The Nature of Credit Constraints and Human Capital*

Lance J. Lochner  
University of Western Ontario and NBER

Alexander Monge-Naranjo  
Northwestern University

February 2, 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, University 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. 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 costs and returns to college in the U.S. over the 1980s and 1990s.\footnote{See College Board (2005) for evidence on rising costs and Katz and Autor (1999) Heckman, Lochner and Todd (2008) for evidence on rising returns.}
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. Both approaches neglect the link between borrowing opportunities and investment decisions that plays a key role in both 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 trade-off 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. By introducing the restriction that borrowing cannot exceed investment, GSL programs effectively increase the population of students who invest the unconstrained optimal amount.

---


3Thus, 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.
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. 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.

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

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.

5 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.
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 available in the U.S. to finance college education. We begin with a description of key institutional features of 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 government student loan (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.\(^6\)

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 these loans to some students under the same rules and terms.\(^7\) 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

\(^6\)See The College Board (2006) for details about financial aid disbursements and their trends over time.

\(^7\)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.
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. 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. 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. Student borrowers can defer loan repayments 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. The relative stability of real GSL limits combined with a near doubling of tuition costs in recent decades (College Board 2005), has pushed recent student borrowing up against these upper limits for many undergraduates. Indeed, the fraction of all undergraduate borrowers that borrowed the maximum limit from the federal Stafford Student Loan Program 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,
Table 1: Borrowing Limits for Stafford and Perkins Student Loan Programs (1993-2007)

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

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 unsubsidized)loan amount is not greater than the independent student limits.
*** Cumulative graduate loan limits include loans from undergraduate loans.
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.\textsuperscript{12} 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.\textsuperscript{13}

\subsection*{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 postgraduation 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.\textsuperscript{14}

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.\textsuperscript{15} 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, they vary substantially in their terms and eligibility requirements.\textsuperscript{16} 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.\textsuperscript{17}

\textsuperscript{12}Formally, a borrower is considered to be in default once a payment is 270 days late.
\textsuperscript{13}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.
\textsuperscript{14}These figures do not include student borrowing on credit cards, which has also increased considerably over this period. See College Board (2005).
\textsuperscript{15}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.
\textsuperscript{16}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.
\textsuperscript{17}For example, the relatively new private lending institution, MyRichUncle, states on its website
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.\textsuperscript{18} 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

\textsuperscript{18}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.
Test, AFQT) for youth from all levels of family income and wealth.\textsuperscript{19}

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.\textsuperscript{20}

More recently, Belley and Lochner (2007) show that family income has become a much more important determinant of college attendance in the early 2000s.\textsuperscript{21} 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 2 shows college attendance rates by AFQT quartiles and either family income or family wealth quartiles in the NLSY79 and NLSY97.\textsuperscript{22} 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

\textsuperscript{19}AFQT scores are widely used as measures of cognitive achievement by social scientists and are strongly correlated with post-school earnings conditional on educational attainment. See, e.g., Cawley, et al. 2000. Appendix A provides additional details.

\textsuperscript{20}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).

\textsuperscript{21}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.

\textsuperscript{22}See Appendix A for a detailed description of the data and variables used here.
(under empirically relevant assumptions about preferences). Therefore, a close examination of the ability – college attendance relationship by family income and wealth is warranted. Figure 2 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.\textsuperscript{23}

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.\textsuperscript{24} Results are reported for separate regressions by family income or wealth quartile. The estimates confirm the general patterns observed in Figure 2: ability has strong positive effects on college attendance for all family income and wealth quartiles in both NLSY samples.

### 4 Modelling Student Credit

We now 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 a framework which incorporates realistic assumptions about public and private student lending sources produces investment – wealth and investment – ability relationships that are qualitatively consistent with the empirical patterns discussed above, while the canonical exogenous borrowing constraints model does not.

#### 4.1 Preferences and Human Capital Production Technology

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

\[ U = u(c_0) + \beta u(c_1), \]  

\textsuperscript{23}We observe similar patterns in the NLSY97 for age 20 enrollment in four-year colleges/universities conditional on attendance at any post-secondary institution. Among youth from the lowest wealth quartile, the enrollment rate in four-year schools (conditional on post-secondary enrollment) is 34% higher for the most able relative to the least able. Among the highest wealth quartile, the difference is 32%. For the lowest family income quartile, the same high - low ability difference is 41%, while it is 52% for the highest income quartile.

\textsuperscript{24}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.
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 initial financial assets \( w \geq 0 \) and ability \( a > 0 \). Initial assets represent all transfers from parents and other family members. 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.\(^{25}\)

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 are complementary.\(^{26}\) 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 determined by first order conditions.

Human capital investment, \( h \), is in units of the consumption good.\(^{27}\) Individuals can borrow \( d \) of these units (or save, which is indicated by \( d < 0 \)) at a gross interest rate \( R > 1 \). Given \( a, h \) and \( d \), consumption in each of the periods is

\[
\begin{align*}
    c_0 &= w + d - h, \quad (2) \\
    c_1 &= af(h) - Rd. \quad (3)
\end{align*}
\]

\(^{25}\)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, 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.

\(^{26}\)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.

\(^{27}\)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.
4.2 Unrestricted Allocations

In the absence of financial frictions, young individuals maximize utility (1) subject to (2) and (3). This maximization can be separated into two steps. The first is to choose \( h \) to maximize the present value of lifetime net resources, \( w + R^{-1}af(h) - h \). Optimal unrestricted investment, \( h^U(a) \), equates the marginal return on human capital with the return on financial assets:

\[
af' [h^U(a)] = R.
\]  

Therefore, \( h^U(a) \) is 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:

\[
u' (w + d^U(a, w) - h^U(a)) = \beta R u' \left( a f[h^U(a)] - R d^U(a, w) \right).
\]  

Therefore, unconstrained borrowing is strictly decreasing in wealth and increasing in ability. Optimal debt is strictly increasing in ability for two reasons. First, more able individuals wish to finance a larger investment. Second, for any given level of investment, more able individuals earn higher net lifetime income and, therefore, wish to consume more during youth. 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, \( \frac{\partial d^U(a, w)}{\partial a} > \frac{\partial h^U(a)}{\partial a} > 0 \) and \( -1 < \frac{\partial d^U(a, w)}{\partial w} < 0 \).

Proofs for all results and other analytical details for the models in this section are provided in Appendix B.

We make repeated use of Lemma 1 to characterize the behavior of investment with borrowing constraints. In all models of constraints discussed below, unconstrained individuals invest \( h^U(a) \) and borrow \( d^U(a, w) \) to smooth consumption optimally. We, therefore, limit our discussion to a characterization of which individuals will be constrained and their investment and borrowing behavior.

4.3 Exogenous Borrowing Constraints

At least since Becker (1975), economists have introduced financial market imperfections in models of human capital formation. Credit constraints are typically introduced by imposing a fixed and exogenous upper bound on the amount of debt.\(^2\) Following this approach,
assume that borrowing is restricted by the exogenous constraint:

\[ d \leq d_0, \quad \text{(EXC)} \]

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 \( w_{\min}^X (a) \) defines who is constrained \((w < w_{\min}^X (a))\) and who is unconstrained \((w \geq w_{\min}^X (a))\). Because this threshold is increasing in ability, persons who are constrained will be of high ability or low wealth. (Appendix B characterizes the threshold \( w_{\min}^X (a) \) and the thresholds defined by other constraints considered in this section.)

Individuals constrained by (EXC) exhaust their ability to bring resources to the early (investment) period. 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 equating its marginal costs and benefits:

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

We highlight four empirically testable results. First, constrained investment never exceeds unconstrained investment. This result holds for all forms of constraints considered in this paper. 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 1** If \( w < w_{\min}^X (a) \) 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, \( af' [h^X (a, w)] \), 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 (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.

The more interesting and novel result is part (iv): the model predicts a negative relationship between ability and investment for an IES below one.\(^{29}\) The relationship between

\(^{29}\) 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.
ability and investment for constrained youth derives from two opposing effects. On the one hand, an increase in ability raises the financial returns to investment, which encourages investment. On the other hand, ability raises lifetime income, which encourages early consumption. Since constrained youth can only increase early consumption by investing less, strong preferences for smoothing (e.g. IES ≤ 1) imply that the second effect dominates. This is a serious shortcoming of the model, because most estimates of the IES are less than one (see Browning, Hansen, Heckman 1999) and, as we showed earlier, a positive relationship between ability and investment is a robust empirical regularity.

4.4 Government Student Loan Programs

Now, consider that student credit is governed by the key features of GSL programs described earlier. First, lending is tied to investment and cannot be used to finance non-schooling related consumption goods or activities:

\[ d \leq h. \]  

(TIC)

In the absence of other sources of credit, (TIC) is equivalent to \( c_0 \geq w \). Second, borrowing is constrained by an upper limit \( 0 < d_{\text{max}} < \infty \) for the total credit to each student:

\[ d \leq d_{\text{max}}. \]  

(6)

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

\[ d \leq \min \{ h, d_{\text{max}} \}. \]  

(GSLC)

Finally, we assume that all government loans must be repaid in full. We use the superscript \( G \) to refer to allocations under a GSL program.

To understand the role of (TIC), first assume that it is the only constraint individuals face.\(^\text{30}\) In this case, individuals are unconstrained as long as desired investment exceeds desired borrowing. Because desired investment is increasing in ability, the (TIC) constraint is less stringent than (EXC) for higher ability individuals while the reverse is true for those of lower ability.

If (TIC) is the only binding constraint, then \( d = h \) and individuals consume their entire initial wealth (and no more) while in school. As such, the maximization problem becomes

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

which is equivalent to maximizing lifetime income and yields optimal investment equal to the unconstrained amount \( h^U(a) \).

\(^{30}\)This would be the case if upper borrowing limits were non-existent or set very high (e.g. PLUS program for students’ parents).
Tying credit dollar-for-dollar to investment avoids any conflict between smoothing consumption and maximizing net lifetime resources. If (TIC) were the only constraint on borrowing, everyone would invest the unconstrained amount, $h^U(a)$, regardless of initial wealth. Only consumption decisions would be distorted by the constraint. Empirical tests based on investment differences by family resources would always conclude that borrowing constraints were non-binding, even when consumption allocations were distorted. *Empirical tests must use measures of consumption over time to detect this constraint.*

Consider now the full GSL constraint (GSLC). For comparison purposes and ease of exposition, assume that $d_{\text{max}} = d_0$, so (EXC) coincides with (6). It is straightforward to determine the wealth threshold $w^G_{\text{min}}(a)$, above which individuals are unconstrained. As with exogenous constraints, the wealth threshold $w^G_{\text{min}}(a)$ is increasing in ability, so more able and less wealthy persons tend to be constrained under the GSL. There are three potential categories of constrained individuals ($w < w^G_{\text{min}}(a)$). The first group 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 more for consumption purposes. The other two groups are composed of more able individuals who would like to invest more than $d_{\text{max}}$. The second group is constrained only by the upper limit on borrowing (6). These individuals borrow $d_{\text{max}}$ and invest more than this using some of their initial assets to help finance schooling. Investment coincides with $h^X(a,w)$, because (TIC) is slack. The third group (very poor high ability youth) is constrained by both (6) and (TIC). They borrow and invest the maximum amount, $d_{\text{max}}$, consuming their initial wealth while in school. This discussion is formally summarized in the following proposition.

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

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

Regardless of the IES, there is always a region in which $h^G(\alpha, w)$ is increasing in ability and independent of initial wealth; there may also be a region in which it is constant and equal to $d_{\text{max}}$. If utility has an IES $\leq 1$, 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.

---

31If $h^X(\alpha, w)$ is always increasing in $\alpha$ (e.g. for an IES sufficiently greater than one), then $h^G(\alpha, w)$ is globally increasing in both arguments. The characterization is as follows: $h^G(\alpha, w) = h^U(\alpha)$, for $a \leq \bar{\alpha}$ or $w \geq w^X_{\text{min}}(\alpha)$; $h^G(\alpha, w) = d_{\text{max}}$ for $a > \bar{\alpha}$ and $w < \bar{w}(\alpha)$ and $h^G(\alpha, w) = h^X(\alpha, w)$ otherwise. The flat region where investment equals $d_{\text{max}}$ may not exist.
Figures 3 and 4 illustrate the behavior of $h^G(a, w)$, $h^X(a, w)$, and $h^U(a)$ for the empirically relevant case of $\text{IES} \leq 1$. These figures also display unconstrained borrowing as a function of ability for different levels of wealth. Figure 3 displays investment and borrowing behavior for two low levels of wealth, $\bar{w}$ and $w_L < \bar{w}$, while Figure 4 illustrates investment behavior for a higher level of wealth $w_H > \bar{w}$.\textsuperscript{32}

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 3). 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 $d_{\text{max}}$. 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 4). 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 4). 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

\textsuperscript{32}Note that $\bar{w} \equiv w^G_{\text{min}}(\bar{a})$ reflects the level of wealth below which agents of ability $\bar{a}$ are constrained, where $\bar{a}$ is the ability level at which unconstrained investment equals the upper limit on borrowing (i.e. $h^U(\bar{a}) = d_{\text{max}}$).
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 2**  Impose \( d_0 = d_{\text{max}} \) and let \( \{h^m, c^m_0, c^m_1, U^m\} \), denote the optimal allocations and attained utilities in models \( m = U, X, G \) for arbitrarily fixed \((a, w) \in \mathbb{R}_+^2\). Then:

\[
h^U \geq h^G \geq h^X, \quad c^U_0 \geq c^X_0 \geq c^G_0, \quad c^G_1 \geq c^U_1 \geq c^X_1, \quad U^U \geq U^X \geq U^G,
\]

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

### 4.5 GSL Programs and Private Lenders

Private lending has become an important source of funding for many college students in the U.S. In this section, we introduce private credit markets to our GSL framework for public lending. While we assume that loans from the GSL are fully enforced by the government, we recognize that private lenders face limited repayment incentives due to the inalienability of human capital and the lack of other forms of collateral held by most potential students. The superior repayment enforcement of GSL loans discussed in Section 2 supports this difference in assumptions about the repayment of these loans.

We begin by characterizing limits on private credit that arise endogenously from limited commitment, briefly discussing human capital investment behavior in the presence of private
lending alone. Then, we discuss investment decisions in the presence of both public and private lending.

A rational borrower repays private loans if and only if the cost of repaying is lower than the cost of defaulting. The incentive to default can be foreseen by rational lenders who, in response, limit their supply of credit.\textsuperscript{33} 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. Since earnings are determined by ability and investment, private credit limits and investments will be co-determined in equilibrium.

For the sake of concreteness, in this section we assume that private lenders can punish defaulting borrowers by garnisheeing a fraction $\bar{\kappa} \in (0, 1)$ of their earnings.\textsuperscript{34} (In the next section, we incorporate additional punishments for default in a richer lifecycle model.) Given the punishment for default, 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 $\bar{\kappa}af(h)$. Foreseeing this, borrowing from private lenders is limited to:

$$d_p \leq \kappa af(h),$$

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

Consider, briefly, allocations in the presence of private lenders alone, so individuals only face constraint (7) on their borrowing. As with all previous models, one can define a wealth threshold $w_{\text{min}}^L(a)$ above which individuals are unconstrained and below which they are constrained. In contrast with earlier models; however, more able persons may be unconstrained while the least able are constrained if punishments are severe enough. Furthermore, there may be some levels of ability for which all individuals are unconstrained regardless of their wealth.

Now, consider both public and private lending together. Borrowing from the GSL, denoted by $d_g$ here, must satisfy (GSLC). If, regardless or whether he defaults or not on private loans, a person must pay GSL loans, then the constraint on private borrowing is still given by (7). Therefore, the constraint on total borrowing constraint is given by

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

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

As with the previous two models, for each ability level $a$ there is a threshold $w_{\text{min}}^{G+L}(a)$ of assets above which individuals are unconstrained and below which they are constrained. Because of the endogeneity of availability of private lending this threshold can be decreasing in $a$ which, if $\kappa$ is large, might lead to the possibility that the least able may be constrained.

\textsuperscript{33}Gropp, Scholz, and White (1997) empirically support this form of response by private lenders.

\textsuperscript{34}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 $\bar{\kappa}$. 
and the most able unconstrained. Moreover, the following Assumption 3 provides sufficient conditions under the resulting optimal investment, $h^{G+L}(a, w)$ in this model is qualitatively consistent with the empirical evidence presented earlier as long as punishments are severe enough.

**Assumption 3.** Either of the following assumptions 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 3** If $w < w^{G+L}_{\text{min}}(a)$ so constraint (8) binds, then: (i) $h^{G+L}(a, w) = h^U(a)$ for $a \leq \bar{a}$; otherwise, $h^{G+L}(a, w) < h^U(a)$; (ii) $h^{G+L}(a, w)$ is strictly increasing in $w$ for $a > \bar{a}$; (iii) If Assumption 3 holds, then $h^{G+L}(a, w)$ is strictly increasing in $a$ under Conditions 1 or 2.

The GSL ensures unconstrained optimal investment levels for all lower ability individuals regardless of wealth. For those agents, investments in human capital increase with their ability. Private loans matter because they would be better-off as they would be able to increase consumption during youth. For all the other agents with higher abilities and for which the GSL 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 lifecycle economy calibrated any empirically reasonable values of the IES.

The availability of public and private lending both play an important role in determining investment as the following proposition shows.

**Proposition 4** Let $h^{G+L}(a, w; d_{\text{max}}, \kappa)$ denote optimal investment with both the GSL (characterized by upper loan limit $d_{\text{max}} > 0$) and private markets with limited commitment (characterized by enforcement level $\kappa > 0$). Then: (i) $w^{G+L}_{\text{min}}(a) < \min \{ w^{G+L}_{\text{min}}(a), w^L_{\text{min}}(a) \}$; and (ii) $\frac{\partial h^{G+L}(a, w; d_{\text{max}}, \kappa)}{\partial d_{\text{max}}} \geq 0$ and $\frac{\partial h^{G+L}(a, w; d_{\text{max}}, \kappa)}{\partial \kappa} \geq 0$ with strict inequality for both if $a > \bar{a}$ and $w < w^{G+L}_{\text{min}}(a)$.

Both sources of credit help reducing the set of persons whose investments are constrained, but the expansion of either has differential impacts on individuals depending on their ability levels. For example, for high ability individuals an expansion of the limit $d_{\text{max}}$ of GSL loans might have a more than proportional impact on their investment, because in addition to the one-to-one impact of a higher GSL loan it would lead private lenders to increase their offer of...
credit for both consumption and further investments. However, a higher $d_{\text{max}}$ would have no impact on the consumption or investments of lesser ability individuals. On the other hand, an increase in private credit (i.e. an increase in $\kappa$) would allow for greater consumption for all individuals during the investment period, but it would only affect the investments of the most able ones.

The interaction between public and private lending is likely to lead to greater responses in investment (among some individuals) to changes in GSL policies than would otherwise be expected in the absence of private lenders. Moreover, as we will see in the following section, the repayment terms of GSL loans might also impact the amount of credit available from private lenders.

5 Quantitative Analysis

In this section, we quantitatively explore the implications of our model with public and private lending for schooling in the U.S. To this end, we extend the model to a multi-period setting and incorporate education subsidies. We calibrate the model using data on schooling costs, earnings, and other features of the U.S. economy in the early 1980s. Then, we simulate an increase in both the returns to and costs of education (reflecting key changes taking place in the 1980s and 1990s) in order to see whether the model can explain the rising importance of family resources as a determinant of schooling and the increase in student borrowing from private lenders. Overall, the model performs well when compared against the empirical patterns discussed in Section 3.

5.1 A Multiperiod Model

Assume all individuals have lives of length $T > 1$ divided into three stages: “Youth”, $t \in [0, 1]$, when individuals invest $x(t)$ in school but do not work; “maturity,” $t \in [1, P]$, when they work full-time earning $y(t)$; and “retirement,” $t \in [P, T]$, when they consume from accumulated savings.

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

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

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

We assume competitive financial markets and that the market interest rate equals the individual’s discount rate $\rho$. Consumption and investment decisions are restricted by the lifetime budget constraint

$$\int_{0}^{T} e^{-\rho t} c(t) dt + \int_{0}^{1} e^{-\rho t} x(t) dt \leq w + \int_{1}^{P} e^{-\rho t} y(t) dt.$$  (10)
Here $w \geq 0$ indicates initial financial wealth as of $t = 0$. Individuals are also endowed with an ability $a \geq 0$, which may reflect genetic traits, early family investments, and other characteristics that affect the returns on investment.

Labor earnings $y(t)$, depend on ability $a$, the discounted present value of all investments made during youth $h$, and post-school labor market experience $t - 1$:

$$y(t) = ah^{\alpha}e^{gt(t-1)},$$  \hspace{1cm} (11)

where $g \geq 0$ is the rate of return to experience. \hspace{1cm} 36

Total human capital investment, $h$, reflects both private individual investments and government investments or subsidies. We define the discounted value of accumulated private investment in human capital (valued as of time $t = 1$) $h_{\text{priv}} \equiv e^{\rho} \int_0^1 e^{-\rho t} x(t) dt$, where $x(t)$ reflects out-of-pocket individual investment flows. We assume that the government provides a free lump-sum investment $h_{\text{pub}} \geq 0$ (valued as of time $t = 1$) and also matches every unit of privately financed investment with a subsidy rate of $s \geq 0$. Altogether, this implies that total investment in human capital at the time of labor market entry is given by

$$h = h_{\text{pub}} + (1 + s) h_{\text{priv}}.$$

In what follows, we ignore the timing of investment flows $x(t)$ and focus on accumulated private ($h_{\text{priv}}$) and total ($h$) human capital investment. We are interested in how total human capital investment varies across individuals and responds to economic or policy changes.

\subsection*{5.2 Unrestricted Allocations}

With frictionless and competitive financial markets, individuals will fully smooth consumption. Optimal unconstrained total investment, $h^U(a)$, maximizes lifetime resources net of investments costs and is given by

$$h^U(a) = \max \left\{ h_{\text{pub}}, [\alpha (1 + s) a \Phi]^{\frac{1}{1-a}} \right\},$$  \hspace{1cm} (13)

where $\Phi > 0$ is a constant that depends on $P$, $g$ and $\rho$. The constant $\Phi$ converts initial earnings into the present value of life-time earnings as of $t = 1$, and its expression is given in Appendix C.

\footnote{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.}

\footnote{Our main theoretical results readily extend to the case where $g$ is increasing in $a$ (i.e. more able individuals have steeper wage profiles).}

\footnote{Implicitly, our analysis assumes that investments during youth are perfectly substitutable over time.}
As with the two-period model, unrestricted investment is independent of initial financial wealth and increasing in ability. Individuals with ability $a \leq a_0 \equiv \frac{[h_{pub}]^{1-\alpha}}{\alpha(1+s)\Phi}$ do not find it worth investing above the publicly provided amount, so $h^U_{priv}(a) = 0$ and $h^U(a) = h_{pub}$ for them. Those with $a > a_0$ invest $h^U_{priv}(a) = \left\{ \alpha (1 + s) a^\Phi \right\}^{1/\alpha} - h_{pub}$, equating the marginal return on human capital investment with its (private) marginal cost.

The unrestricted optimal allocations imply that when an individual enters the labor market ($t = 1$), total debt is equal to $d^U(a, w) = \left( \frac{1 - e^{-\rho}}{1 - e^{-\rho T}} \right) a^\Phi \left[ h^U(a) \right]^{\alpha} + \left( \frac{e^{-\rho} - e^{-\rho T}}{1 - e^{-\rho T}} \right) \left( h^U_{priv}(a) \right) (1 + s) - e^\rho w$. (14)

Following the analysis for the two-period model, it can be verified that $-1 < \frac{\partial d^U(a, w)}{\partial w} < 0$ and $\frac{\partial d^U(a, w)}{\partial a} > \frac{\partial h^U(a)}{\partial a} \geq 0$.

### 5.3 Exogenous Borrowing Constraints

We now introduce exogenous credit constraints. As in the two-period model, assume that there is an upper bound on the amount of credit that an individual can accumulate while in school:

$$d \leq d_0,$$ (15)

where $d$ is the accumulated amount of debt as of $t = 1$ and $0 \leq d_0 < \infty$. As noted earlier, we assume that credit after $t = 1$ is unconstrained.

The budget constraint during youth is $\int_0^1 e^{-\rho t} [c(t) + x] dt \leq w + e^{-\rho}d$, since own resources, $w$, plus debt, $d$, finance the flows of investment, $x$, and consumption, $c(t)$, for $t \in [0, 1]$. During youth, consumption will be constant, denoted $c_0$, since the interest rate is equal to the discount rate and the constraint (15) does not distort the intertemporal allocation of consumption within the interval $[0, 1]$. Using these results, the budget constraint during youth simplifies to

$$c_0 + x \leq \mu e^\rho \left[ w + e^{-\rho}d \right].$$ (16)

After school (i.e. in the time interval $[1, T]$), consumption is also constant at the (potentially different) level $c_1$, since post-schooling financial markets are frictionless. The value of $c_1$ is determined by $\mu^{-1} a^\Phi h^\alpha - d$, the difference between the present value of lifetime earnings and the financial liabilities carried from youth. In Appendix C, we show that $V(h, d; a)$, the person’s utility as of $t = 1$, is

$$V(h, d; a) = \Theta \left[ \mu^{-1} a^\Phi h^\alpha - d \right]^{1-\sigma},$$

where $\Theta \equiv \left[ \frac{1 - e^{-\rho(T-1)}}{\rho} \right]^{\sigma} > 0$. Discounted lifetime utility at $t = 0$ is

$$U(c_0, x, d; a) \equiv \frac{e^{-\rho} c_0^{1-\sigma}}{\mu^{1-\sigma}} + e^{-\rho} V([i_{pub} + (1 + s) x], d; a).$$ (17)
Individuals choose \( x \geq 0, \ c_0 \geq 0 \) and \( d \) to maximize \( U (c_0, x, d; a) \) subject to (15) and (16). Aside from the possibility of \( x = 0 \), this problem is analytically identical to the corresponding problem in the two-period model, and Proposition 1 holds.

### 5.4 Government Student Loan Programs

As with exogenous constraints, an analysis of GSL programs in this environment is straightforward and follows that of our two-period model. Instead of (15), cumulative debt as of \( t = 1 \), \( d \), is restricted to satisfy:

\[
d \leq \min \{ x, d_{\text{max}} \}. \tag{18}\]

Note that borrowing is tied to out-of-pocket investment, \( x \), and not to total investment, \( h \).

Individuals choose \( x \geq 0, \ c_0 \geq 0 \), and \( d \) to maximize \( U (c_0, x, d; a) \) subject to (18) and (16). Aside from the link of \( d \) to \( x \) instead of \( h \), and the possibility of \( x = 0 \), this problem is analytically identical to the corresponding problem in the two-period model. Proposition 2 holds. See details in Appendix C.

### 5.5 Private Lending with Limited Commitment

Now, consider private loans for schooling that are subject to limited enforcement. As in the two-period model, loans are repaid if and only if the cost of defaulting is higher than the cost of repaying. The incentives to repay at dates \( t \geq 1 \) define the amount of credit lenders are willing to supply during the schooling period.

We consider two penalties for default. First, defaulting borrowers are reported to credit bureaus, an action that is assumed to prevent the borrower from accessing formal credit markets for some period. This penalty does not reduce earnings, but it disrupts the ability to smooth consumption and can be quite costly if labor earnings grow quickly with age or if the IES is low. Second, the borrower forfeits a fraction \( \gamma \in [0, 1) \) of his labor earnings. The fraction \( \gamma \) 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 \( 0 < \pi < P - 1 \) that starts the moment default takes place.\(^{38}\)

Consistent with our earlier assumptions about post-investment credit markets, we assume that loans contracted after schooling are fully enforceable and that loans contracted while in school can only be defaulted on at age \( t = 1 \). We explicitly focus on credit constraints directly related to the financing of schooling.\(^{39}\)

The amount of debt that a person can credibly commit to repay depends on the discounted utility associated with default. Consider a person with ability \( a \) and human capital \( h \) that

---

\(^{38}\)Livshits, MacGee, and Tertilt (2007) make a similar set of assumptions in modelling U.S. bankruptcy regulations.

\(^{39}\)Monge-Naranjo (2007) considers a continuous time model in which the agent can default in any period.
defaults at $t = 1$ on debt $d$. Since punishments are not reduced by partial re-payment, all defaults would be on the entire debt. During the punishment period $[1, 1 + \pi]$, consumption is $c(t) = (1 - \gamma) \mu^{-1} ah^\alpha e^{\theta(t-1)}$. From $t = 1 + \pi$ onwards, a fresh start allows the person to fully smooth consumption, including the time after retirement. The maximized $t = 1$ discounted utility of a person who chooses to default at the end of the schooling period is

$$V^D(h; a) = \hat{\Theta}_{\gamma, \pi} \frac{[\mu^{-1} ah^\alpha]^{1-\sigma}}{1-\sigma},$$

where $\hat{\Theta}_{\gamma, \pi}$ is a positive constant, the expression for which is shown in Appendix C.

Rational lenders foresee the repayment incentives of borrowers and restrict credit to avoid triggering default. Given penalties $(\pi, \gamma)$ a borrower with ability $a$ and human capital investment $h$ is better off repaying a level of debt $d$ when $V^D(h; a) \leq V(h, d; a)$. In our setting, this is equivalent to:

$$d \leq \kappa \mu^{-1} a \Phi [i_{pub} + (1 + s) x]^\alpha,$$

where $\kappa \equiv 1 - \left[ \hat{\Theta}_{\gamma, \pi} / \Theta \right]^{1-\sigma} \geq 0$ incorporates the effects of both the garnishment and distortions to consumption profiles. Notice that credit limits are proportional to labor earnings, just as we imposed in the two-period model. However, in this model, the value of $\kappa$ is determined by preferences $(\rho, \sigma)$ and institutions $(\gamma, \pi)$. Below we show that $\kappa$ can be large even if wage garnishments, $\gamma$, are negligible.

Some aspects of the determination of $\kappa$ are worth mentioning. First, $\kappa$ is increasing in $\gamma$ and $\pi$. The option to default is less tempting with harsher punishments. Second, $\kappa > 0$ as long as $\pi > 0$, even if $\gamma = 0$. The exclusion from financial markets alone suffices to sustain lending. Third, if $\pi = 0$, the model boils down to an exogenous constraint model with $d_0 = 0$, since no lending can be sustained in equilibrium, i.e. $\kappa = 0$.

With credit limits determined by limited commitment, a person chooses investment, consumption and borrowing $(x \geq 0, c_0 \geq 0$ and $d)$ to maximize $U(c_0, x, d; a)$ subject to (16) and (19). Given the endogenously determined $\kappa$ and ignoring the possibility of $x = 0$, this problem is analytically identical to the two-period case and a parallel to Proposition ?? holds.

**Proposition 5** Let ability and financial assets of a young individual be $(a, w)$, and let $h^L(a, w)$ and $h^U(a)$ indicate, respectively, the optimal investments in human capital with private lenders with limited commitment and in the unrestricted allocation. If $a \leq a_0$, then $h^L(a, w) = h^U(a) = i_{pub}$. If instead $a > a_0$ and constraint (19) binds, then: (i) $h^L(a, w) < h^U(a)$; (ii) $h^L(a, w)$ is strictly increasing in $w$; and (iii) $h^L(a, w)$ is strictly increasing in $a$ if $\kappa \geq \kappa(\sigma) \equiv [(\sigma - 1) / \sigma] [(1 - e^{-\rho}) / (1 - e^{-\rho T})].$

This proposition is central to our quantitative analysis of credit constraints, given its empirically verifiable predictions.
Recall from our discussion in Section 4.3 that strong preferences for smooth consumption (i.e. a high $\sigma$ or low IES) generate a negative ability–investment relationship when credit constraints are exogenously fixed. This tendency also exists when constraints are endogenous since $\frac{\partial \kappa(\sigma)}{\partial \sigma} \geq 0$, which implies that a stronger link between investment and credit limits (i.e. a larger $\kappa$) is needed to generate 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, $\kappa$ is also increasing in $\sigma$ under limited commitment. Below, we show that for empirically plausible punishment parameters ($\gamma, \pi$) the effect of $\sigma$ on $\kappa$ often dominates its effect on $\kappa(\sigma)$, and a higher value of $\sigma$ (or lower IES) makes it more rather than less likely that condition (iii) of this proposition holds. Most importantly, we show that $\kappa > \kappa(\sigma)$ for empirically plausible parameters, so that the model implies a positive relationship between investment and ability for all ability and initial wealth levels.

5.6 GSL Programs Plus Private Lenders

Finally, we consider the coexistence of a GSL program with private lenders. As before, we assume that the repayment of loans from the GSL program is fully enforced. As noted earlier, this assumption is in line with the fact that GSL programs in the U.S. are better protected against default than private unsecured loans. First, private loans can be cleared in bankruptcy proceedings while GSL loans cannot. This implies a longer (potentially unlimited) punishment period for government loans. Second, wage garnishments of up to 15% are explicitly incorporated in GSL programs, whereas no explicit rate exists for private unsecured loans. Third, GSL programs include a wide array of additional punishments for those who default that have no counterpart in the private sector (e.g. governments can seize income tax returns for those defaulting on a GSL program loan).

With both public and private sources of credit, a young person with $(a, w)$ chooses out-of-pocket investment $x$, consumption during youth $c_0$, borrowing from GSL $d_g$, and borrowing from private lenders $d_p$ to maximize $U(c_0, x, d; a)$ subject to (16) and

$$d \leq d_p + d_g$$

$$d_p \leq \kappa \mu^{-1} a \Phi \left[i_{pub} + (1 + s) x\right]^\alpha$$

$$d_g \leq \min \{x, d_{max}\}.$$  

The threshold level of initial assets $w_{G+L}^{G+L}(a)$ above which individuals with ability $a$ are unconstrained is given in Appendix C. As with the two-period model, this threshold is lower and fewer people are constrained than under the GSL alone or under private markets alone. Proposition ?? of the two-period model also applies in this setting.
Table 3: Baseline Model Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>To match:</th>
<th>Parameter</th>
<th>Value</th>
<th>Estimates for:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$</td>
<td>6.125</td>
<td>Retirement at 65</td>
<td>$g$</td>
<td>0.369</td>
<td>Experience</td>
</tr>
<tr>
<td>$T$</td>
<td>8</td>
<td>Lifespan of 80</td>
<td>$\alpha$</td>
<td>0.432</td>
<td>Schooling investment</td>
</tr>
<tr>
<td>$\pi$</td>
<td>0.875</td>
<td>7 yrs., U.S. Bankruptcy</td>
<td>$\rho$</td>
<td>0.3138</td>
<td>Annual rate = 4%</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.3138</td>
<td>Annual rate = 4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma$</td>
<td>2</td>
<td>IES = 0.5</td>
<td></td>
<td></td>
<td>Ability Levels:</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.1</td>
<td>Garnishment &amp; other costs</td>
<td>$\tilde{a}_1$</td>
<td>106.70</td>
<td>AFQT quartile 1</td>
</tr>
<tr>
<td>$a_{pub}$</td>
<td>65,239</td>
<td>Educ. costs through grade 9</td>
<td>$\tilde{a}_2$</td>
<td>137.83</td>
<td>AFQT quartile 2</td>
</tr>
<tr>
<td>$s$</td>
<td>1.19</td>
<td>Educ. subsidy grades 10+ = 54%</td>
<td>$\tilde{a}_3$</td>
<td>157.38</td>
<td>AFQT quartile 3</td>
</tr>
</tbody>
</table>

5.7 Calibration

We now calibrate parameter values to explore the quantitative implications of the model. Some of the parameters are estimated using data on earnings and educational attainment from the random sample of males in the NLSY79. Other parameter values are calibrated to replicate features of the U.S. economy. We use AFQT quartiles to measure ability. All dollar amounts are denominated in 1999 dollars using the Consumer Price Index (CPI-U). Table 3 reports the value of all parameters used in our baseline simulations.

We assume that youth (investment period) begins at age 16 and ends at age 24. Maturity (labor market participation period) runs from age 24 until age 65. Retirement runs from age 65 until death at age 80. Since youth lasts one period in the model, each interval of unit length corresponds to 8 years of life. Dates for retirement and death are, respectively, $P = (65 - 16)/8 = 6.125$, and $T = (80 - 16)/8 = 8$.

To match an annual interest rate of 4%, we set $\rho = 8 \times \ln(1.04) = 0.319$ for the value of the discount rate in the model. We set $\sigma = 2$ as our baseline value to match an IES of 0.5, an intermediate value in the estimates in Browning, Hansen, and Heckman (1999). Values of $\sigma$ inside the interval $[1.5, 3]$ yield similar results.

Based on U.S. bankruptcy regulations, we set $\pi = 7/8 = 0.875$ (7 year penalty period) as the baseline length of penalties. 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 suboptimally employed, renting instead of owing a house, and paying subprime interest rates for short-term transactions, etc.) Since the implied $\kappa$ varies little with $\gamma$, our results are not sensitive to reasonable variations in this parameter.
We assume all investment through age 16 is publicly provided for free. After age 16, schooling entails direct costs (i.e. tuition and public expenditures on primary, secondary, and post-secondary schooling) and indirect costs (i.e. foregone earnings). To compute direct costs we use an annual government expenditure of $5,928 for primary and secondary schooling. Annual direct expenditures for college and graduate education are assumed to equal $16,838.\(^{40}\) (College expenditures include government expenditures as well as tuition and fees paid by students.) We set \(i_{pub} = 65,239\), which is equal to the discounted value of all direct schooling expenditures through grade nine.\(^{41}\) This is consistent with our focus on investments made from age 16 onwards.

Because of the laws on compulsory schooling and minimum work age, we only include foregone earnings as part of investment costs for grades ten and above. To estimate foregone earnings, we use data from the NLSY79 to regress log earnings on indicators for each possible year of completed schooling from grade 10 through six years of post-secondary studies, indicators for AFQT quartiles, total years of potential work experience and experience-squared.\(^{42}\) From this regression, we compute foregone earnings for \(S \geq 10\) years of schooling using the predicted earnings of someone with nine years of completed schooling, \(S - 10\) years of potential work experience, and the desired AFQT quartile. These foregone earnings estimates are included in our estimates of total schooling expenditures and are reported in Table D1 of the Appendix.

We calibrate the government subsidy rate \(s\) as follows. We assume that the private costs of investment are foregone earnings plus a fraction of post-secondary tuition and other direct costs. Table 333 of the Digest of Education Statistics (2003) reports that tuition and fees accounted for 20% of total current-fund revenue for degree-granting higher education institutions in 1980. Assuming a ratio of 0.20 for private to total direct expenditures for college, the subsidy rates for investment education beyond \(i_{pub}\) range from 0.47 to 0.6 for completed schooling levels 12–16 depending on the AFQT quartile and completed years of schooling. As our baseline, we use a subsidy rate of 54% (i.e. \(s = 1.19\)) based on the average government subsidy rate for individuals completing 2 years of college; however, our results are very similar when we use other reasonable values.

We estimate the parameters \(g\) and \(\alpha\) of the earnings equation using data from the

\(^{40}\)Annual expenditure for education through grade twelve (for college and graduate education) is the average of annual current expenditures per pupil for public primary and secondary schools (for all two and four year colleges) over the academic years 1979-80 through 1988-89 as reported in Table 170 (Table 342) of the Digest of Education Statistics, 1999. These years roughly reflect the years our NLSY79 sample respondents made their final schooling decisions.

\(^{41}\)We use a 4% annual discount rate, reporting the value discounted to the end of grade nine. Less than 0.2% of our NLSY79 sample acquired less than 10 years of school.

\(^{42}\)This regression uses all available earnings observations for male respondents with at least nine years of completed schooling when they were ages 16-24 and no longer enrolled in school. Potential work experience is measured as age - years of completed schooling - 6. The estimates (available upon request) suggest that earnings for these young workers are generally increasing in years of completed schooling and increasing and concave in potential work experience.
From the model, the wage earnings of someone with ability \( a \) who invested \( h \) and has been working \( \tau \) periods is
\[
y(\tilde{a}, h, \tau) = \tilde{a}h^\alpha e^{g\tau}
\]
where \( \tilde{a} \equiv a/\mu \). This is log-linear, so we regress log earnings for individual \( i \) on AFQT quartile indicators \( (A_i) \), estimated total schooling expenditures \( (h_i) \) as reported in Table D1, and years of experience \( (\tau = \text{age} - 24) \):
\[
\ln[y_{i\tau}] = \beta_0 A_i + \beta_1 \ln(h_i) + \beta_2 \tau + \nu_{i\tau},
\]
where \( \nu_{i\tau} \) is a mean zero idiosyncratic earnings shock, i.e. \( E(\nu_{i\tau}|A_i, h_i, \tau) = 0 \). The implied estimates for \( \alpha \) and \( g \) are, respectively, \( \hat{\alpha} = \hat{\beta}_1 \) and \( \hat{g} = 8\hat{\beta}_2 \) (recall that a unit time interval in the model corresponds to 8 years in the data.) Even though \( \nu_{i\tau} \) is mean zero, \( E[e^{\nu_{i\tau}}] > 1 \), so we adjust the coefficient vector \( \hat{\beta}_{0q} \) on AFQT quartiles using the sample average \( \overline{e^{\nu_{i\tau}}} \). That is, our ability estimate for quartile \( q \) is
\[
\tilde{a}_q = e^{\hat{\beta}_{0q}\overline{e^{\nu_{i\tau}}}}.
\]
These ability estimates range from 106.7 for the least able to 158.3 for the most able, suggesting that, for the same schooling, the most able are on average about 50% more productive than are the least able.

Ideally, we would like to specify a joint distribution of wealth and ability to simulate our model and compute the distribution of investment in the economy. Unfortunately, this is not feasible. We do not directly observe the assets available to youth, because they not only depend on their parents net worth and income but also on intra-family transfers, which are not always observed. Modelling these transfers is beyond the scope of this paper. Instead, we analyze investment behavior for our estimated ability types and a range of potential wealth levels. While we cannot compare moments for investment implied by the model with those observed in the data, we can explore whether investment is increasing in ability and wealth over reasonable ranges of wealth, how investment behavior depends on the type of constraints we assume, and how investment and borrowing (as functions of ability and wealth) respond to changes in the economy.

Because foregone earnings are an important part of investment expenditures in our calibration, an individual’s initial wealth, \( w \), includes at least the amount he could earn if he left school after grade 9 and began working. This amount depends on ability, since foregone earnings depend on ability (see Appendix Table D1). The relevant range of initial wealth, therefore, begins at $52,000 for the least able, $74,000 for AFQT quartile 2, and $80–84,000 for the top two quartiles. Any wealth levels above these amounts must come from parents or other outside sources.

### 5.8 Baseline Simulations

We are primarily interested in the implied cross-sectional relationship between ability, wealth, and investment in our model with both the GSL and private lenders. We begin with a discussion of borrowing constraints in this environment and then discuss investment behavior.

Figure 5 shows a very strong result: for any value of \( \sigma \geq 0 \), the limited commitment model implies a positive relationship between investment and ability given our values for
Figure 5: Sufficient Condition for a positive ability-investment relationship in the $L$ model.

The figure displays the value of $\kappa$ (the fraction of future earnings that can be borrowed from private lenders) associated with different values of $\sigma$ under our baseline parameterization (thick green line) and under alternative assumptions about punishments $\gamma$ and $\pi$. The figure also displays $\kappa(\sigma)$ as defined in Proposition 5 (dashed line). When $\kappa \geq \kappa(\sigma)$, investment is increasing in ability. Notice, $\kappa$ exceeds $\kappa(\sigma)$ for any value of $\sigma$ under our baseline parameters. This is also true if the only penalty for default is a seven-year exclusion from financial markets ($\pi = 7/8$, $\gamma = 0$) or if the exclusion period lasts only one year ($\pi = 1/8$) and $\gamma = 0.1$. When no penalties for default exist (i.e. $\pi = \gamma = 0$), the limited commitment model is equivalent to an exogenous constraint model with $d_0 = 0$, and there is only a positive relationship between ability and investment if $0 \leq \sigma < 1$ (IES $> 1$).

As discussed earlier, both $\kappa$ and $\kappa(\sigma)$ are increasing in $\sigma$. Over most of the empirically relevant range where $\sigma > 1$ (IES $< 1$), an increase in $\sigma$ makes it more rather than less likely that the condition in Proposition 5 is met and investment is increasing in ability.

Figure 6 shows the implied borrowing constraints as a function of individual investment (not including government subsidy amounts) for the GSL program. The figure also shows the private lending constraints as a function of individual investment for all four ability groups. While offering sizeable loans to students, private lenders do not offer as much as the GSL program over a wide range of investment amounts. At any level of investment, the difference in private lending limits between the most and least able is sizeable, ranging from about

$43^{43}$Since the IES equals the inverse of $\sigma$, the empirically relevant range is $\sigma \geq 1$. 

29
$5,000 to $10,000. These limits are also strongly increasing in investment (more so for the most able).

The amount of borrowing from private lenders as a function of ability and initial assets is shown in Figure 7 for our baseline economy with both private lending and a GSL program. We assume that individuals borrow first from the GSL program and then, when this source is exhausted, they may borrow from private lenders. As one can see, our calibrated model implies that private borrowing should be negligible during the baseline period (i.e. for NLSY79 cohorts). As noted earlier, initial assets for all individuals are at least as high as the potential earnings they would receive if they quit school after grade 9: $52,000 for the lowest AFQT quartile, $74,000 for the second lowest, and $80–84,000 for the top two quartiles. For all ability quartiles, borrowing from private lenders is zero for asset levels above potential school-period earnings.

For all ability types, individuals with initial assets above $40,000 (far below potential school-period earnings) invest the unconstrained optimal amounts. As a result, investment is increasing in ability and independent of initial assets, consistent with reported schooling patterns in the NLSY79 data. Optimal total human capital investment ranges from roughly $85,000 for the least able to $130,000 for the second ability quartile to around $165,000 for the top two ability quartiles.\footnote{With private lending markets alone, the most able are constrained up through initial asset levels of around $85,000. With the GSL alone or with exogenous constraints, the most able are constrained through}
It is noteworthy that the investment amounts implied by the model are fairly close to average total expenditures by AFQT quartile in the NLSY79 data, even though we did not target these values. This external validity provides additional confidence in our model and the baseline parameterization.

### 5.9 A Rise in the Costs of and Returns to Schooling

We now simulate the effects of an increase in the costs of and returns to schooling — two major economic changes that took place between the early 1980s and early 2000s. We aim to see whether the model can reproduce the observed rise in private lending as well as the increased effects of family income on educational attainment. We also compare the investment and consumption allocations under different assumptions about credit markets. This sheds light on the importance of a GSL program for investment, as well as the role played by private lenders today. We also compare these environments with the standard model, which assumes borrowing constraints are exogenous.

We model an increase in the wage returns to education by assuming that $\alpha$ increases initial asset levels of around $70,000$. In all cases, investment is unconstrained for asset levels above potential school-period earnings.

45Combining the total costs by AFQT quartile and schooling level reported in Table D1 with the distribution of educational attainment by AFQT in the NLSY79, we obtain average total investment amounts ranging from $88,000$ for the least able to $178,000$ for the most able.
by 0.02 (from the baseline estimate of 0.432). This change produces a modest increase in the college − high school log wage differential. We model the rise in net tuition costs by assuming that the government subsidy rate, $s$, falls from 1.19 to 1.05. This reduction reflects the increased importance of tuition and fees as a fraction of total current-fund revenue for public and private universities in the U.S. Finally, we incorporate the stability of maximum GSL loan limits by assuming that $d_{max}$ remains unchanged at $35,000.

Figure 8 graphs the new private lending limits against the unchanged GSL limits. For all investment and ability levels, private credit limits increase by at least $3,000 over the baseline amounts, with much larger increases at higher investment amounts. This increase is entirely driven by the increased return to investment, which raises the monetary costs of default and, therefore, the amount students can commit to repay.

Youth wish to borrow more in response to increases in the costs and returns to investment. With no increase in credit available from the GSL, private lenders must cover the increased demand for credit. Indeed, Figure 9 shows that private lending increases substantially. While the least able youth still do not borrow from private lenders, high ability youth with low asset levels now borrow as much as $50,000.

Figure 10 shows investment behavior under the higher assumed costs and returns to school (i.e. $\alpha = 0.453$ and $s = 1.05$) for the four different models of credit constraints (GSL plus private lenders, private lenders alone, GSL alone, and exogenous constraints with $d_0 = d_{max}$). First, panel (a) considers investment in our preferred model with a GSL program
and private lending. Investment increases noticeably relative to the baseline economy. This increase would have been even greater if private costs had not risen as well. Borrowing constraints now appear to be binding for a broad range of initial asset levels among the higher ability types. Individuals in the top two ability quartiles with assets below $100,000 ($15-20 thousand above potential school-period earnings) are constrained. Individuals in the second lowest ability quartile with initial assets below $60,000 (less than potential school-period earnings) are constrained. The least able are unconstrained for all reported levels of wealth. Consistent with the NLSY97 data, family resources have become more important for investment among a broader set of individuals.

Now, contemplate eliminating either source of credit. Panel (b) of Figure 10 considers private lending alone, while panel (c) considers the GSL program alone. In both cases, individuals from a much broader range of initial assets and abilities are constrained and invest less than in panel (a) when both sources are present. For most initial asset and ability levels, the private lending market and GSL yield fairly similar investment levels; however, this is not true for those with very low initial assets (amounts below potential school-period earnings). Under the GSL, these youth would invest the maximum amount they can borrow, $35,000, above the publicly provided amount. Investing less does not provide them with any more consumption while in school, since they would also be required to borrow less. Private lenders do not impose this tight restriction, so very poor youth would invest less even though they may be able to borrow more than under the GSL.
Figure 10: Total Investment (‘Year 2000’) with Different Credit Market Assumptions

(a) GSL and private lending

(b) Private Lending

(c) GSL

(d) Exogenous borrowing limits
Notice that both models which incorporate private lending (panels a and b) imply a positive relationship between ability and investment for all levels of assets; although, investment is quite similar across ability types for very low asset levels (below potential school-period earnings). This is not the case for the GSL alone (panel c), since the upper limit on borrowing is the binding constraint for a broad range of initial asset levels and ability types. The perverse relationship between ability and optimal investment is even worse for the exogenous constraint model as shown in panel (d).

Finally, we show consumption during the investment period under all four credit market assumptions in Figure 11. Consumption is substantially higher when both the GSL and private lending markets are available than when either is not. As expected from our discussion of the GSL program in Section 4, consumption while in school is quite low for those with low initial assets. The fact that borrowing cannot be used to finance consumption under the GSL can be quite costly for the poor in the absence of private lending. All other forms of credit constraints allow for more intertemporal consumption smoothing, even if it is at the expense of lower investment.

6 Conclusions

This paper develops a lifecycle 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 hu-
Figure 11: Consumption during Investment Period (‘Year 2000’) with Different Credit Market Assumptions

(a) GSL and private lending  (b) Private Lending

(c) GSL  (d) Exogenous borrowing limits
man 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. 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. 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. 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.

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.

---

46 See Belley and Lochner (2007) for additional details on the sample and variables used in this paper.
47 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).
48 Family income includes government transfers (e.g. welfare and unemployment insurance), but it does not subtract taxes. Net wealth measures the value of all assets (e.g. home and other real estate, vehicles, checking and savings, and other financial assets) less any loans and credit card debt.
49 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.
Exogenous Constraints: The threshold \( w^X_{\text{min}}(a) \) is defined by \( d^U(a, w^X_{\text{min}}(a)) = d_0 \), so it is increasing in \( a \). Note that \( w^X_{\text{min}}(a) \geq h^U(a) - d_0 \), the wealth level needed to finance \( h^U(a) \) given maximum borrowing. Consumption smoothing further implies that \( w^X_{\text{min}}(a) \) is steeper than \( h^U(a) \) as a function of \( a \), since \( \frac{dw^X_{\text{min}}}{da} = \frac{\partial d/U(a, w_{\text{min}})}{\partial a}(a) \frac{\partial d/U(a, w_{\text{min}})}{\partial w} > \frac{\partial d/U(a, w_{\text{min}})}{\partial a} > 0 \) by implicit differentiation.

GSL Programs: The threshold \( w^{G}_{\text{min}}(a) \equiv \max\{w^X_{\text{min}}(a), \tilde{w}_{\text{min}}(a)\} \), where \( \tilde{w}_{\text{min}}(a) \) is defined by \( h^U(a) = d^U(a, \tilde{w}_{\text{min}}(a)) \). It is increasing in \( a \) because \( d^U(\cdot, w) \) is steeper than \( h^U(\cdot) \). To see that \( w^X_{\text{min}}(a) \) is steeper than \( \tilde{w}_{\text{min}}(a) \), use implicit differentiation to obtain \( \frac{dw^X_{\text{min}}}{da} = \frac{dw^X_{\text{min}}}{da} + \frac{\partial h^U}{\partial a} \frac{\partial d^U}{\partial w} < \frac{dw^X_{\text{min}}}{da} \) because \( \frac{\partial h^U}{\partial a} \leq 0 \).

Private Lending with Limited Commitment: The threshold \( w^{L}_{\text{min}}(a) \) is defined by \( d^U(a, w^L_{\text{min}}(a)) = \kappa f[h^U(a)] \). This threshold increases at a slower rate in \( a \) than does \( w^X_{\text{min}}(a) \), and it may even be decreasing in \( a \) if \( \kappa \) is large enough. To see this, use implicit differentiation and obtain \( \frac{dw^L_{\text{min}}}{da} = \frac{dw^X_{\text{min}}}{da} + \kappa \left( f(h^U) + R \frac{\partial h^U}{\partial a} \right) \frac{\partial d^U}{\partial w} < \frac{dw^X_{\text{min}}}{da} \) because \( \frac{\partial h^U}{\partial a} \leq 0 \).

GSL Programs Plus Private Lenders: The threshold \( w^{G+L}_{\text{min}}(a) \) is defined by \( d^U(a, w^{G+L}_{\text{min}}(a)) = \kappa f[h^U(a)] + \min \{h^U(a), d_{\text{max}}\} \). Direct inspection implies that \( w^{G+L}_{\text{min}}(a) < \min \{w^G_{\text{min}}(a), w^L_{\text{min}}(a)\} \). As with \( w^L_{\text{min}}(a) \), the threshold \( w^{G+L}_{\text{min}}(a) \) can be decreasing in \( a \) and may even be negative.

### B.2 Proofs

#### Proof of Lemma 1.\footnote{From the FOC define \( F \equiv u'(w + d - h^U(a)) - \beta Ra' \left[ f[h^U(a)] - Rd \right] = 0. \)} Implicit differentiation of (4) yields \( \frac{dh^U(a)}{da} = \frac{f'[h^U(a)]}{a f''[h^U(a)]} > 0. \) Using expression (5), define

\[
F \equiv u'[w + d - h^U(a)] - \beta Ra'[f[h^U(a)] - Rd] = 0.
\]

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

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

where we have used \( a f'[h^U(a)] = R \). Similarly,

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

Since the denominator is greater than one, the argument is complete. \( \blacksquare \)

#### Proof of Proposition 1. From the FOC define

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

From the second order condition \( \partial^2 F/\partial h < 0 \) and then, from implicit differentiation \( \text{sign} \{ \frac{\partial h}{\partial w} \} = \text{sign}\{ \frac{\partial w}{\partial a} \} = \text{sign}\{ \frac{\partial F}{\partial w} \} \). First, \( \frac{\partial h}{\partial w} > 0 \) since \( \frac{\partial F}{\partial w} = -u''(w + d_0 - h) > 0 \). Second,

\[
\frac{\partial F}{\partial a} = \beta f'[h] u'[af(h) - Rd_0] \left\{ 1 + af(h) \frac{u''[af(h) - Rd_0]}{u'[af(h) - Rd_0]} \right\}
\]

\[
< \beta f'[h] u'[af(h) - Rd_0] \left\{ 1 + [af(h) - Rd_0] \frac{u''[af(h) - Rd_0]}{u'[af(h) - Rd_0]} \right\}
\]

\[
= \beta f'[h] u'[af(h) - Rd_0] \left\{ 1 - 1/\eta[af(h) - Rd_0] \right\},
\]

39
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 \( \eta(\cdot). \) 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 2.** Using the FOC of the exogenous constraint model,
\[
\hat{a}(w) \equiv \sup \{ \hat{a} : u'(w) \geq \beta \hat{a} f' [d_{\text{max}}] u' [\hat{a} f (d_{\text{max}}) - Rd_{\text{max}}] \},
\]
which in principle could be \( +\infty. \) If \( u(c) = c^{1-\sigma}/(1 - \sigma), \) then a finite \( \hat{a}(w) \) would be given by
\[
\hat{a} : w \left( \beta f' [d_{\text{max}}] \right)^{\frac{1}{\sigma}} = \left( \hat{a} \right)^{\frac{\sigma - 1}{\sigma}} f (d_{\text{max}}) - Rd_{\text{max}} \left( \hat{a} \right)^{-\frac{1}{\sigma}}.
\]
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 2.** Straightforward and omitted. ■

**Proof Proposition 3.** (NOTE: This is only a partial proof based on an earlier version of the paper – an updated version coming soon.) The proof of part (i) follows naturally from previous proofs and similar arguments. Part (ii) follows from the discussion of the GSL and implicit differentiation of the following first order condition for investment as in Proposition 1.

We prove (iii) when \( d_{\text{max}} = 0 \) based on earlier results, but we will update this soon for general \( d_{\text{max}} > 0. \) From the FOC for investment when the constraint binds, define
\[
F \equiv (\kappa a f'(h) - 1) u'(w + \kappa a f(h) - h) + \beta a f'(h) (1 - \kappa R) u' [a f(h) (1 - \kappa R)] = 0.
\]
From the second order condition, \( \partial F/\partial h < 0 \) and \( \text{sign}\{\partial h/\partial a\} = \text{sign}\{\partial F/\partial a\}. \) After some simplification:
\[
\frac{\partial F}{\partial a} = \kappa f'(h) u'(c_0) + (1 - \kappa a f'(h)) \kappa f(h) [u''(c_0)] + \beta (1 - \kappa R) f'(h) \{u'(c_1) + c_1 u''(c_1)\}.
\]
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
\[
\frac{\partial F}{\partial a} = u'(c_1) \left[ \left( \frac{u'(c_0)}{u'(c_1)} \right) \kappa f'(h) + (1 - \kappa a f'(h)) \kappa \left( \frac{-u''(c_0)}{(1 - \kappa a f'(h)) u'(c_0)} \right) f(h) + \beta (1 - \kappa R) f'(h) [1 - \sigma(c_1)] \right],
\]

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

\[
\geq \beta u'(c_1) f'(h) \left[ \kappa c_1 \frac{1}{c_0 \eta(c_0)} + 1 - \frac{1}{\eta(c_1)} (1 - \kappa R) \right],
\]
where the second uses \( c_1 = (1 - \kappa R) a f(h) \) 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 \( \kappa c_1 \frac{1}{c_0 \eta(c_0)} \) 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,
\[
\frac{\partial F}{\partial a} \geq \beta u'(c_1) f'(h) \left[ 1 - \frac{1}{\eta(c_0)} [1 - \kappa (1 + R)] \right],
\]
which is strictly positive if \( \kappa \geq [1 - \eta(c_0)] / (1 + R). \) ■
Proof of Proposition 4. 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 $F(h, d_{\text{max}}, \kappa)$ as

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

The first order condition that determines $h^{G+L}(a, w; d_{\text{max}})$ is $F = 0$ and $\partial F / \partial d_{\text{max}} = -[\partial F / \partial h] / [\partial F / \partial h]$. Since $h$ is optimally chosen, $\partial F / \partial h < 0$ and $\text{sign} \{ \partial h^{G+L}(a, w; d_{\text{max}}) / \partial d_{\text{max}} \} = \text{sign} \{ \partial F / \partial d_{\text{max}} \}$, where

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

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

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

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

\[\blacksquare\]

C Proofs and Other Aspects of the Quantitative Model

C.1 Unrestricted Allocations

Given $(a, w)$ an individual maximizes the $t_0 = 0$ value of utility (9) subject to

$$\int_0^T e^{-\rho t} c(t) \, dt + \int_0^1 e^{-\rho t} x(t) \, dt \leq w + \int_1^P e^{-\rho t} y(t) \, dt. \quad (23)$$

The definition of $\Phi$ is

$$\Phi \equiv \begin{cases} \left[ e^{(g-\rho)(P-1)} - 1 \right] / (g - \rho) & \text{if } g \neq \rho \\ P - 1 & \text{if } g = \rho. \end{cases}$$

Optimal out-of-pocket investment is

$$x^U(a) = \arg\max_{x \geq 0} \left\{ w + \frac{e^{-\rho}}{\mu} [a \Phi [i_{\text{pub}} + (1 + s) x]^\alpha - x] \right\}.$$

Since total schooling investment is given by (13), it clear that if $a \leq a_0 \equiv \frac{[i_{\text{pub}}]^{1-\alpha}}{\alpha(1+s)\Phi}$, then $x^U(a) = 0$ and $h^U(a) = i_{\text{pub}}$. Finally, the optimal unconstrained consumption is constant and equal to

$$c^U(t, a, w) = \frac{\rho}{1 - e^{-\rho T}} \left\{ w + \frac{e^{-\rho}}{\mu} [a \Phi [i_{\text{pub}} + (1 + s) x^U(a)]^\alpha - x^U(a)] \right\}.$$

C.2 Exogenous Constraint Model

The expression for $w^X(a)$ is given by

$$w^X(a) = \frac{1 - e^{-\rho}}{\mu (1 - e^{-\rho(T-1)})} a \Phi [h^U(a)]^\alpha + \frac{h^U(a) - i_{\text{pub}}}{\mu (1 + s)} - d_0 \left( \frac{1 - e^{-\rho T}}{1 - e^{-\rho(T-1)}} \right).$$

Everything else is the same as in the basic model.
C.3 GSL Model

The threshold level of initial assets $u^G_{\text{min}}(a)$ above which individuals with ability $a$ are unconstrained satisfies $d^U(a, u^G_{\text{min}}(a)) = \min \{ d_{\text{max}}, \max \{ 0, (h^U(a) - i_{\text{pub}}) / (1 + s) \} \}$, where $d^U(a, w)$ is given by expression (14) and $h^U(a)$ by expression (13).

C.4 Private Lending with Limited Commitment

The highest discounted utility that can be attained by an individual that defaults at $t = 1$ is $V^D(a, h) = \Theta(\gamma, \pi)^{1-\sigma} \left( \frac{1}{\Phi} \right)^{1-\sigma} \left( \frac{\left[ e^{\rho(1-\sigma)} - 1 \right] \gamma}{g(1-\sigma) - \rho} \right) + e^\rho \left( \frac{e^{\rho(1+\pi)} - e^{-\rho T}}{\rho} \right) \gamma^{1-\sigma} \left( \frac{e^{(g-\rho)(1+\pi)} - e^{(g-\rho)T}}{\Phi(g - \rho)} \right)^{1-\sigma}.

Claims about $\kappa$ follow directly from: (i) $\Theta(\gamma, \pi) < \Theta$ for $\gamma, \pi > 0$; (ii) $\Theta(\gamma, \pi)$ is decreasing in $\gamma$; (iii) for all $\gamma \in (0, 1)$, $\Theta(\gamma, \pi)$ converges to $\Theta$ as $\pi \to 0$.

As of $t = 0$, the maximization problem consists of choosing a consumption $c_0$ for all $t \in [0, 1]$, and investment and borrowing levels $(x, d)$, such that

\[ BC : \frac{e^{-\rho}}{\mu} [c_0 + x] \leq w + e^{-\rho}d, \]

\[ CC : d \leq \kappa \left[ \mu^{-1} a \Phi[i_{\text{pub}} + (1 + s) x] \right]^\alpha. \]

Aside from government subsidies $(s, i_{\text{pub}})$ and the determination of $\Theta, \Phi,$ and $\kappa$, this problem is equivalent to the two-period model of Section ??.

The value $w^L_{\text{min}}(a)$ defined by $d^U(a, w^L_{\text{min}}(a)) = \kappa \mu^{-1} a \Phi_a [h^U(a)]^\alpha$ is the threshold of wealth above which an agent is unconstrained. It is equal to

\[ w^L_{\text{min}}(a) = \min \begin{cases} \frac{a \Phi [i_{\text{pub}}] \alpha [1 - e^{-\rho} - \kappa (1 - e^{-\rho T})]}{\mu (1 - e^{-\rho T - 1})} & \text{for } a \leq a_0 \\ h^U(a) \left[ \frac{1 - \kappa (1 - \alpha) e^{-\rho} + \kappa (1 - \alpha) e^{-\rho T}}{\mu \alpha (1 + s) (1 - e^{-\rho T - 1})} \right] - \frac{e^{-\rho}}{\mu} \left( \frac{i_{\text{pub}}}{1 + s} \right) & \text{for } a > a_0. \end{cases} \]

Individuals with $w \geq w^L_{\text{min}}(a)$ attain the unrestricted allocations. For those with $w < w^L_{\text{min}}(a)$, constraint (19) holds with equality and we can use it to eliminate $d$. With this, the problem becomes

\[ \max_{\{x : x \geq 0\}} \left\{ \frac{e^{-\rho} [e^\rho \mu w + \kappa a \Phi_a [i_{\text{pub}} + (1 + s) x]^\alpha] - x [1 - \sigma]}{1 - \sigma} + e^{-\rho} \left[ \frac{(1 - \kappa) \mu^{-1} a \Phi_a [i_{\text{pub}} + (1 + s) x]^\alpha}{1 - \sigma} \right] \right\}. \]

**Proof of Proposition 5.** To shorten notation define:

\[ A \equiv a \Phi, \quad c_0 \equiv e^\rho \mu w + \kappa Ah^\alpha - x, \quad m_1 \equiv (1 - \kappa) \mu^{-1} Ah^\alpha, \quad \delta \equiv \alpha Ah^{\alpha - 1} (1 + s). \]

Optimality requires that either $F < 0$ and $x = 0$, or $F = 0$ and $x > 0$, where

\[ F \equiv \left[ \kappa \delta - 1 \right] [c_0]^{-\sigma} + \Theta [m_1]^{-\sigma} (1 - \kappa) \delta. \]
We first prove part (i). If the credit constraint binds, then \(e^{-\sigma} > \Theta [m_1]^{-\sigma}\). If \(F < 0\), then \(h^L(a, w) = i_{pub}\), and the result is trivial. If \(F = 0\), then \([1 - \kappa \delta] < (1 - \kappa)\delta\), implying that \(\delta > 1\). For the unconstrained case define

\[
\begin{align*}
    c^U_0(a, w) &= \mu e^\rho w + \mu d^U(a, w) - x^U(a), \\
    m_1^U(a, w) &= \mu^{-1} A [h^U(a)]^\alpha - d^U(a, w), \\
    \delta^U(a) &= \alpha A [h^U(a)]^{\alpha - 1}(1 + s).
\end{align*}
\]

Given that \([c^U_0(a, w)]^{-\sigma} = \Theta [m_1^U(a, w)]^{-\sigma}\), the first order condition implies that \(\delta^U(a) \leq 1\). Thus, \(\delta > \delta^U(a)\) and hence \(h^L(a, w) < h^U(a)\). We now prove part (ii). From maximization, we have the condition \(\frac{\partial F}{\partial \alpha} < 0\) and therefore \(\text{sign} \left\{ \frac{\partial h}{\partial A} \right\} = \text{sign} \left\{ \frac{\partial F}{\partial A} \right\}\). The latter derivative is

\[
\frac{\partial F}{\partial A} = \left[ \alpha \kappa Ah^{\alpha - 1} (1 + s) - 1 \right] \left\{ -\sigma [c_0]^{-\sigma - 1} \frac{\partial c_0}{\partial A} \right\} + \alpha \kappa h^{\alpha - 1} (1 + s) [c_0]^{-\sigma} + \Theta [m_1]^{-\sigma} \alpha (1 - \kappa) h^{\alpha - 1} (1 + s) + \alpha (1 - \kappa) Ah^{\alpha - 1} (1 + s) \left\{ -\sigma \Theta [m_1]^{-\sigma - 1} \frac{\partial m_1}{\partial A} \right\}.
\]

First, from the first order condition \([\alpha \kappa Ah^{\alpha - 1} (1 + s) - 1] [c_0]^{-\sigma} = -\Theta [m_1]^{-\sigma} \alpha (1 - \kappa) Ah^{\alpha - 1} (1 + s)\) and then taking \(\Theta [m_1]^{-\sigma} \alpha h^{\alpha - 1} (1 + s) > 0\) as a common factor gives

\[
\frac{\partial F}{\partial A} = \left\{ \Theta [m_1]^{-\sigma} \alpha h^{\alpha - 1} (1 + s) \right\} \left\{ \sigma \kappa \mu \frac{m_1}{c_0} + \kappa \frac{[c_0]^{-\sigma}}{\Theta [m_1]^{-\sigma}} + (1 - \kappa) - \sigma (1 - \kappa) \right\} = \left\{ \Theta [m_1]^{-\sigma} \alpha h^{\alpha - 1} (1 + s) \right\} \left\{ \sigma \kappa \mu \frac{m_1}{c_0} + \kappa \frac{[c_0]^{-\sigma}}{\Theta [m_1]^{-\sigma}} + (1 - \kappa) - \sigma + \sigma \kappa \right\}
\]

where we also have multiplied and divided by \(\mu\) and used the definition of \(m_1\). For constrained individuals, we have that \([c_0]^{-\sigma} \Theta [m_1]^{-\sigma} \geq 1\), and therefore \(\frac{m_1}{c_0} \geq \Theta \frac{1}{\sigma}\). With these inequalities we can find a lower bound to \(\partial F/\partial A\):

\[
\frac{\partial F}{\partial A} \geq \left\{ \Theta [m_1]^{-\sigma} \alpha h^{\alpha - 1} (1 + s) \right\} \left\{ \sigma \kappa \left( \frac{1 - e^{-\rho T} e^\rho}{e^\rho - 1} \right) + 1 - \sigma (1 - \kappa) \right\} = \left\{ \Theta [m_1]^{-\sigma} \alpha h^{\alpha - 1} (1 + s) \right\} \left\{ 1 - \sigma \left[ 1 - \kappa \left( \frac{1 - e^{-\rho T}}{1 - e^{-\rho}} \right) \right] \right\},
\]

where in the first line we have used the expressions for \(\Theta\) and \(\mu\) and the second we have simplified. As claimed in the text, the last expression is positive if \(\kappa \geq [(\sigma - 1)/\sigma] \left[ (1 - e^{-\rho}) / (1 - e^{-\rho T}) \right]\).

\[\blacksquare\]

### C.5 GSL Programs Plus Private Lenders

The threshold level of initial assets \(w_{min}^{G+L}(a)\) above which individuals with ability \(a\) are unconstrained satisfies

\[
d^U \left( a, w_{min}^{G+L}(a) \right) = \kappa \mu^{-1} a \Phi [h^U(a)]^\alpha + \min \left\{ d_{max}, \max \left\{ 0, \left( h^U(a) - i_{pub} \right) / (1 + s) \right\} \right\},
\]

where \(d^U(a, w)\) is given by expression (14) and \(h^U(a)\) by expression (13). Since borrowers combine both sources of credit, \(w_{min}^{G+L}(a) \leq \min \left\{ w_{min}^G(a), w_{min}^L(a) \right\}\), where \(w_{min}^G(a)\) is the threshold under private lending alone and \(w_{min}^L(a)\) is the threshold under the GSL alone.
References


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


