})^{\frac{-1}{\sigma}}.$$

If $\sigma > 1$ (IES < 1), the RHS is strictly increasing and unbounded and, hence, $\hat{a}(w)$ is finite. The rest is direct upon examination of optimality conditions under the three different cases. ■

Proof of Lemma 4. Straightforward and omitted. ■

Proof Proposition 5. The fact that $h^{G+L}(w,a) \geq h^U(a)$ for a<ā is trivial, while fact that for low enough (a,w); that the inequality is strict follows from the previous subsection of this appendix. For the second part of the proposition: Items (i) and (ii) follow from the discussion of the GSL and implicit differentiation of the following first order condition for investment as in Proposition 2. We now prove (iii). From the FOC for investment when the constraint binds, define

$$F \equiv (\kappa af'(h) - 1)\, u'(w + \kappa af(h) - h) + \beta af'(h)(1-\kappa R)\, u'[af(h)(1-\kappa R) - Rd_{\max}] = 0.$$

From the second order condition, $\partial F/\partial h < 0$ and sign$\{\partial h/\partial a\}$=sign$\{\partial F/\partial a\}$. After some simplification:

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

where $c_1 = af(h)(1-\kappa R) - Rd_{\max}$ and $c_0 = w + \kappa af(h) - h$. In the expression, the first two terms are always positive, while the third term can be either positive or negative. Multiply and divide $\partial F/\partial a$ by $u'(c_1)$, to obtain

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

where the second uses $c_1 = (1-\kappa R)\, af(h)$ and the definition of $\eta(\cdot)$, the IES. The last line follows from $u'(c_0)/u'(c_1) \geq \beta R$ and then simplifying. Since $\kappa \frac{c_1}{c_0}\frac{1}{\eta(c_0)}$ is non-negative, the condition $\eta(c_1) > 1-\kappa R$ implies that $\partial F/\partial a > 0$. Finally, if $\beta R \geq 1$ then $c_1 \geq c_0$, and with $\eta(\cdot)$is non-decreasing, then $\eta(c_1) \geq \eta(c_0)$. Therefore,

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

which is strictly positive if $\kappa \geq [1-\eta(c_0)]/(1+R)$. ■

Proof of Proposition 6. Part (i) is from direct inspection based on the thresholds defined at the beginning of this appendix. From the first order condition for investment (imposing the borrowing constraint), define $F(h, d_{\max}, \kappa)$ as

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

The FOC for $h^{G+L}(a, w; d_{\max})$ is $F=0$ and the SOC $\partial F/\partial h < 0$. Then, $sign\left\{\partial h^{G+L}(a, w; d_{\max})/\partial d_{\max}\right\} = sign\left\{\partial F/\partial d_{\max}\right\}$, and

$$\frac{\partial F}{\partial d_{\max}} = [1-\kappa af'(h)]\left[-u''(c_0)\right] + \beta af'[h](1-\kappa R)\, R\left[-u''(c_1)\right] > 0.$$

Similarly, $sign\left\{\partial h^{G+L}(a, w; \kappa)/\partial \kappa\right\} = sign\left\{\partial F/\partial \kappa\right\}$, where

$$\frac{\partial F}{\partial \kappa} = af'(h)[u'(c_0) - \beta Ru'(c_1)] + af(h)\left[(\kappa af'(h)-1)u''(c_0) + \beta R(\kappa R - 1)u''(c_1)af'(h)\right] > 0,$$

since $u'(c_0) > \beta Ru'(c_1)$ and $af'(h) < R$ for constrained agents. ■

C Proofs and Other Aspects of the Quantitative Model

C.1 Thresholds

For the sake of brevity, let

$$m^U(a) \equiv \Phi_{[P,R]}a\left[h^U(a)\right]^{\alpha}\frac{\Theta_{[S,P]}}{\Theta_{[S,T]}} + e^{r(P-S)}\frac{\Theta_{[P,T]}}{\Theta_{[S,T]}}h_I^U(a).$$

With this definition, $d^U(a,w) = m^U(a) - e^{r(P-S)}\Theta_{[P,T]}/\Theta_{[S,T]}w$, and:

$$w^X(a) \equiv e^{-r(P-S)}\frac{\Theta_{[S,T]}}{\Theta_{[P,T]}}\left[m^U(a) - d_{\max}\right],$$

$$w^L(a) \equiv e^{-r(P-S)}\frac{\Theta_{[S,T]}}{\Theta_{[P,T]}}\left[m^U(a) - \kappa_L \Phi_{[P,R]}a\left[h^U(a)\right]^{\alpha}\right],$$

$$w^G(a) \equiv e^{-r(P-S)}\frac{\Theta_{[S,T]}}{\Theta_{[P,T]}}\left[m^U(a) - \min\left\{e^{r(P-S)}h^U(a), d_{\max}\right\}\right].$$

For our baseline model:

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

C.2 Formulas for Credit Constraints in the Model

A non-defaulting individual retains access to formal credit markets, is able to optimally smooth consumption and attains an post-graduation lifetime utility given by

$$V_P^R(a, h, d_g, d_p) = \Theta_{[P,T]}\frac{\left\{\left[\Phi_{[P,R]}ah^{\alpha} - d_g - d_p\right]/\Theta_{[P,T]}\right\}^{1-\sigma}}{1-\sigma}. \tag{21}$$

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

$$V_P^D(a,h,d_g,r(\cdot;d_g))=\int_P^{P+\pi} e^{-\rho(t-P)}\frac{\left[(1-\gamma)\,ah^{\alpha}E\,(t-P)-r\left(t;d_g\right)\right]^{1-\sigma}}{1-\sigma}dt$$
$$+e^{-\rho\pi}\Theta_{[P+\pi,T]}\frac{\left\{\left[\Phi_{[P+\pi,R]}ah^{\alpha}-e^{r\pi}\left(d_g-R\left(P+\pi,d_g\right)\right)\right]/\Theta_{[P+\pi,T]}\right\}^{1-\sigma}}{1-\sigma}, \quad (22)$$

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 or longer that the length $\pi$of default punishment, the repayments to GSL loans is given by $r\left(t,d_g\right)=\delta ah^{\alpha}E\,(t-P)$, i.e. the individual must pay a constant fraction of his earnings.[48] 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. In one extreme is the "fastest" repayment case when $\delta=\delta_{fast}=d_g/\left(\Phi_{[P,P+\pi]}ah^{\alpha}\right)$and all the GSL debt must be repaid while the agent is being punished. This is the most disruptive case and is only relevant when the earnings are high enough to cover the debt and leave positive consumption during the punishment (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 $\Delta\equiv\int_P^{P+\pi}e^{-\rho(t-P)}E\,(t-P)^{1-\sigma}\,dt$. In 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 being rolled-over in the mean time, leading to a balance of $e^{r\pi}d_g$at time $P+p$. This case is relevant only if $\Phi_{[P+\pi,R]}ah^{\alpha}>e^{r\pi}d_g$, leads to utility equal to

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

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

In general, for intermediate values of δwe can use the resulting (22) expression with the condition $V^R\geq V^D$ to obtain a closed form for the constraint on private credit:

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

with $M_0\equiv\Delta\,[1-\gamma-\delta]^{1-\sigma}/\left(\Theta_{[P,T]}\right)^{\sigma}$, $M_1\equiv e^{-\rho\pi}\left(\frac{\Theta_{[P+\pi,T]}}{\Theta_{[P,T]}}\right)^{\sigma}$, 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-schooling 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 limits $d_{\max}$could lead to an overall expansion in available credit.

[48] Given our assumptions, the timing and structure of repayments does not matter if the agent does not default. With access to perfect credit markets he can attain any timing of consumption as long as its present value equates the net present value of earnings minus debt.

For our baseline case we set $\delta = \delta^*$, where

$$\frac{\delta^*}{1-\gamma}=\frac{e^{r\pi}\left(d_g-\delta\Phi_{[P,P+\pi]}ah^{\alpha}\right)}{\Phi_{[P+\pi,R]}ah^{\alpha}},$$

which simply equates the service of the debt, as a fraction to the the income (net of garnishments) between the punishment and post-punishment periods. Imposing this equality we can write

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

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

$$V_P^D(a,h,d_g,\delta^*)=\Theta_D\frac{\left[ah^{\alpha}-\frac{d_g}{\left[\Phi_{[P,R]}-\gamma\Phi_{[P,P+\pi]}\right]}\right]^{1-\sigma}}{1-\sigma}. \tag{23}$$

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

$$\kappa_1\equiv 1-\frac{\Theta_{[P,T]}}{\Phi_{[P,R]}}\left(\frac{\Theta_D}{\Theta_{[P,T]}}\right)^{\frac{1}{1-\sigma}} \text{ and } \kappa_2=\left(\frac{\Theta_D}{\Theta_{[P,T]}}\right)^{\frac{1}{1-\sigma}}\frac{\Theta_{[P,T]}}{\Phi_{[P,R]}-\gamma\Phi_{[P,P+\pi]}}-1. \tag{24}$$

Direct inspection of this formulas verifies that $0<\kappa_1<1$, $\kappa_2>-1$, as well as the other properties stated in the text.

C.3 Proofs

Proof of Proposition ??. TWe only prove part (iii) of the second part, since everything else is virtually the same as for the proposition in the two-period model. To keep the algebraic expressions manageable, we simply use κ to denote κ_1 and define

$$\begin{aligned}&A\equiv a\Phi_{[P,R]}, \qquad x=h_I, \qquad g\equiv\alpha Ah^{\alpha-1}\left(1+s\right),\\&c_0\equiv\left(w+e^{-r(P-S)}\kappa Ah^{\alpha}-x\right)/\Theta_{[S,P]}, \qquad c_1\equiv\left(1-\kappa\right)\left[Ah^{\alpha}/\Theta_{[P,T]}\right].\end{aligned}$$

Then, when the constraint binds, the maximization of the individual is solely in terms of x (equivalently in terms of h):

$$\max_{\{x:x\geq 0\}} L=\left\{\Theta_{[S,P]}\frac{[c_0]^{1-\sigma}}{1-\sigma}+e^{-\rho(P-S)}\Theta_{[P,T]}\frac{[c_1]^{1-\sigma}}{1-\sigma}\right\},$$

with a first order condition equal to

$$F\equiv(c_0)^{-\sigma}\left(e^{-r(P-S)}\kappa g-1\right)+(c_1)^{-\sigma}\left(1-\kappa\right)ge^{-\rho(P-S)}. \tag{25}$$

By the optimality of h, $\partial F/\partial h<0$, and by the implicit function theorem $sign\left\{\frac{\partial h}{\partial a}\right\}=sign\left\{\frac{\partial F}{\partial a}\right\}$. Direct derivation on F implies that

$$\begin{aligned}\frac{dF}{da}=&(c_0)^{-\sigma}\left(e^{-r(P-S)}\kappa\frac{\partial g}{\partial a}\right)-\sigma(c_0)^{-\sigma}\left(e^{-r(P-S)}\kappa g-1\right)\frac{1}{c_0}\frac{\partial c_0}{\partial a}\\&+(c_1)^{-\sigma}\left(1-\kappa\right)\frac{\partial g}{\partial a}e^{-\rho(P-S)}-\sigma\left[c_1\right]^{-\sigma-1}\left(1-\kappa\right)ge^{-\rho(P-S)}\frac{\partial c_1}{\partial a}.\end{aligned}$$

The first order condition (25) implies that $-(c_0)^{-\sigma}\left(e^{-r(P-S)}\kappa g-1\right)=(c_1)^{-\sigma}(1-\kappa)\,ge^{-\rho(P-S)}$, which used in the previous equation leads to

$$\frac{dF}{da}=(c_0)^{-\sigma}e^{-r(P-S)}\kappa\frac{\partial g}{\partial a}+\sigma(c_1)^{-\sigma}(1-\kappa)\,ge^{-\rho(P-S)}\left(\frac{1}{c_0}\frac{\partial c_0}{\partial a}\right)$$
$$+(c_1)^{-\sigma}(1-\kappa)\,ge^{-\rho(P-S)}\left(\frac{\partial g}{\partial a}\frac{1}{g}\right)-\sigma(c_1)^{-\sigma}(1-\kappa)\,ge^{-\rho(P-S)}\left(\frac{\partial c_1}{\partial a}\frac{a}{c_1}\right).$$

Now, taking $T\equiv(c_1)^{-\sigma}(1-\kappa)\,ge^{-\rho(P-S)}/a>0$ as a common factor and rearranging, we obtain

$$\frac{dF}{da}=T\left[\left(\frac{c_1}{c_0}\right)^{\sigma}e^{(\rho-r)(P-S)}\left(\frac{\kappa}{1-\kappa}\right)\left(\frac{\partial g}{\partial a}\frac{a}{g}\right)+\sigma\left(\frac{c_1}{c_0}\frac{\partial c_0}{\partial a}\frac{a}{c_1}\right)+\left(\frac{\partial g}{\partial a}\frac{a}{g}\right)-\sigma\left(\frac{\partial c_1}{\partial a}\frac{a}{c_1}\right)\right],$$

where the first term within the brackets resulted from multiplying and dividing it by c_1.Notice that

$$\frac{\partial g}{\partial a}\frac{a}{g}=\frac{\partial c_1}{\partial a}\frac{a}{c_1}=1,\text{ and }\frac{\partial c_0}{\partial a}\frac{a}{c_1}=\frac{e^{-r(P-S)}\kappa\Theta_{[P,T]}}{(1-\kappa)}.$$

Using these three equalities with the inequality $c_1/c_0\geq e^{\left[\frac{r-\rho}{\sigma}\right](P-S)}$ that holds when the borrowing constraint binds, then, after some rearranging we obtain:

$$\frac{dF}{da}\geq T\left\{\left(\frac{\kappa}{1-\kappa}\right)\left(1+\sigma e^{\left[\frac{r-\rho}{\sigma}-r\right](P-S)}\Theta_{[P,T]}\right)+1-\sigma\right\}.$$

From direct inspection, the term inside brackets is positive if κ is above $\underline{\kappa}$ as defined in the text, and the argument is complete. ■

D Estimating the Human Capital Production Function

To be added.

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. 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.

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

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.

Dick, A., A. Edlin, and E. Emch (2003), “The Savings Impact of College Financial Aid", *Contributions to Economic Analysis and Policy*, 2(1).

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.

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).

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.

P. Kehoe 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.

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

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

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

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. XXXX

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.