Human Capital and Economic Opportunity:
A Global Working Group

Working Paper Series

Working Paper No. 2011-036

Credit Constraints in Education

Lance Lochner

Alexander Monge-Naranjo

October, 2011

Human Capital and Economic Opportunity Working Group
Economic Research Center
University of Chicago
1126 E. 59th Street
Chicago IL 60637
humcap@uchicago.edu

Credit Constraints in Education*

Lance Lochner
University of Western Ontario

Alexander Monge-Naranjo
Pennsylvania State University

August 16, 2011

Abstract

We review studies of the impact of credit constraints on the accumulation of human capital. Evidence suggests that credit constraints are increasingly important for schooling and other aspects of households' behavior. We highlight the importance of early childhood investments, since their response largely determines the impact of credit constraints on the overall lifetime acquisition of human capital. We also review the intergenerational literature and examine the macroeconomic impacts of credit constraints on social mobility and the income distribution.

A common limitation across all areas of the human capital literature is the imposition of ad hoc constraints on credit. We propose a more careful treatment of the structure of government student loan programs as well as the incentive problems underlying private credit. We show that endogenizing constraints on credit for human capital helps explain observed borrowing, schooling, and default patterns and offers new insights about the design of government policy.

JEL: D14, H52, I22, I23, J24.

Keywords: *Human Capital, Incentive Problems, Government Loans, Early Investments, Social Mobility.*

*Lance Lochner: Department of Economics and CIBC Centre for Human Capital and Productivity, University of Western Ontario, London, Ontario N6A 5C2, Canada; email: llochner@uwo.ca. Alexander Monge-Naranjo: Department of Economics, Pennsylvania State University, University Park, PA 16802-3306; email: aum26@psu.edu.

Contents

1 Introduction

Education and other human capital investments are central to both individual and economy-wide development. By limiting the incentives and capacity to invest in human capital, credit constraints play an important role in determining aggregate productivity, national income distributions, social mobility, and economic growth and development (Becker 1975). This article reviews recent research in both the micro and macro literatures on human capital investment and credit constraints.

Using a simple two-period model, in Section 2 we derive frequently tested implications of constraints for schooling. U.S.-based evidence on the impacts of credit constraints on college-going, as well as consumption and work during college, is reviewed in Section 3. Evidence suggests that the increases in college costs and returns over the last two decades have increasingly pushed more youth up against their credit limits.

Recent U.S. studies suggest that borrowing constraints may be more harmful for investments in young children. We review this evidence in Section 4 and discuss the benefits of considering multi-period investments in human capital. The high estimated degree of complementarity between early and late investments suggests that post-secondary aid policies may come too late to help many youth from disadvantaged families.

Section 5 reviews intergenerational studies in which borrowing constraints determine social mobility and the income distribution. Some of these studies also quantify the impacts of education-based government policies on these outcomes. While recent studies are pessimistic about the benefits of additional subsidies for higher education, new efforts to help finance earlier investments offer more promise.

Ad hoc assumptions about credit constraints constitute a common limitation across all areas of the human capital literature. In Section 6, we propose a more careful treatment of government loan programs and the incentive problems underlying private credit. We show that endogenizing credit constraints for human capital helps explain certain features of the data. We also demonstrate how the modern literature on optimal contracts under limited commitment and private information can help provide new insights about the behavior of human capital investments and the design of government programs.

2 Human Capital with Exogenous Borrowing Constraints

In this section, we use a simple two-period model of human capital investment to examine the key economic trade-offs and empirical relationships studied in the literature on education and borrowing constraints.

2.1 A Basic Model

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

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

where c_t is consumption in periods $t \in \{0, 1\}$, $\beta > 0$ is a discount factor, and $u(\cdot)$ is strictly concave and increasing and satisfies standard Inada conditions.

Each person is endowed with financial assets $W \geq 0$ and ability $a > 0$. Initial assets capture all familial transfers while ability reflects innate factors, early parental investments and other characteristics that shape the returns to investing in schooling. We take (W, a) as given to focus on schooling decisions that individuals make largely on their own; however, central results generalize naturally to an intergenerational environment in which parents endogenously make transfers to their children (see Lochner and Monge-Naranjo 2011b).

During the schooling period, individuals make human capital investments h that increase post-school labor earnings $y = w_1 a f(h)$. Each unit of h entails foregone wages $w_0 \geq 0$ and tuition costs $\tau > 0$; w_1 is the price of human capital and $f(\cdot)$ is positive, strictly increasing and concave. A higher ability a increases total and marginal returns to investment.

Young individuals can borrow d (or save, in which case $d < 0$) at a gross interest rate $R > 1$. Consumption levels in each period are

$$c_0 = W + w_0(1 - h) - \tau h + d, \tag{2}$$

$$c_1 = w_1 a f(h) - Rd. \tag{3}$$

Unrestricted optima. In the absence of credit market frictions, individuals maximize utility (1) subject to (2) and (3). This maximization can be conveniently separated into two problems. First, human capital investment maximizes the present value of net lifetime income, equating its marginal return with that of financial assets:

$$\frac{w_1 a f'\left[h^U(a)\right]}{w_0+\tau} = R. \tag{4}$$

Optimal unrestricted investment $h^U(a)$ is strictly increasing in ability a and independent of initial assets W.

Second, individuals optimally smooth consumption over time. Unconstrained optimal borrowing $d^U(a,w)$ satisfies the Euler equation:

$$u'\left[W + w_0 + d^U(a,W) - (w_0+\tau)h^U(a)\right] = \beta R u'\left[w_1 a f\left[h^U(a)\right] - R d^U(a,W)\right], \tag{5}$$

where $W + w_0$ reflects 'full wealth', i.e. assets plus potential earnings if no time is devoted to schooling. Unconstrained borrowing strictly decreases in wealth and increases in ability.

A higher ability increases borrowing for two different reasons: (i) more able individuals wish to finance a larger investment; and (ii) for any given level of investment, more able individuals earn higher net lifetime income and wish to consume more in the first period. Because of (ii), unrestricted borrowing increases more steeply in ability than does unrestricted expenditure on human capital investment: $\frac{\partial d^U}{\partial a} > \frac{\partial[(w_0+\tau)h^U]}{\partial a}$. All else equal, the more able a person is, the more he wants to borrow relative to his investment.

A Canonical Exogenous Constraints Model. Now, consider a fixed upper limit on the amount of debt that individuals can accumulate:

$$d \leq \bar{d}, \tag{6}$$

where $0 \leq \bar{d} < \infty$. Let λ denote the LaGrange multiplier on this restriction in the utility maximization problem. The first order condition for d becomes:

$$u'(c_0) = \beta R u'(c_1) + \lambda, \tag{7}$$

where $\lambda > 0$ when the constraint binds and $\lambda = 0$ otherwise. The equation $d^U(a,W) = \bar{d}$ defines a threshold level of assets $W_{\min}(a)$ determining who is constrained ($W < W_{\min}(a)$) and who is unconstrained ($W \geq W_{\min}(a)$).

When constraints do not bind, optimal investment and borrowing are given by the unconstrained amounts $d^U(a,w)$ and $h^U(a)$. Otherwise, borrowing is at the limit $\bar{d}$ and optimal investment $h^X(a,W)$ satisfies

$$\frac{w_1 a f'\left[h^X(a,W)\right]}{w_0+\tau} = R+\lambda^*,$$

where $\lambda^* = \frac{\lambda}{\beta u'\left(w_0 a f[h^X(a,W)]-R\bar{d}\right)}$ is positive and decreasing in the borrowing limit, $\bar{d}$.[1]

Constrained persons have high ability relative to their wealth, since $W_{\min}(a)$ is increasing in ability. It is worth noting that being 'unconstrained' may require much higher wealth W than is necessary to cover tuition (i.e. $W+w_0>\tau h$ does not ensure that $d^U(a,W)<\bar{d}$), since individuals also borrow to smooth consumption. When the borrowing constraint binds, all possibilities to bring future resources to the early (investment) period have been exhausted. Then, the optimality condition for human capital investment h^X is

$$(w_0+\tau)u'\left[W+w_0-(w_0+\tau)h^X+\bar{d}\right] = \beta u'\left[\left(w_1 a f\left(h^X\right)-R\bar{d}\right)w_1 a f'\left(h^X\right)\right]. \qquad (8)$$

The implied function $h^X(a,W)$ strikes a balance between increasing lifetime earnings and smoothing consumption, yielding a number of predictions that have been extensively examined in the empirical literature.

Empirical Predictions. Assume constraint (6) binds when referring to $h^X(a,W)$. Then:

1. Constrained individuals under-invest in their human capital: $h^X(a,W)<h^U(a)$.

2. Unconstrained investment $h^U(a)$ is independent of wealth W, while constrained investment $h^X(a,W)$ is strictly increasing in wealth and the borrowing limit $\bar{d}$.

3. The marginal return on human capital $\frac{w_1 a f'[h]}{w_0+\tau}$ is equal to the return on savings R for unconstrained individuals and is strictly greater than R and strictly decreasing in wealth W for constrained individuals.

4. Constrained investment $h^X(a,W)$ responds more negatively to an increase in direct costs, τ, than to an increase in opportunity costs, w_0 (i.e. $-\partial h^X/\partial w_0 < -\partial h^X/\partial\tau$); unconstrained investment responds equally to both costs (i.e. $\partial h^U/\partial w_0 = \partial h^U/\partial\tau$).

[1] This formulation draws a parallel with models that assume individuals face different interest rates, R. A higher $\bar{d}$ or lower W is analogous to a higher R. Assuming an increasing interest rate schedule yields similar predictions to those discussed here.

These results follow from implicit differentiation of equations (4) and (8). The first three are well-known since Becker (1967). They derive from the fact that the marginal cost of investment is higher for constrained individuals, since they cannot borrow to smooth consumption over time. This causes constrained individuals to invest less, stopping school when the marginal return is still relatively high. The fourth implication is derived by Cameron and Taber (2004) in a slightly different setting. Here, it derives from the fact that an increase in opportunity costs also raises 'full wealth' levels, while an increase in direct costs does not.[2] We discuss empirical evidence related to these predictions in Section 3.

Lochner and Monge-Naranjo (2011b) discuss an additional prediction of this model. They show that human capital investment, $h^X(a, W)$, will be *decreasing* in ability a for constrained individuals if the consumption intertemporal elasticity of substitution (CIES), $-u'(c) / [cu''(c)]$, is less than or equal to one (as most estimates in the literature suggest, e.g. see Browning, Hansen, Heckman 1999).[3] This result not only implies perverse cross-sectional investment patterns, but it also implies that an increase in the price of human capital w_1 should lead to aggregate reductions in investment among constrained individuals, since a change in the skill price is analogous to increasing the ability of everyone in the economy. In Section 6.1, we show that important features of government student loan programs and private lending generate a more positive relationship between constrained investment and ability (and w_1).

2.2 Other Margins

Credit constraints are likely to affect other choices. As equation (7) makes clear, early consumption is reduced when borrowing constraints bind. In fact, government student loans link credit to educational expenditures, shifting the impact of borrowing constraints onto

[2]This asymmetry is more easily seen when investment can take only two values, $h \in \{0, 1\}$. In this case, an increase in opportunity costs lowers resources in the no-schooling case when consumption is relatively high, while an increase in tuition reduces resources in the schooling case when consumption is relatively low.

[3]The relationship between ability and constrained investment is driven by two opposing forces. On the one hand, more able individuals earn a higher return on human capital investment, so they would like to invest more. On the other hand, more able individuals have higher lifetime earnings, which increases their desired consumption at all ages. Since constrained borrowers can only increase consumption during the initial period by investing less, the latter effect discourages investment. With strong preferences for intertemporal consumption smoothing (i.e. CIES$\leq$1), the second effect dominates.

consumption rather than schooling investments (see Section 6.1).

The model above abstracts from leisure, so labor supply varies inversely one-to-one with investment. More generally, constrained youth may also substitute leisure for work in order to help alleviate the negative impacts of constraints on consumption and investment. Alternatively, constrained youth may choose to delay college (and its labor market rewards) for a few years to accumulate savings.

Finally, youth may adjust on the school quality margin given any level of attendance. The models above do not explicitly distinguish between school quality and quantity; however, abstracting from opportunity costs (i.e. $w_0 = 0$), one can simply re-interpret h in the model above as the quality of school conditional on school attendance. With this interpretation, constrained youth should attend lower quality institutions, with quality increasing in wealth and the borrowing limit. This implies that wage returns from college attendance should be lower for constrained youth, since they effectively invest less at lower quality schools. As noted by Carneiro and Heckman (2002), this prediction contrasts sharply with the prediction above that the marginal wage return to investment is higher for constrained youth.

2.3 Adding Tastes for Schooling

Much of the empirical literature on college attendance incorporates unobserved heterogenous 'tastes' for education. Augmenting utility (1) to include school taste ξh and restricting human capital investment choices to $h \in \{0, 1\}$ (non-attendance vs. attendance) produces a discrete choice schooling model similar to that of Belley and Lochner (2007). In this environment, individuals choose whether or not to attend college $\max\{U_0(a, W), U_1(a, W) + \xi\}$, where optimal borrowing/consumption would deliver $U_h(a, W) \equiv \max_d \{u(W + w_0(1 - h) - \tau h + d) + \beta u(w_1 a f(h) - Rd)\}$ given schooling choices $h \in \{0, 1\}$, and the individual's ability a and wealth W. Because individuals may enjoy ($\xi > 0$) or dislike school ($\xi < 0$), schooling choices do not necessarily maximize lifetime income. This, along with the discrete nature of schooling, generates some important differences with the model above regarding the relationship between schooling and initial resources.

The observed probability that someone with ability a and wealth W attends college is given by $\Pr[-\xi < \Delta(a, W)]$ where $\Delta(a, W) \equiv U_1(a, W) - U_0(a, W)$. Although the probability of attendance is lower when the borrowing constraint binds for any given (a, W) (analogous to the model above), the probability of attendance is not generally independent of wealth in the absence of borrowing constraints. As discussed in Belley and Lochner (2007), if the net financial return to college is positive and schooling tastes are independent of wealth, then the probability of attending college should be *decreasing* in wealth (conditional on ability) when borrowing constraints are non-binding.[4] Need-based grant aid makes this relationship even more negative. Of course, unobserved tastes for schooling may be positively correlated with W, so this prediction is not particularly powerful on its own. More importantly, in the absence of borrowing constraints, the relationship between wealth W and the probability of attendance (conditional on ability) should become more negative (or less positive) as the net financial returns to college increase, regardless of the underlying relationship between ξ and W.[5] Intuitively, an increase in the net returns to college raises the relative value of college less for individuals with high initial wealth due to diminishing marginal returns to consumption. This need not be true when borrowing constraints limit the consumption of low-wealth individuals. Constrained youth may benefit little from an increase in future labor market returns to school, since additional post-school earnings cannot be used to increase consumption during school when it is most valuable. As discussed below, these results are important for interpreting recent changes in family income – college attendance relationships in light of the contemporaneous increase in returns to college.

3 Evidence on Borrowing Constraints and College

The empirical literature on borrowing constraints and higher education has primarily focused on measuring the population of youth constrained and on the effects of borrowing constraints on education decisions. A few studies also evaluate the impacts of potential constraints on other behaviors at college-going ages (e.g. work in school, consumption allocations). We summarize the recent empirical literature on borrowing constraints and post-secondary ed-

[4]The net financial returns are defined as $N(a) \equiv -\tau + R^{-1}w_1 af(1) - [w_0 + R^{-1}w_1 af(0)]$. When $N(a) < 0$, the probability of attending college is increasing in W.

[5]This result assumes that the density for ξ is relatively flat in the population. Otherwise, if more low-wealth individuals are on the margin of attending, it is possible that $\frac{\partial^2 Pr(-\xi < \Delta(a,W))}{\partial W \partial N} > 0$ even though $\frac{\partial^2 \Delta}{\partial W \partial N} < 0$.

ucation, distinguishing studies by their general approach.

3.1 Family Income/Wealth – Attendance Patterns

Many economists have examined the wide disparities in education by parental income, education, and race to gauge the impact of borrowing constraints on education decisions.

Studies based on the 1979 Cohort of the National Longitudinal Survey of Youth (NLSY79) generally find that family income played little role in college attendance decisions during the early 1980s. Cameron and Heckman (1998, 1999) find that, after controlling for family background, adolescent cognitive achievement, and unobserved heterogeneity, family income had little effect on college enrollment rates for this cohort of youth. Carneiro and Heckman (2002) reach a similar conclusion.

Using data for the late 1990s and early 2000s (1997 Cohort of the NLSY, NLSY97), Belley and Lochner (2007) show that family income has become a much more important determinant of college attendance over time.[6] Youth from high income families in the NLSY97 are 16 percentage points more likely to attend college than are youth from low income families, conditional on adolescent cognitive achievement, family composition, parental age and education, race/ethnicity, and urban/rural residence. This is roughly twice the effect observed in the NLSY79. Belley and Lochner further show that the increased importance of income was primarily focused on lower and middle ability youth.

The NLSY79 do not contain data on wealth; however, the combined effects of family income and wealth in the NLSY97 are substantially greater than the effects of income alone. Comparing youth from the highest family income and wealth quartiles to those from the lowest quartiles yields an estimated difference in college attendance rates of nearly 30 percentage points after controlling for ability and family background. In an attempt to address concerns about the endogeneity of family wealth, Lovenheim (2011) uses data from the Panel Survey of Income Dynamics to estimate the impacts of exogenous changes in housing wealth (driven by local housing booms and busts) on post-secondary enrollment decisions. His estimates suggest that an additional $10,000 in housing equity raises college enrollment by

[6]Ellwood and Kane (2000) argue that college attendance differences by family income were already becoming more important by the early 1990s.

0.7 percentage points, with much larger effects among lower income families. He also finds that the impacts of housing wealth have become more important in the 2000s; however, it is unclear whether this is due to the increased liquidity of housing wealth or a general increase in the effect of family resources on educational attainment.

Belley and Lochner (2007) also use the NLSY79 and NLSY97 to examine the changing role of family income as a determinant of work during college, college delay, and the type of institution attended (two-year vs. four-year).[7] Among lower ability groups, they estimate weak effects of income on weeks worked and hours worked per week in both NLSY cohorts. In contrast, family income becomes a more important determinant of work during school for the most able in the recent cohort. Among the most able NLSY97 college attendees, those from low-income families work more weeks and nearly twice as many hours per week during the school year as those from high-income families. While the growing effects of income on attendance are largely focused on lower ability groups, the growing effects of income on work are concentrated among the most able. Interestingly, Belley and Lochner estimate weak effects of family income on college delay in both NLSY cohorts.

The relationship between family income and the type of post-secondary institution individuals attend has changed since the early 1980s. While family income had little effect on the choice of two-year vs. four-year institutions in the NLSY79, students from the highest income quartile in the NLSY97 are 11 percentage points more likely to be attending a four-year institution than their counterparts from the bottom quartile (Belley and Lochner 2007).[8] By contrast, the relationship between family income and attendance at selective high quality institutions appears to have weakened over this same period. Kinsler and Pavan (2010) estimate that moving from the bottom to top income quartile increased the probability of attending a top quality college by about 25 percentage points in the NLSY79 and by only 16 percentage points in the NLSY97. Among top (often private) schools, the sharp increases in tuition since the early 1980s were generally accompanied by increases in financial aid for

[7]Their estimated effects of income on college delay and institution type for the NLSY79 are consistent with those of Carneiro and Heckman (2002), who also examine these margins.

[8]Lovenheim and Reynolds (2011) also use the two NLSY cohorts to explore more detailed trends in college enrollment by institution type. They estimate an important shift in enrollment from four-year to two-year schools among men from high income/low ability and low income/high ability backgrounds. Among women, college enrollment increases were largely focused in four-year institutions.

lower income students. This effectively increased the price of college quality more for high-income students relative to their lower-income counterparts. As such, it is unclear whether these findings signal that low-income youth wishing (and able) to attend selective colleges are less constrained now than in the early 1980s or whether these findings simply reflect changes in relative prices.

As emphasized by Carneiro and Heckman (2002), adolescent cognitive achievement has much stronger effects on college-going than does family income. This is true in both NLSY samples. Still, the fact that family income has become so much more important for attendance in recent years suggests that credit constraints may have become more salient for many American youth. As Belley and Lochner (2007) and Lochner and Monge-Naranjo (2011b) discuss, both college costs and returns have risen substantially since the early 1980s. These forces should have led to an increase in demand for credit among students; however, real government student loan limits changed very little. The fraction of undergraduate borrowers 'maxing out' their federal Stafford loans nearly tripled over the 1990s to 52% (Berkner 2000 and Titus 2002).

Many factors shape the relationship between family income and schooling besides borrowing constraints. Need-based financial aid is an important feature of American higher education. In fact, tuition net of non-repayable aid is generally negative for very low-income American youth attending in-state public universities (Belley, Frenette and Lochner 2011). By lowering the net price of college for low-income youth relative to high-income youth, need-based grants and scholarships tend to reduce income – attendance gradients through price effects alone. Of course, the effects of aid depend critically on the extent to which low-income youth are aware of available aid. This awareness, as well as tastes for schooling more generally, may depend on social networks and peers. This would tend to amplify any relationship between family income and schooling through social multiplier effects.

Finally, it is sometimes argued that higher income families place greater value on education and that this may explain the positive relationship between family income and schooling. If true, it is not clear why this relationship should have strengthened so much since the early 1980s.[9] As the model of Section 2.3 shows, the well-documented increase in net returns to

[9]Nor, is it obvious why the income – attendance relationship should be so much stronger in the U.S.

schooling since the 1980s should have *weakened* the income – attendance relationship in the absence of borrowing constraints if the relationship between 'tastes' for college and family income had remained stable.

3.2 Differential Marginal Returns to Schooling

As Card (1999) notes, many instrumental variables (IV) estimates of the wage return to schooling exceed ordinary least squares (OLS) estimates by 20-30%. Lang (1993) and Card (1995, 1999) have conjectured that borrowing constraints may explain this finding. They point out that many IV studies use various institutional details affecting the marginal cost of schooling (often college). If these instruments largely impact the decisions of low-income and constrained youth, then these IV estimates reflect the relatively high marginal return to schooling for those that are constrained (based on the 'local average treatment effect' interpretation of IV).[10] On the other hand, OLS estimates may more closely reflect average returns in the population, which may be lower.

Carneiro and Heckman (2002) raise concerns with this interpretation. First, many of the instruments used are either 'weak' or correlated with (typically unobserved) cognitive ability. Second, IV estimates may exceed OLS estimates even in the absence of borrowing constraints due to heterogeneity in returns to schooling and self-selection into different schooling levels. Third, once school quality differences are considered, borrowing constraints may lead to lower returns per year of schooling. Moreover, the marginal cost of schooling may differ for reasons other than borrowing constraints, e.g. heterogeneity in tastes for schooling. High IV estimates based on tuition variation or college proximity may simply reflect the fact that these instruments affect attendance decisions most for youth who dislike school.

Cameron and Taber (2004) push further on this issue. Based on the prediction that investment responds more to direct costs than opportunity costs when individuals are borrowing constrained (result 4 in Section 2.1), they estimate the returns to schooling using separate instruments related to each of these costs. They argue that the set of individu-

than it is in Canada. Despite greater targeting of aid to low-income families in the U.S. relative to Canada, post-secondary attendance gaps by family income are nearly twice as large in the U.S. compared to Canada (Belley, Frenette and Lochner 2011).

[10] See Imbens and Angrist (1994), Heckman and Vytlacil (1998), or Card (1999) for more detailed discussions of the interpretation of IV estimators in this context.

als whose college-going is affected by a change in direct costs (measured by whether there is a college in the individual's county of residence) should disproportionately include more credit constrained youth than the set of individuals affected by a change in opportunity costs (measured by local low-skill wage rates). Ignoring differences in college quality, if borrowing constraints are important for college-going, then using 'college in county' to instrument for schooling should yield a larger estimate for the returns to schooling than using local low-skill wage rates. Based on men from the NLSY79, they find the opposite, leading them to conclude that borrowing constraints are not important for college-going. By addressing issues related to weak and invalid instruments and by explicitly comparing different IV estimates, their results overcome the first two criticisms raised by Carneiro and Heckman (2002). However, college quality differences can make it difficult to draw strong conclusions about the significance of borrowing constraints.

3.3 Structural Models

A few studies estimate lifecycle schooling models that exploit data on schooling choices, earnings, and in some cases, assets and family transfers, to identify the role of borrowing constraints. By estimating preferences, human capital production technology, and other important factors determining educational choices, this approach enables researchers to evaluate a wide range of potential policies. We discuss three important papers in this literature.

Cameron and Taber (2004) estimate a lifecycle model with a discrete set of schooling options to test different discount rates in schooling choices. Evidence that some individuals face high interest rates relative to others would imply that borrowing constraints distort their education decisions. Using data on men from the NLSY79, Cameron and Taber estimate discount rates that are consistent with optimal schooling choices given observed schooling costs and earnings functions for different types of individuals. The main source of identification for differences in discount rates is the differential roles played by opportunity costs and direct costs as discussed above. Consistent with their IV analysis, they find no evidence of discount rate heterogeneity in their sample.

Keane and Wolpin (2001) estimate a dynamic model of schooling, work, and consumption behavior to explore the importance of borrowing constraints for all of these choices. Their framework incorporates (exogenously determined) parental transfers, which depend on both parental education as well as an individual's own schooling enrolment choices. They use panel data on schooling and work (full-time and part-time), wages, and assets for white males in the NLSY79. Importantly, Keane and Wolpin allow for unobserved heterogeneity in the ability to acquire human capital, tastes for work and school, and borrowing limits.

Estimated borrowing limits are very tight (ranging from $600 to $1000 across individuals, in 1987 dollars) — less than one-third the estimated cost of a single semester of school (about $3,700). Not surprisingly, then, simulations suggest an important role for parental transfers and part-time work in enabling school attendance. The estimates suggest that parents provide between $3,300 and $10,000 in transfers while enrolled in school, where transfers are increasing in parental education. Because transfers are estimated to be substantially lower when students are not enrolled in school, a sizeable portion of parental transfers effectively acts as a subsidy for education — a subsidy that is much larger for children with more educated parents. Based on a series of simulations, Keane and Wolpin conclude that nearly all of the (sizeable) differences in educational attainment by parental education are accounted for by higher enrollment-contingent parental transfers and unobserved heterogeneity. Although they estimate tight (often binding) borrowing limits, increases in available credit have negligible effects on schooling. Instead, increasing loan limits tends to reduce work and increase consumption during school.

The model of Section 2.1 is useful for interpreting these results and understanding identification. Ideally, one would identify who is constrained by their consumption profiles; however, Keane and Wolpin have no data on consumption nor do their data allow them to directly infer consumption during school for most youth.[11] As noted earlier, schooling patterns in the NLSY79 are largely consistent with unconstrained investment; however, the low levels of debt taken on by most youth suggest that borrowing limits are quite low. Since debt levels appear to vary little with ability, many youth must be borrowing constrained; otherwise, (unconstrained) borrowing should increase sharply with ability as shown in Section 2.1.[12] If

[11]Asset measures are not generally available during most college-going years and there are no measures of schooling costs or parental transfers. It is, however, possible to directly infer consumption from reported income levels and changes in assets for older individuals that are no longer in school.

[12]Unfortunately, Keane and Wolpin (2001) do not report distributions of debt by ability type; however,

many youth are in fact constrained, then the CIES would need to be greater than one in order to generate a positive relationship between ability and schooling. Their high estimated CIES of 2 implies that distorted consumption profiles are not particularly costly in utility terms. Thus, borrowing constraints will have weak effects on schooling, and heterogeneous income-contingent transfers are needed to explain the positive relationship between parental education and schooling (conditional on ability type).

Using data on recent male high school graduates in the NLSY97, Johnson (2010) estimates a similar decision model with a few important differences. He explicitly models government student loan programs as well as a private credit limit, allows for differences in tuition across states, incorporates need- and merit-based grants, and allows for exogenous unemployment. Most importantly, he exploits additional data on average tuition by state and data on reported grant aid and parental transfers in the NLSY97.[13] This better enables him to infer consumption during and after school, which helps in identifying who may or may not be constrained. His data allow him to directly estimate parental transfer functions and student aid by parental income, while Keane and Wolpin (2001) have to infer parental transfers indirectly from schooling and work choices (and asset levels in later years).

Some of Johnson's main findings are similar to those of Keane and Wolpin (2001): parental transfers (especially the fact that schooling-contingent transfers are greater for higher-income families) and unobserved heterogeneity are important determinants of schooling. Johnson also estimates modest borrowing limits relative to college costs. However, his estimated credit limits are substantially greater than those of Keane and Wolpin (2001).[14] Despite greater borrowing opportunities, Johnson estimates a stronger, though modest, impact of increasing loan limits. Simulations suggest that an additional $1,500 in credit per year in school would raise BA completion rates by 2.3%. Allowing students to borrow up to the total costs of schooling would increase completion rates by nearly 4%. Given the low cost

estimated borrowing limits are similar across types suggesting little variation in debt along that dimension.

[13]Like Keane and Wolpin (2001), he also uses data on schooling, work, assets, and wages. Since many of his respondents are still quite young, Johnson (2010) uses wages at ages 25+ from the NLSY79 cohort in estimation. This effectively yields estimates that average the returns to schooling and experience across the two NLSY cohorts.

[14]Youth attending college for four-years can borrow up to $23,000 from the Stafford Loan Program plus as much as an estimated $7,000 in private loans for some types. Average annual tuition for Johnson's sample is about $15,000. All figures in 2004 dollars.

of extending government student loan programs, Johnson (2010) estimates that increasing loan limits would have a greater impact on college outcomes than an increase in education subsidies costing the same amount. However, Johnson (2010) argues that subsidies are necessary to generate large increases in college completion.

Borrowing constraints have small to modest impacts on schooling choices in these two studies for very different reasons. As discussed above, estimates from Keane and Wolpin (2001) suggest that most students are constrained but that consumption and leisure are distorted rather than schooling. That schooling is unaffected by borrowing constraints is not surprising given other evidence based on the NLSY79. It is more surprising that Johnson (2010) estimates that increasing borrowing limits would have only modest effects on college-going given the increased importance of family income in the NLSY97. Despite the fact that credit opportunities plus parental transfers allow for, at best, modest consumption during school, Johnson estimates that few youth borrow up to their limit. This is almost certainly due to risk aversion and the possibility of very low income associated with post-school unemployment in his model.[15] Because of this, his estimates suggest that few individuals are willing to take on much debt. Indeed, his estimates suggest that very few choose to borrow more than $10,000 (compared to nearly 20% of men in his data).[16]

Assumptions about minimal income (or consumption) levels are crucial for the importance of borrowing limits in dynamic schooling models with uncertainty such as Johnson's. The demand for credit can be much higher with explicit insurance mechanisms or implicit ones such as bankruptcy, default, or other options (e.g. deferment and forgiveness in government student loans). Despite their importance, the empirical literature has generally given little attention to risk and insurance, issues we discuss further in Section 6.

Much of the relationship between socioeconomic background (parental education or income) and college-going in Keane and Wolpin (2001) and Johnson (2010) is driven by

[15] If $\lim_{c \to 0} u'(c) = \infty$, individuals that must honor all debts would never choose to borrow more than the minimum value of the present discounted lifetime income, i.e. the 'natural' borrowing limit of Aiyagari (1994). Given Johnson's assumptions on unemployment income, the 'natural' borrowing limit in his model would be around $17,000 at college-going ages. However, Johnson (2010) does not fully solve the model through the end of life, instead assuming a terminal value function at age 40 that depends positively on remaining assets and skill levels at that age. As such, there is no actual 'natural' limit in his framework.

[16] In related work, Navarro (2010) estimates that simultaneously removing both uncertainty and borrowing constraints would substantially increase college attendance.

schooling-dependent parental transfers: more advantaged parents provide greater schooling 'subsidies' to their children. This raises the question: why do wealthier parents effectively subsidize so much schooling if their children are not willing to pay for it themselves? (Their estimates suggest that many children would not attend without parental 'subsidies' even if credit were abundant.) Taken at face value, these results suggest that many parents must value their children's education more than their children do. This gives rise to three potential explanations for the strong positive relationship between parental income/education and schooling-contingent subsidies: (i) All parents may value schooling the same, but poor parents may be constrained in what they can afford to pass on to their children. (ii) All parents may value schooling the same, but wealthier parents prefer to buy more of it like they do other consumption goods. (iii) Wealthier parents may value schooling more than poor parents. Ironically, these explanations mirror the earlier discussion of the wealth – schooling relationship, only for parents rather than potential students.[17]

While the results of Keane and Wolpin (2001) and Johnson (2010) suggest that expansions in student loan programs are likely to have limited effects on college-going (an important result by itself), they effectively shift the 'constrained' question up a generation. As such, it is not clear how these results help explain the dramatic increase in family income – attendance gaps over the past few decades. Efforts to endogeneize parental transfer decisions in these models are needed to make more progress on this question.

Adolescent 'endowments' or abilities also play a central role in determining the relationship between socioeconomic background and education (and earnings) outcomes in both Keane and Wolpin (2001) and Johnson (2010). This is also true in studies explicitly analyzing education gaps by family income (e.g. Cameron and Heckman 1998, Carneiro and Heckman 2002, Belley and Lochner 2007). Yet, these 'endowments' are typically treated as exogenous and invariant to policy. Recent work discussed in Section 4 endogenizes these endowments through early investments by families and schools.

[17]Explanation (i) is consistent with the findings of Brown, Scholz and Seshadri (2011) and Caucutt and Lochner (2011). Explanations (ii) and (iii) are difficult to reconcile with the strong increase in the returns to college for reasons discussed earlier: unless the relationship between parental income and parental tastes for schooling strengthened over time, the increased returns should have weakened the link between parental income and schooling-contingent transfers.

3.4 Other Approaches to Identifying Constraints

Stinebrickner and Stinebrickner (2008) take a novel approach to measuring borrowing constraints by directly asking students enrolled at Berea College in Kentucky whether they would like to borrow more if they could (at a 'fair' interest rate). They, therefore, measure the share of enrolled students that are constrained and the impact of being constrained on college dropout rates. It is important to note that the typical student at Berea College comes from a low-income family; however, the college is unique in that it effectively charges zero tuition and offers large room and board subsidies. Despite these unique institutional features, college dropout rates are quite similar to those for low-income students in the U.S. as a whole. While Stinebrickner and Stinebrickner (2008) find that many Berea students live on a very tight budget, only about one-in-five reports that they would like to borrow more if they could. Interestingly, they further estimate that college drop out rates (by the beginning of year two) are about 13 percentage points higher (or roughly double) for those youth deemed to be 'constrained' relative to those that are not. Adjusting for other potential factors reduces this difference to about 11 percentage points. While factors other than borrowing constraints explain more than 85% of the dropout rate at Berea College, the inability to borrow appears as an important determinant for those that are constrained.

Brown, Scholz and Seshadri (2011) explicitly model intergenerational relationships and derive a new way of identifying which youth may be affected by borrowing constraints. Their model assumes that youth would be borrowing constrained if they did not receive help from their parents. Parents are assumed to be able to borrow freely, but they cannot write enforceable loan contracts with their children. While parents may make transfers to their children due to altruism, they may not want to transfer enough resources to satisfy the child's full demand for consumption and schooling at college ages. In this case, parents would provide all transfers to their children when they are college-age, but children would still invest less than the unconstrained optimal amount. By contrast, unconstrained families will transfer enough resources to their children when they are young and will continue to make transfers after their children leave school. These results suggest that one can distinguish between 'constrained' youth and 'unconstrained' youth based on the presence of post-school parental transfers. Brown, Scholz and Seshadri further show that in their framework, total human capital investment should be more sensitive to a tuition subsidy among constrained youth than among unconstrained youth.[18]

Based on these insights, Brown, Scholz and Seshadri (2011) use intergenerational data on educational attainment and family transfers from the Health and Retirement Survey (HRS) to estimate the effects of borrowing constraints on schooling in the U.S. during the 1970s, 1980s, and 1990s. Identifying 'constrained' youth as those receiving no post-school family transfers, they find that roughly 50% of all youth in their sample are 'borrowing constrained'. Because the HRS do not contain information on educational subsidies/aid, they use sibling spacing as an instrument for student aid. Families with multiple children in college at the same time generally qualify for more aid than families with children attending at different times. Their estimates suggest that among 'constrained' youth, an additional $3,600 in aid (associated with having a twin in college at the same time vs. no sibling overlap in college) increases average schooling levels by 0.2 years. They estimate negligible effects of additional aid on those youth who are 'unconstrained' by their measure.

3.5 Summarizing the Evidence

For the most part, there is general agreement regarding the extent to which borrowing constraints affect college decisions. Most studies analyzing the NLSY79 data find little evidence that borrowing constraints affected college-going in the early 1980s. However, the evidence suggests that constraints have become more salient in recent years: the rising costs of and returns to college, coupled with stable real government student loan limits, are the likely cause for stronger family income – college attendance gradients among recent cohorts.

Borrowing constraints affect more than college attendance. For example, they can affect the quality of school attended. Family income has become a more important determinant of attendance at four-year (relative to two-year) schools, while it has become less important

[18] As Carneiro and Heckman (2002) discuss, this result does not necessarily generalize to other models of schooling choice. It is particularly difficult to derive strong predictions on 'tuition sensitivity' in models with discrete schooling choices. For example, consider the college attendance choice. Even if tuition has a large effect on the value of college for constrained individuals, it is possible that very few constrained youth are near the margin of indifference. By contrast, even small changes in the value of college among unconstrained youth may cause many to change their attendance decision if they are all largely indifferent.

for attendance at very selective institutions. Borrowing constraints could also delay college attendance, but the evidence suggests little impact on this margin. Instead, constrained students appear to work more while in school than those that are unconstrained. In recent years, this distortion appears to have become more important for higher ability youth. Lastly, there is widespread agreement that consumption is quite low for constrained youth enrolled in college.

4 Early Investments in Children

Despite evidence that adolescent skill levels are important in determining subsequent schooling and lifetime earnings (see, e.g., Cameron and Heckman 1998, Keane and Wolpin 1997, 2001, and Carneiro and Heckman 2002), only recently has the literature begun to examine the impacts of borrowing constraints on early investments in young children.

Indirect evidence suggests that constraints at early ages play a more important role in determining human capital investment than constraints at later ages. First, most empirical studies find high lifetime returns for early childhood programs, especially for the most disadvantaged children (e.g., see Karoly et al. 1998, Blau and Currie 2006, or Cunha, et al. 2006). Second, empirical studies find that family income received at early childhood ages has a greater impact on achievement and educational attainment when compared with income received at later ages (e.g. Duncan and Brooks-Gunn 1997, Duncan, et al. 1998, Levy and Duncan 1999, Caucutt and Lochner 2006, 2011). More generally, recent studies show that exogenous increases in family income lead to improvements in early child development (e.g. Løken 2010, Løken, Mogstad and Wiswall 2010, Duncan, Morris and Rodrigues 2011, Milligan and Stabile 2011, and Dahl and Lochner, forthcoming).

Credit constraints are natural candidates to explain why most low-income children do not participate in quality preschool programs despite the high economic returns. First, while (generous) government student loan programs are available for college in the U.S. and other developed countries, neither governments nor private lenders typically offer loans to parents to help finance human capital investments in younger children. Second, even though elementary and secondary education is publicly provided, the quality of public schools available to poor American families is often low, while high quality private schools and preschool programs are typically quite expensive. Parental time is also an important input in a young child's education that poor parents may be unable to afford. Finally, most parents of young children are young themselves, in the early stages of their labor market careers and without a solid credit history. Even young college-educated parents may be constrained by mortgages, their own schooling loans, and other liabilities.

To better understand the role of borrowing constraints at early and late childhood ages, it is useful to generalize the human capital production function in Section 2 to include multiple periods of investment. For simplicity, suppose human capital upon labor market entry H depends on early childhood investment h_1, adolescent investment h_2, and ability a:

$$H = af(h_1, h_2). \tag{9}$$

As discussed in Cunha, et al. (2006) and Cunha and Heckman (2007), the dynamic complementarity between early and late investments (as measured by $\frac{\partial^2 f}{\partial h_1 \partial h_2}$ or the elasticity of substitution) is crucial for the accumulation of human capital over the lifecycle. With strong complementarity, it is difficult to compensate for a lack of early investment at later ages. In this case, inadequate early investments lead to low returns in later investments, consistent with Keane and Wolpin (2001) and Cameron and Heckman (1998) who find small returns to higher education for those with low adolescent skills. Caucutt and Lochner (2011) also show that when h_1 and h_2 are sufficiently substitutable, relaxing the borrowing constraint in either investment period increases investment in that period but reduces investment in the other; when h_1 and h_2 are strong complements, relaxing borrowing limits in either period increases investment in both periods.

The estimates of Cunha, Heckman, and Schennach (2010) suggest that investments are quite complementary over time, with the degree of complementarity growing with age for cognitive skills.[19] They find that it is optimal to invest relatively more in young children with investment declining with age. This is particularly true for children with low initial endowments. This optimal path of declining investment contrasts sharply with the typical

[19] Cunha, Heckman, and Schennach (2010) estimate elasticities of substitution ranging from 0.4 to 1.5. They use data from the Children of the NLSY and exploit a dynamic non-linear factor structure and multiple measurements for cognitive and non-cognitive skills and family investments.

pattern of increasing parental earnings over the lifecycle. To the extent that borrowing constraints limit early investments in some children, those early deficits are likely to be compounded over time.

Caucutt and Lochner develop and calibrate a dynastic overlapping-generations model to quantitatively assess the importance of borrowing constraints and policy interventions over different stages ('early' and 'late') of child development.[20] They consider four levels of 'late' human capital investment, corresponding to high school dropout, high school completion, some college, and college completion, calibrating their model using intergenerational micro data from the Children of the NLSY and national data on schooling expenditures and revenues for the U.S. Despite their very different approach, they identify a similar degree of complementarity between early and late human capital investment to that estimated by Cunha, Heckman and Schennach (2010).

Caucutt and Lochner find that many young and old parents are borrowing constrained, especially those with higher education who took out loans to finance their own education and who tend to have high ability children. However, like Keane and Wolpin (2001) and Johnson (2010), their model suggests that there would be little impact on human capital investment ('early' or 'late') from relaxing borrowing constraints on college-age youth or their parents. At least in the short-run, relaxing constraints on young parents would substantially increase both 'early' investments in young children and 'late' investments in older children (e.g. high school completion and college). For example, they find that a modest increase in the borrowing limit faced by young parents would increase early investment by about 8% and college graduation rates by 7%. Interestingly, the effects are greater for families with more educated parents, since these families are constrained and want more credit for investment in their children. Less educated parents appear to want more credit primarily for current consumption.

Caucutt and Lochner also explore the long-run impacts of permanently relaxing borrowing constraints, allowing the distribution of assets and human capital to change in response. Here, the results are quite different. Since relaxing the borrowing constraint for young par-

[20]Restuccia and Urrutia (2004) also calibrate a dynastic model of human capital production with early and late investments; however, they abstract from borrowing and saving. We discuss this work in Section 5.

ents causes families to accumulate more debt over time, future generations find themselves constrained to nearly the same extent that initial generations were before the constraint was relaxed. On average, this shift in assets results in negligible long-run effects of relaxing the constraint on average human capital levels. Modest increases can be a double-edged sword, increasing human capital in the short-run but lowering family assets in the long-run.[21]

Finally, Caucutt and Lochner obtain two interesting results on the impact of subsidies for 'early' vs. 'late' investment. First, subsidies for investment at either stage raise investments at both stages, calling into question traditional analyses of policies targeting college-age that ignore the response of early investment; this omission would cause one to under-estimate the final impact on human capital levels by about 75%. Second, they find that subsidies for early investment produce much greater short- and long-run gains in human capital than (fiscally equivalent) subsidies for late investment. Dynamic complementarity implies that families that are constrained when their children are young do not fully capitalize on subsidies at later ages, because it is too costly to adjust early investments. Those that receive inadequate early investments do not find it worthwhile to make additional later investments (especially college) even if it is heavily subsidized. By contrast, early investment subsidies enable families to increase investments in their young children without sacrificing current consumption or borrowing more. Those investments can then be matched with later investments when constraints are less severe.

5 Macroeconomic Perspectives

By shaping the capacity and the incentives of individuals to invest in their own human capital and that of their children, credit constraints become a major determinant of social (intergenerational) mobility and of the overall distribution of skills, income and wealth in a country. Unlike the micro empirical literature discussed in Section 3, the macro literature has largely considered 'intergenerational' constraints arising from the inability of parents to borrow against their children's future earnings.

[21] Of course, welfare of the dynasty is improved by relaxing the constraint; however, initial generations capture most of this gain.

5.1 Inequality and Persistence of Skills and Income

Becker and Tomes (1979, 1986) and Loury (1981) pioneered the development of fully consistent economic models of the income distribution based on intergenerational transfers and investments in human capital. These papers show how an economy's income distribution is endogenously determined by preferences and market opportunities. Credit constraints play a central role in that they limit the capacity of poor parents to invest in their children.

In these models, human capital for generation t, H_t, depends on the investments, h_t, and ability, a_t, for that generation: $H_t = f(a_t, h_t)$. It may also depend on shocks to the production of human capital as well as the human capital of one's parents H_{t-1}. Ability is typically assumed to follow a first order Markov process across generations, and earnings generally depend on human capital levels, independent idiosyncratic market shocks, and the economy-wide price of human capital.

Three types of intergenerational preferences are commonly assumed in this literature. 'Altruistic' preferences assume that parents directly value the welfare of their children. Parents of generation t value their own consumption and the utility of their future generations according to the recursive value function $V_t = u(c_t) + \beta E_t(V_{t+1})$, where $u(c_t)$ is the utility of consumption of the family at time t, $\beta > 0$ is an altruism parameter reflecting the weight placed on children, and $E_t(V_{t+1})$ is the expected utility for children in the next period. A closely related form of 'impure' altruism may arise when parents live for two or more periods: $V_t = u(c_t) + \beta[u(c_{t+1}) + \theta E_t(V_{t+1})]$ with $\beta > 0$ reflecting time discounting and $\theta \geq 0$ the degree of intergenerational altruism. Altruism is said to be 'impure' when $\theta < 1$. Notice that V_t are value functions that must satisfy a recursive relationship, a fixed point problem defined by a Bellman equation. This is not the case for the other two form of preference assumed in the literature. 'Paternalistic' preferences assume that parents directly value human capital investments, human capital outcomes, or even earnings, e.g. $V_t = u(c_t) + E_t U(h_{t+1})$. 'Warm-glow' preferences assume that parents directly value any transfers/bequests b_t made to their children: $V_t = u(c_t) + U(b_t)$. In this case, it does not matter what children do with the money.

Early theoretical papers by Becker and Tomes (1979) and Loury (1981) assume parents invest in their children's human capital while ruling out intergenerational financial transfers. Becker and Tomes (1979) assume 'paternalistic' preferences over child income, while Loury assumes a very general form of altruism. These papers derive useful conditions that ensure that the economy converges to a unique invariant income (and skill) distribution that depends on preferences and technology. Importantly, the economy is 'ergodic' in the sense that the impact of the initial conditions for a dynasty progressively washes out with the passing of time. After many generations, the income levels of all dynasties tend to exactly the same distribution as that for the economy as a whole. There tends to be 'regression to the mean' in the sense that richer (poorer) than average parents tend to have richer (poorer) than average children, but the gaps tend to close over time. Becker and Tomes (1979) show how social mobility is driven by intergenerational persistence in ability, the variance of labor market shocks, and the extent to which parents value the income of their children. In Loury (1981) a positive intergenerational persistence in income arises even when ability is not correlated across generations.

Becker and Tomes (1986) extend their earlier analysis to incorporate complementarity between ability and investment in the production of human capital, non-negative parent-to-child financial transfers, and altruistic preferences. They show that constrained families will not leave financial bequests; instead, all their transfers will be in the form of human capital investment. Moreover, they suggest that the relationship between ability and investment might be negative for constrained families. Beyond these lessons, the aggregate implications of models like this can only be examined numerically.

A few more recent studies consider the implications of *indivisibilities* in human capital investment (e.g. Galor and Zeira 1993, Caucutt and Kumar 2003, Restuccia and Urrutia 2004). Galor and Zeira (1993) study the determination of the income distribution when human capital investments consist of a single discrete choice: whether or not to attend college. They assume warm-glow preferences and abstract from individual shocks to earnings, abilities, and human capital production. In the presence of credit market imperfections (modelled as a positive gap between borrowing and lending interest rates), Galor and Zeira show that non-convexities can greatly affect the long-run determination of the income distribution. In particular, they show that there can be multiple steady states (potentially infinitely many) that fall into three categories: (i) global poverty traps (the entire population is unschooled), (ii) a perfect caste system with "individual poverty traps" where some dynasties are forever unschooled while the others are forever schooled; or (iii) a fully developed country/skilled population equilibrium if the initial capital is high enough and everyone is above some threshold. Which long-run steady states arise depends entirely on the initial distribution of skills, a sharp contrast with the ergodicity in Loury (1981) and Becker and Tomes (1979).

The non-convexity in schooling choices also produces very different predictions for social mobility (compared to the earlier studies). Galor and Zeira (1993) show that, in the limit, their model produces zero social mobility (i.e. infinite persistence in family status) in all steady state types. The lack of ergodicity is likely to hold even with ability shocks, as long as abilities are always high enough so that rich individuals always find it worthwhile to invest in college. One way to induce ergodicity is to introduce large (and uninsured) post-investment income shocks that consistently move dynasties away from the attraction of 'unschooled' and 'schooled' resting points. If so, unschooled rich (impoverished poor) parents may (not) transfer enough resources for the child to go to school.

Assumptions about the form of intergenerational preferences also have important consequences. For example, taxes that transfer resources from rich to poor dynasties reduce average human capital levels when parents are paternalistic as in Becker and Tomes (1979) while they increase human capital with warm-glow preferences as in Galor and Zeira (1993). Implications of impure altruism are studied by Aiyagari, Greenwood and Seshadri (2002). Their numerical exercises suggest important general equilibrium interactions between impure altruism and credit constraints with incomplete insurance markets. Most notably, a higher stock of physical capital can lead to higher steady state levels of human capital when parents are less than perfectly altruistic towards their children.

In addition to Aiyagari, Greenwood and Seshadri (2002), Caucutt and Kumar (2003) provide an important dynastic framework for quantitatively studying the formation of human capital and the evolution of earnings across generations. They assume altruistic preferences, a form of persistence in schooling, and lumpy human capital investments with uncertain payoffs (i.e. students may fail to complete school). By ruling out financial transfers, they impose intergenerational credit constraints. As in most of this literature, Caucutt and Kumar assume that families cannot insure against the different risks they face, including the possibility of school failure (which depends on ability) and uncertainty in the ability levels of grandchildren and subsequent generations.

Caucutt and Kumar (2003) find similar types of steady states to those of Galor and Zeira (1993). Global or individual poverty traps arise (even with altruistic preferences) due to indivisibilities in human capital investment. With their preferred calibration, their relatively simple model captures the share of college educated workers, the college wage premium, and the enrollment and dropout rates of children conditional on parental education as observed in the U.S.[22] To fit intergenerational schooling relationships in the data, it is important that they allow parental education to directly enter the production of children's human capital (i.e. graduation probabilities are higher for those whose parents completed college).

Restuccia and Urrutia (2004) extend the dynastic framework of Caucutt and Kumar (2003) to include a period of early investment in children along with a college attendance decision at later ages. Early investments (and innate abilities) are assumed to increase earnings associated with college attendance as well as the likelihood of finishing college. Borrowing and saving, as well as intergenerational financial transfers, are ruled out. Families must consume and make investments out of current incomes. Calibrating their model to U.S. data, they argue that differences in early investments by parental income are largely responsible for observed levels of intergenerational persistence, since the lack of credit availability is particularly problematic for poor young parents (for reasons discussed in Section 4).

5.2 Government Policies

In general, human capital investments are more efficient if done earlier in life. Therefore, if credit constraints limit the ability of younger generations to invest, aggregate output will suffer. Under these circumstances, private market allocations are inefficient, and government programs that transfer resources from older to younger generations could increase overall welfare. Unfortunately, such transfer schemes may not be politically implementable, because they entail a net loss for older generations.

[22] It is worth noting, however, that Caucutt and Kumar (2003) present numerical examples in which changes in the value of the coefficient of relative risk aversion or in the cost of college lead to qualitatively different sets of steady states.

As discussed by Boldrin and Montes (2004), intergenerational conflict can be averted if public schooling policies are tied to other government transfers. Specifically, (mid-aged) workers could be taxed to help finance the human capital of youth and the pensions of retirees. This simple point of Boldrin and Montes' — linking government human capital policies with other transfer programs to improve upon allocations — seems robust to much richer environments than their three-period deterministic overlapping-generations model with representative generations and exogenous credit constraints. A secondary result of their analysis, namely the ability of the government to fully restore efficiency with simple (homogeneous) policies, is much more fragile. Still, it provides a useful reference point for discussing three key limitations in the design of government policies: (i) heterogeneity; (ii) endogeneity of private credit constraints; and (iii) risks and incentive problems.

Heteregeneity in abilities and family resources can be a major limitation for the efficacy of government programs. Merit-based programs may be imprecise in differentiating by ability, especially at younger ages when investments may have high returns and credit constraints may be most severe. Need-based programs may be more precisely targeted, but they may lead to inefficient over-investment by lower ability individuals. The implications of these and other forms of government policies are studied by Caucutt and Kumar (2003), Restuccia and Urrutia and (2004) and Hanushek, Leung and Yilmaz (2003).

Caucutt and Kumar (2003) explore a number of policies in the form of subsidies/earnings taxes: (i) need-based policies that seek equality of opportunity; (ii) policies that maximize the (steady state) fraction of college educated; and (iii) merit-based subsidies. Caucutt and Kumar compare the implications of these policies (relative to the laissez faire) on two measures: (a) welfare (steady state and with transition dynamics) and (b) efficiency of the schooling sector (number of successful graduates relative to the resources used in education). Overall, their results are relatively pessimistic about the benefits of further government subsidies. Policy types (i) and (ii) lead to negligible welfare gains and reductions in the resource efficiency of schooling. Type (iii) policies improve school efficiency but reduce average welfare. Hanushek, Leung and Yilmaz (2003) are also skeptical about the beneficial effects of educational subsidies, arguing instead that wage subsidies dominate in most regards.

Enriching the analysis with early investments significantly changes the implications for government policies. Restuccia and Urrutia (2004) find that increasing government funding for early schooling substantially increases social mobility (i.e. reduces intergenerational persistence), aggregate human capital, consumption and output. By contrast, increasing subsidies to college (late) education has negligible effects on social mobility and produces smaller increases in aggregate human capital, output, and consumption. Moreover, while this policy increases enrollment rates, it also increases dropout rates, reducing the efficiency of the college sector.

Incorporating early investment essentially endogenizes the formation of — and heterogeneity in — ability. This essentially moves these dynastic human capital models closer to the simpler homogeneous agents model of Boldrin and Montes (2004). Indeed, Restuccia and Urrutia (2004) report that calibrating their model without early education requires a much greater exogenous dispersion in innate abilities.

The endogenous response of private market arrangements is another major consideration for the analysis of government policies that is typically neglected. As we stress in the next section, credit constraints arise from repayment incentive problems and institutional features of the economy. These incentive problems are affected by taxes and subsidies that governments impose on the different actions and outcomes of individuals. Indeed, Andolfatto and Gervais (2006) show that the analysis of Boldrin and Montes (2004) may not be robust when credit constraints are endogenously driven by limited commitment. In this case, the pension transfers advocated by Boldrin and Montes would reduce the cost of default and repayment incentives. The net effect can be a *lower* — not higher — supply of resources for youth to invest in human capital.[23]

Finally, the risky nature of human capital can give rise to many incentive problems, including imperfect observability and moral hazard during and after school. Much of the research on human capital has yet to incorporate lessons from the literature on optimal contracts with dynamic incentive problems. We discuss some of these issues in Section 6.

[23] Yang (2011) further examines the conditions under which full efficiency can be restored with endogenous credit constraints.

5.3 Cross-Country Differences in Schooling

It is useful to look beyond the U.S. — the focus of this article — to appreciate the macroeconomic consequences of credit constraints in human capital formation. A large literature examines cross-country differences in income and educational attainment; yet, most of this literature abstracts from borrowing constraints entirely.[24]

Recent work by Cordoba and Ripoll (2011) shows that introducing credit constraints significantly improves the ability of a Ben-Porath (1967) model to explain the cross-country variation in the average years of schooling and the gap between the returns to schooling and the returns on riskless assets. Interestingly, they show that intergenerational constraints and lifecycle borrowing constraints for students yield similar implications. Contrary to a frictionless model, the model of Cordoba and Ripoll implies that parental lifetime income, family size, and the supply of public education are important determinants of education investments. A calibration of their model with low discount rates does a good job of explaining the observed data on educational attainment.

Cordoba and Ripoll (2011) assume exogenous constraints on credit that are uniform across countries. Exploring cross-country differences in access to credit would likely lead to interesting insights, given the evidence (e.g. Filmer and Pritchett 1999) on large cross-country dispersion in the effect of household wealth on educational attainment in developing countries. Moreover, models with endogenous constraints would not only capture the direct impact of different institutions and policies on human capital, but they would also incorporate additional effects due to responses in credit.

6 The Nature of Borrowing Constraints for Education

Despite all the attention paid to credit constraints in the market for human capital, little attention has been paid to the nature of those constraints, i.e. the underlying institutions and incentive problems associated with credit to young individuals with little collateral to pledge while in school. Instead, nearly all studies, theoretical and empirical, assume that

[24] See, e.g., Mankiw, Romer and Weil (1992), Klenow and Rodriguez (1997), Bils and Klenow (2000), Kaboski (2007), Manuelli and Seshadri (2010).

individuals face limits on borrowing as in Section 2 or arbitrary differences in interest rates based on family income. Such simple assumptions are at odds with the actual operation of public and private sources of credit for education.

This section shows that more realistic assumptions about public and private lending can be useful in understanding the behavior of human capital investments, altering key predictions discussed in Section 2.1. We begin by discussing individual behavior when future incomes are certain, then introduce uncertainty about returns on human capital investment.

6.1 Government Student Loans and Limited Commitment

As discussed in Lochner and Monge-Naranjo (2011b), government student loan (GSL) programs explicitly link credit to educational expenditures, while private lenders extend credit to students based on their prospects of repayment and projected future earnings. The following discussion borrows heavily from their analysis. We use the same notation and preferences as in Section 2.1; however, we consider different constraints on borrowing that incorporate central features of existing GSL programs and private lending available for higher education.

GSL programs. Lending programs supported by the federal U.S. government generally have three salient 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, and other expenses directly related to schooling) less any other financial aid they receive in the form of grants or scholarships. Thus, GSL programs do not finance non-schooling related consumption expenses. Second, GSL programs set upper loan limits on the total amount of credit available for each student. Third, loans covered by GSL programs typically have extended enforcement rules compared to unsecured private loans.

To capture these key features of GSL programs, Lochner and Monge-Naranjo (2011b) assume that individuals face two constraints on government loans. First, lending is tied to investment and cannot be used to finance non-schooling related consumption goods or activities, so government borrowing d_g must satisfy $d_g \leq \tau h$.[25] They refer to this as the tied-to-investment constraint (TIC). Second, GSL borrowing is constrained by a fixed upper

[25] In the U.S., GSL programs do not allow for borrowing against foregone earnings costs; however, they do allow for borrowing against certain 'living expenses' (e.g. room and board).

limit $d_g < \bar{d}$ like the exogenous limit of Section 2. Combining these two constraints yields actual credit limits imposed by typical GSL programs:

$$d_g \leq \min\left\{\tau h, \bar{d}\right\}. \tag{10}$$

For now, we assume that GSL repayment is fully enforced. In the next section, we discuss models with default.[26]

Private Lending. Students have increasingly turned to private lending markets to finance their schooling: private student loan amounts skyrocketed from $1.3 billion in 1995-96 to almost $14 billion in 2004-05 (nearly 20 percent of all student loan dollars distributed). Credit card debt among students also rose considerably over this period (College Board 2005).

Private lenders possess weaker enforcement mechanisms than the government to ensure repayment of loans. Indeed, this is the central justification for assuming credit market imperfections in the education sector (Becker 1967). In modeling private lending, Lochner and Monge-Naranjo (2011b) build on recent work on credit constraints that arise endogenously when lenders have limited mechanisms for enforcing repayment.[27] A rational borrower repays private loans if and only if repaying is less costly than defaulting. These limited incentives can be foreseen by rational lenders who, in response, limit their supply of credit to amounts that will be repaid.[28] Since penalties for default typically impose a larger monetary cost on borrowers with higher earnings and assets — only so much can be taken from someone with little to take — credit offered to an individual is directly related to his perceived future earnings. Because expected earnings are determined by ability and investment, private credit limits and investments are co-determined in equilibrium.

It is possible to derive a simple private lending constraint by assuming that defaulting borrowers lose a fraction $0 < \tilde{\kappa} < 1$ of labor earnings.[29] In this case, borrowers repay if and

[26] In practice, default rates have hovered around 5-10% over the past 15 years.

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

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

[29] This is consistent with wage garnishments and costly penalty avoidance actions like re-locating, working in the informal economy, borrowing from loan sharks, or renting instead of buying a home.

only if the payment Rd_p is less than the punishment cost $\tilde{\kappa} a f(h)$. As a result, credit from private lenders is limited to a fraction of post-school earnings:

$$d_p \leq \tilde{\kappa} R^{-1} a f(h). \tag{11}$$

Private credit is directly increasing in both ability and investment. Moreover, ability may also indirectly affect credit through its influence on investment.

Students can borrow d_g from the GSL (subject to (10)) and d_p from private lenders (subject to (11)). Because GSL repayments are fully enforced and do not affect incentives to repay private loans, total borrowing is constrained by

$$d = d_g + d_p \leq \min\left\{h, \bar{d}\right\} + \tilde{\kappa} R^{-1} a f(h). \tag{12}$$

In this simple framework, government credit does not crowd out private credit. Lochner and Monge-Naranjo (2011b) show that in a similar lifecycle model that includes both temporary exclusion from credit markets and wage garnishments as punishments for default, there will be partial crowd out of private credit with expansions in GSL credit. Crowd out occurs, because increases in GSL debt lower incentives to repay private debt.

Empirical Implications. This framework can explain four important empirical patterns in the U.S. over the past few decades: (1) Schooling is strongly positively correlated with ability over time. (2) The correlation between schooling and family income (conditional on ability and family background) has grown since the early 1980s. (3) There has been a sharp increase in the fraction of undergraduates borrowing the maximum amount from GSL programs since the 1990 (Berkner 2000 and Titus 2002). (4) There has been a dramatic rise in student borrowing from private lenders since the mid-1990s (College Board 2005).

As noted in Section 2.1, the standard 'exogenous' constraints model predicts that constrained human capital investment is *decreasing* in ability for constrained individuals under empirically relevant assumptions about preferences for consumption smoothing. Because GSL programs and private lenders link credit to individual ability and investments in human capital, explicitly modeling these endogenous constraints produces a stronger positive relationship between ability and investment for constrained individuals. In contrast to the prediction of the exogenous constraint model of Section 2.1, if $\tilde{\kappa}$ is large enough so that private credit is sufficiently increasing in future earnings, more able individuals may be unconstrained while the least able are constrained (given any level of family resources W).

Calibrating their lifecycle model to the U.S. in the early 1980s, Lochner and Monge-Naranjo show that patterns (2)-(4) can be explained as equilibrium responses to the observed increase in the returns to and costs of college since the early 1980s (given stable GSL limits). Their quantitative analysis suggests that in the early 1980s, the GSL provided adequate credit so that few students would have needed to turn to private creditors. College attendance was, therefore, largely independent of family resources. The rising college costs and returns over time have encouraged more recent students to invest and borrow more, with many exhausting their GSL loans and borrowing substantially from private lenders. Although private lenders have responded to increases in schooling (and its return) by offering more credit, their results suggest that many students with low family resources are now constrained and unable to invest as much as they would like.

The fact that GSL and private credit limits are linked to investment shifts the distortionary effects of credit constraints onto consumption and away from investment. In fact, Lochner and Monge-Naranjo show that credit constrained individuals may not under-invest in human capital. When private loans are unavailable, students constrained only by the GSL's TIC always invest the unconstrained optimal amount — only consumption profiles are affected. When both public and private loans are available, poor low ability youth may actually *over-invest* in human capital.[30]

Lochner and Monge-Naranjo analyze a number of policy issues that cannot be studied without explicitly endogenizing access to credit. For instance, their framework lends itself naturally to an analysis of the interaction between private credit and GSL programs and other government policies. Simulations suggest that expansions of public credit have only modest crowd out effects on private lending. Increases in GSL limits lead to higher levels of total credit and raise human capital investment among youth constrained by those limits.

[30] Abstracting from foregone earnings, when only the GSL's TIC binds, additional investments (at the margin) can be financed fully from the GSL. Further, increases in investment expand private credit that can be used to augment current consumption. While over-investment is theoretically possible, Lochner and Monge-Naranjo's quantitative analysis indicates that it is not empirically relevant given relatively low current GSL limits.

Additionally, they show that changes in GSL credit tend to have a relatively greater impact on investment among the least able, while changes in private loan enforcement tend to impact investment more among the most able. Not all forms of credit expansion are the same.

Finally, endogenous borrowing constraints make human capital investment more sensitive to government education subsidies. Any policy that encourages investment is met with an increase in access to credit, which further encourages the investment of constrained students. This 'credit expansion effect', absent with fixed constraints, can be quite large. Results in Lochner and Monge-Naranjo (2011b) suggest that investment responds as much as 50% more than in the exogenous constraint model.

6.2 Uncertainty, Default and Other Incentive Problems

A simple model with limited commitment captures interesting responses to individual characteristics and investments. However, other important incentive problems cannot be captured unless uncertainty or risk are introduced. We now add uncertainty to the model of Section 2.1 and discuss the implications of limited insurance and private information for the provision of credit and human capital investment behavior. We show how incorporating ideas from the literatures on optimal contracting with limited commitment, private information, and moral hazard can be helpful for understanding schooling, borrowing, and repayment decisions. Furthermore, this analysis provides useful guidance in designing efficient policies to provide both credit and insurance for youth making schooling decisions in a risky environment. See Lochner and Monge-Naranjo (2011a) for a more detailed analysis.

For simplicity, we abstract from forgone wages and normalize tuition costs to one (i.e. $w_0 = 0$ and $\tau = 1$). Assume now that the second period price of human capital is stochastic and can take on $i = 1, ..., N$ possible realizations. Let $p_i > 0$ denote the probability of realization $w_{1,i}$ which we order so that $w_{1,1} < w_{1,2} < ... < w_{1,N}$. Assume that the individual and potential lenders observe the true probabilities as well as individual ability a and initial assets W. Individuals maximize expected utility

$$U = u(c_0) + \beta \sum_{i=1}^{N} p_i u(c_{1,i}),$$

where $c_{1,i}$ is the second period consumption in realization i.

Let D_i be the (possibly negative) quantity that a person commits to repay in the second period, possibly contingent on the realization i. The budget constraints are

$$c_0 = W - h + \sum_{i=1}^{N} q_i D_i,$$
$$c_{1,i} = af(h)\, w_{1,i} - D_i, \quad i = 1, ..., N.$$

Here, q_i is the (Arrow) price of a contingent claim that pays 1 if realization i takes place and zero otherwise. For cases with complete markets, we follow the standard assumption of risk neutral, arbitrage-free asset prices, i.e. $q_i = \beta p_i$.

Unrestricted optima. A complete markets environment has the same separation property as the basic model in Section 2.1. First, human capital investments $h^U(a)$ maximize the expected net present value of lifetime income by equating the marginal cost of investing with the expected marginal return:

$$\bar{w}_1 a f'\left[h^U(a)\right] = \beta^{-1},$$

where $\bar{w}_1 \equiv \sum_{i=1}^{N} p_i w_{1,i}$ is the expected period 1 price of skill. Neither preferences nor initial wealth W have an effect on investment, because there are no restrictions on asset/debt holdings and there is full insurance. Second, asset/debt holdings D_i are set to optimally smooth consumption over time and across states: $u'(c_0) = u'(c_{1,i})$, for all $i = 1, ..., N$.

Limited Commitment with Complete Markets. To introduce limited commitment, assume that individuals can default on their debts in the second period. Doing so, they attain a 'default' utility of $V^D(w_{1,i}, a, h)$, which would generally be increasing in the realization $w_{1,i}$ and in ability a and human capital investments h. Foreseeing these 'participation' constraints, lenders limit debt obligations so that $u[w_{1,i} af(h) - D_i] \geq V^D(w_{1i}, a, h)$. This limits the set of assets/debts individuals can hold as well as their ability to insure against some future states.

Letting $\lambda_i \geq 0$ denote the (discounted) multiplier on participation constraint $i = 1, ..., N$, optimal debt holdings satisfy $u'(c_0) = (1 + \lambda_i)\, u'(c_{1,i})$. For states $w_{1,i}$ in which the participation constraint does not bind ($\lambda_i = 0$), there is perfect consumption smoothing, $c_{1,i} = c_0$. However, if either a is high and/or W is low, the participation constraint may bind for some states, in which case we should observe positive consumption growth, $c_{1,i} > c_0$.

To explore the implications for human capital accumulation, we now focus exclusively on the case in which a borrower who defaults is penalized by forfeiting a fraction $\tilde{\kappa} \in [0,1]$ of his earnings. This implies $V^D(w_{1i}, a, h) = u[(1-\tilde{\kappa}) w_{1i} a f(h)]$, so participation constraints reduce to simple 'solvency' constraints of the form $D_i \leq \tilde{\kappa} w_{1,i} a f(h)$ for all $i = 1, ..., N$. To ensure repayment, the debts carried into any state cannot exceed the income forfeiture. Solvency constraints bind for *high* realizations of $w_{1,i}$, in which case repayments equal $D_i = \tilde{\kappa} w_{1i} a f'(h)$. There is perfect smoothing across low earnings states but only limited insurance in high earnings states.[31] Optimal human capital investment $h^{LC}(a, W)$ satisfies

$$\bar{w}_1 a f'\left[h^{LC}(a, W)\right]\left[\frac{\sum_{i=1}^{N} p_i w_{1,i}\left(\frac{1+\lambda_i \tilde{\kappa}}{1+\lambda_i}\right)}{\bar{w}_1}\right] = \beta^{-1}.$$

When all $\lambda_i = 0$, the unrestricted allocation is attained. Whenever at least one 'solvency' constraint binds, investment is lower than the unrestricted level. This is because $\sum_{i=1}^{N} p_i w_{1,i}\left(\frac{1+\lambda_i \tilde{\kappa}}{1+\lambda_i}\right) < \bar{w}_1$ when $0 < \tilde{\kappa} < 1$ and $\lambda_i > 0$ for some i.

Other implications for investment are also similar to those discussed earlier in the model with perfect certainty. For example, human capital investments help relax solvency constraints in both models. Here, a marginal unit invested increases next period consumption by $w_{1i} a f'(h)$ units if the individual repays, while it only yields $(1-\tilde{\kappa}) w_{1i} a f'(h)$ units if the person were to default. This encourages investment and implies a 'credit expansion' response to education policies as discussed earlier. Furthermore, default does not occur in equilibrium, since all debt repayments are contingent on future states. With such rich contracts, optimal institutional arrangements would minimize the temptation of default by raising $\tilde{\kappa}$ as high as possible ($\tilde{\kappa} = 1$), in which case the economy attains the unconstrained optimal allocation.

Limited commitment with incomplete markets. We now take the opposite extreme from fully contingent contracts and assume that second period liabilities cannot depend on the state, $w_{1,i}$. Because of the incompleteness of contracts, default may now occur in equilibrium. We assume that punishments for default take the same form of a proportional income forfeiture $\tilde{\kappa} w_{1,i} a f(h)$, which is recovered by lenders.

Let $D > 0$ be the amount of debt individuals 'promise' to repay after school. Of course,

[31] Compared with a simple income-contingent repayment scheme in which individuals always repay a constant fraction of their income, these allocations provide greater insurance in low income states.

individuals will actually repay if and only if $D \leq \tilde{\kappa} w_{1,i} a f(h)$. This defines the threshold for $w_{1,i}$, $\tilde{w}_1(D,a,h) \equiv \frac{D}{\tilde{\kappa} a f(h)}$, below which an individual defaults. The probability of default, $\Pr[w_{1,i} < \tilde{w}_1(D,a,h)]$, is weakly increasing in the level of debt D and decreasing in ability a and human capital h. In exchange for a promise to pay $D > 0$, risk-neutral lenders would be willing to extend credit in an amount equal to

$$Q(D,a,h) = \beta \left\{ D - \sum_{w_{1,i} < \tilde{w}_1} p_i \left[D - \tilde{\kappa} w_{1,i} a f(h) \right] \right\}.$$

From the full repayment D, this expression substracts the expected losses $D - \tilde{\kappa} w_{1,i} a f(h)$ from defaulting loans. Expected payments, $Q(D,a,h)$, are not monotonically increasing in debt, since increasing debt can more than proportionally reduce the probability of repayment.[32] A 'hard' borrowing constraint is given by $\sup_D \{Q(D,a,h)\} < \infty$, the maximum value a lender could possibly expect to extract from someone with ability a investing h.

For simplicity, assume that $\tilde{w}_1$ falls outside the support of $w_{1,i}$ and, therefore, ignore jumps in the default probabilities.[33] Under this assumption, marginal changes in D and h do not affect the probability of default, and the necessary first order condition for D is simply

$$u'(c_0) = E[u'(c_{1,i}) \,|w_{1,i} \geq \tilde{w}_1].$$

Optimal borrowing trades-off the gains on consumption c_0 with the costs on future consumption *only* in higher income states of the world in which there is repayment. The necessary condition for optimal h is

$$\bar{w}_1 a f'(h) \left[\frac{\sum_{i=1}^{N} p_i u'(c_{1,i}) w_{1,i} - \tilde{\kappa} \sum_{w_{1,i} < \tilde{w}_1} p_i u'(c_{1,i}) w_{1,i}}{\bar{w}_1 u'(c_0)(1 - Q_h)} \right] = \beta^{-1},$$

where $Q_h > 0$ is the partial derivative (subgradient) of Q with respect to h and must be strictly less than 1 at the optimum.[34] This equation reveals three important differences between investment here and under full insurance. First, additional investment increases

[32]As a function of D, only the increasing region of $Q(\cdot,a,h)$ is relevant.

[33]See Lochner and Monge-Naranjo (2011a) for a complete analysis of the general case.

[34]For a saver, $D < 0$ and $Q(D,a,h) = \beta D$. Thus, $Q_h = 0$ and $Q_D = \beta$.

expected payments, thereby expanding credit. This 'credit expansion' effect encourages investment. Second, some benefits of investment are lost in the event of default since $0 < \tilde{\kappa} < 1$. This new effect arises only because of default and discourages investment. Third, the lack of insurance implies a precautionary motive for investment; however, the riskiness of human capital can also reduce investments as discussed in Krebs (2003).

The absence of repayment contingencies has a number of important consequences. First, default can occur in equilibrium. Second, if default happens, it is for low realizations of $w_{1,i}$ when earnings and consumption are low. Third, the option to default serves an insurance role: given the same liabilities D, the consumption of borrowers would be even lower if they had to fully repay. As a result, eliminating default may be inefficient and could reduce investment. The policy trade-offs in this model are more interesting than in previous models.

Interest rates, implicitly given by $R(D, a, h) \equiv D/Q(D, a, h)$, contain a premium for the possibility of default. Higher $R(\cdot)$ must cover for states in which borrowers default. Ability directly impacts interest rates and credit limits, since $Q_a > 0$; for the same investments h and credit amount Q, more able individuals are asked to repay less. This effect would lead more able persons to invest further in human capital (especially since $Q_{ah} > 0$). Of course, higher investments in human capital would be coupled with higher liabilities, which has the potential to increase the probability of default. In ongoing work, Lochner and Monge-Naranjo (2011a) explore the extent to which this type of model can reproduce observed default rates by ability, debt, and post-school earnings.

Ionescu (2008, 2009, 2011) analyzes models similar to this in order to study college enrollment, borrowing, and default decisions when credit is provided by GSL programs. Her results suggest that default rates are not higher among individuals that are most financially constrained. Most interestingly, she considers the impact of repayment flexibility (e.g., lock-in low interest rates, switching to income contingent repayments, or alternative bankruptcy discharges) in calibrated versions of her models. Overall, she finds that the degree to which contingencies can be incorporated into repayment schemes can have significant effects on schooling. Her analysis suggests that more than hard borrowing constraints, the lack of insurance can sometimes be the limiting factor for schooling decisions. This general conclusion is consistent with the quantitative analysis of Krebs (2003) as well as the structural estimates of Johnson (2010).

Private Information and Limited Insurance. Conceptually, the lack of insurance assumed above is better seen as arising from imperfect information. As such, it is natural to consider some of the lessons and modeling approaches from the vast literature on optimal contracting under private information.

First, consider ex-post asymmetric information. Lenders may not be able to offer income contingent repayments if they cannot observe the ex-post circumstances of a borrower. Yet, when outcomes can be observed at a cost, the possibility of partial insurance arises. In this case, it is natural to adapt the model of *costly state verification* (Townsend 1979) to our human capital setting. This framework is appealing in that it replicates important features of actual bankruptcy institutions as well as some features of income contingent student loan programs.

If a cost must be incurred for lenders to observe the post-schooling earnings of a borrower, the optimal contract is remarkably simple. For high realizations of $w_{1,i}$, borrowers would simply repay a fixed amount (avoiding any verification costs), while an audit would take place for lower realizations. Observing the actual outcome (through verification), a risk-neutral lender would provide a constant consumption level (i.e. full insurance) to the borrower in 'low' states of the world. Thus, the worst ex post outcomes would be fully insured against (as opposed to partial insurance implicit in basic income-contingent loan programs.)

Given a uniform cost of verification, the fact that higher ability implies higher earnings suggests that the probability of verification will be lower for more able individuals, while their consumption is likely to be higher when verification occurs. Higher family resources would imply lower leverage and, hence, a lower probability of verification. These effects on the terms of insurance would tend to produce more positive ability – investment and family resources – investment relationships.

Next, consider *moral hazard problems in investment.*[35] Suppose that in addition to observable investment h, young individuals must exert unobservable costly effort that affects post-schooling earnings (with higher returns to effort for more able individuals). It is well-known that to induce effort, the degree of insurance must be reduced. This basic result

[35]See Chatterjee and Ionescu (2010) for a model with graduation failure with moral hazard.

suggests important limits on credit programs. Some higher ability individuals may not obtain adequate credit, because lenders foresee (correctly) the toll that high debts impose on effort incentives. Still, more able individuals are likely to make greater observed investments and unobserved effort in equilibrium.

Finally, consider *post-schooling moral hazard problems.* Effort must be exerted to seek, keep and improve one's job after school is over. If these efforts are costly for the borrower and unobserved by the creditor, a high debt may affect labor market outcomes. In turn, those effects determine access to credit in the first place.

In the last two decades, an extensive literature on optimal unemployment insurance has been developed (e.g. Hopenhayn and Nicolini, 1997, Acemoglu and Shimer, 1999 and Shimer and Werning, 2008); however, this literature generally considers the welfare of workers once human capital has been formed. Unfortunately, little is known about the joint design of optimal policies that provide both access to credit for education and insurance against post-schooling labor market uncertainty when moral hazard is a problem.

7 Conclusions

Our review of the evidence suggests that, in recent years, credit constraints have become more important for higher education decisions in the U.S. The significant rise in the costs of and returns to college have increased the demand for credit well beyond the supply available from government programs. As such, the rapid expansion in private lending over the past fifteen years should not come as a surprise. Providing credit for human capital, however, requires repayment enforceability and raises other incentive problems. As in Lochner and Monge-Naranjo (2011b), we have argued that explicitly incorporating these incentive problems in models of human capital formation can help explain observed cross-sectional patterns and shed new light on schooling responses to policies and economic changes.

The importance of credit constraints extends beyond their impacts on college-going. Distortions in student consumption and leisure have been documented even during periods when college outcomes were not (e.g. the early 1980s). More importantly, recent evidence highlights the adverse impacts family borrowing constraints can have on early investments in children. There are good reasons to believe that these early constraints are more pervasive and harmful than constraints at college ages. Recent work on the dynamic complementarity in investments suggests that under-investment at early ages may explain why relaxing constraints at later ages often has little impact. Instead, government policies targeting younger ages can have much larger effects.

Credit constraints affect the degree of social mobility, the evolution of the income distribution, and aggregate output and overall welfare. Quantitative macro studies have been successful in replicating important cross-sectional and intergenerational patterns in the data. However, considerable work remains to incorporate dimensions of heterogeneity and the life-cycle, as emphasized in the applied micro literature.

It is unfortunate that most of the human capital literature has ignored the vast literature on optimal contracts with incentive constraints. We have shown how standard results in this literature can be easily adapted to models of human capital formation, leading to new insights on the way abilities and family resources affect investments in human capital and a better understanding of how to best design government policies.

8 Summary Points

- Evidence suggests that borrowing constraints have become more severe for college attendance in recent years.
- In addition to college attendance, borrowing constraints affect consumption and work/leisure while in school.
- Evidence suggests borrowing constraints may be more salient for family investments in younger children than at college ages.
- Early borrowing constraints and complementarity between early and late investments suggest that policies aimed at earlier ages offer more promise.
- Credit constraints shape the degree of social mobility, income distribution and overall development and welfare of countries.
- Government student loan programs link borrowing to educational investments, while private lenders offer credit based on future earnings, which depends on ability as well as investments.
- The link between government and private credit and schooling generates a private 'credit expansion effect' which strengthens educational investment responses to many education policies.
- Lack of insurance can be a major deterrent to human capital investments. Optimal lending would provide insurance considering incentive problems arising from limited observability and limited enforceability.

9 Future Issues

- Additional work is needed to measure the extent to which early family credit constraints inhibit early childhood investments and affect later educational outcomes and earnings.

- Future empirical studies are needed to better understand the skill production technology, especially with respect to the dynamic complementarity of investments from birth through early adulthood.

- Given improvements in computing power, additional margins of heterogeneity and realistic life-cycle dynamics can be readily introduced in quantitative general equilibrium models of human capital.

- To better understand cross-country differences in aggregate human capital, additional work is needed to consistently measure differences in access to and prices of credit for education.

- Additional empirical studies are needed to better understand the extent to which different individual characteristics and choices, as well as government policies, affect repayment of government and private student loans.

- Adapting well-known results from the optimal contracts literature to human capital accumulation problems should lead to interesting insights about the impacts of ability and family wealth on schooling as well as the optimal design of government lending programs.

- Little is known about the impact of student debt on post-school labor market performance. Future studies in this are can shed light on the importance of moral hazard in the design of optimal student loan contracts.

- A promising avenue of research is integrating the optimal unemployment insurance literature with the optimal design of credit programs for human capital accumulation.

References

Acemoglu, D. and Shimer, R. (1999), 'Efficient unemployment insurance', *Journal of Political Economy* **107**(5), 893–928.

Aiyagari, S. R. (1994), 'Uninsured idiosyncratic risk and aggregate saving', *Quarterly Journal of Economics* **109**(3), 659–684.

Aiyagari, S. R., Greenwood, J. and Seshadri, A. (2002), 'Efficient investment in children', *Journal of Economic Theory* **102**(2), 290–321.

Albuquerque, R. and Hopenhayn, H. A. (2004), 'Optimal lending contracts and firm dynamics', *Review of Economic Studies* **71**(2), 285–315.

Alvarez, F. and Jermann, U. J. (2000), 'Efficiency, equilibrium, and asset pricing with risk of default', *Econometrica* **68**(4), 775–797.

Andolfatto, D. and Gervais, M. (2006), 'Human capital investment and debt constraints', *Review of Economic Dynamics* **9**(1), 52–67.

Andreoni, J. (1989), 'Giving with impure altruism: Applications to charity and ricardian equivalence', *Journal of Political Economy* **97**(6), 1447–1458.

Becker, G. S. (1967), Human capital and the personal distribution of income: An analitical approach, *in* 'Woytinsky Lecture No. 1', University of Michigan, Institute of Public Administration, Ann Arbor.

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

Becker, G. S. and Tomes, N. (1979), 'An equilibrium theory of the distribution of income and intergenerational mobility', *Journal of Political Economy* **87**(6), 1153–1189.

Becker, G. S. and Tomes, N. (1986), 'Human capital and the rise and fall of families', *Journal of Labor Economics* **4**(3), S1–S39.

Belley, P., Frenette, M. and Lochner, L. (2011), Post-secondary attendance by parental income in the U.S. and Canada: What role for financial aid policy?, Working Paper 17218, National Bureau of Economic Research.

Belley, P. and Lochner, L. (2007), 'The changing role of family income and ability in determining educational achievement', *Journal of Human Capital* **1**(1), 37–89.

Ben-Porath, Y. (1967), 'The production of human capital and the life cycle of earnings', *Journal of Political Economy* **75**(4), 352–365.

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

Bils, M. and Klenow, P. J. (2000), 'Does schooling cause growth?', *American Economic Review* **90**(5), 1160–1183.

Blau, D. and Currie, J. (2006), Pre-school, day care, and after-school care: Who's minding the kids?, *in* E. Hanushek and F. Welch, eds, 'Handbook of the Economics of Education', Vol. 2, Elsevier, Amsterdam, chapter 20, pp. 1163–1278.

Board, C. (2005), *Trends in College Pricing 2005*, Colleger Board Publications, New York, NY.

Boldrin, M. and Montes, A. (2004), The intergenerational state: Education and pensions, Staff Report 336, Federal Reserve Bank of Minneapolis.

Brown, M., Scholz, J. and Seshadri, A. (2011), A New test of Borrowing Constraints for Education. University of Wisconsin, Working paper.

Browning, M., Hansen, L. P. and Heckman, J. J. (1999), Microdata and general equilibrium models, *in* J. B. Taylor and M. Woodford, eds, 'Handbook of Macroeconomics', Vol. 1, Elsevier, Amsterdam, chapter 8, pp. 543–633.

Cameron, S. and Heckman, J. J. (1998), 'Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males', *Journal of Political Economy* **106**(2), 262–333.

Cameron, S. and Heckman, J. J. (1999), Can tuition policy combat rising wage inequality?, *in* M. H. Kosters, ed., 'Financing College Tuition: Government Policies and Educational Priorities', American Enterprise Institute Press, Washington, pp. 76–124.

Cameron, S. and Taber, C. (2004), 'Estimation of educational borrowing constraints using returns to schooling', *Journal of Political Economy* **112**(1), 132–82.

Card, D. E. (1995), Earnings, schooling, and ability revisited, *in* S. W. Polachek, ed., 'Research in Labor Economics', Vol. 14, JAI Press, Greenwich, CT, pp. 23–48.

Card, D. E. (1999), The causal effect of education on earnings, *in* O. C. Ashenfelter and D. Card, eds, 'Handbook of Labor Economics', Vol. 3, Part 1, Elsevier, Amsterdam, chapter 30, pp. 1801 – 1863.

Carneiro, P. and Heckman, J. J. (2002), 'The evidence on credit constraints in post-secondary schooling', *Economic Journal* **112**(482), 705–734.

Caucutt, E. and Kumar, K. (2003), 'Higher education subsidies and heterogeneity: A dynamic analysis', *Journal of Economic Dynamics and Control* **27**(8), 1459–1502.

Caucutt, E. and Lochner, L. J. (2006), Borrowing constraints on families with young children, *in* 'Innovation in Education', Federal Reserve Bank of Cleveland, Cleveland, pp. 39–48.

Caucutt, E. and Lochner, L. J. (2011), Early and late human capital investments, borrowing constraints, and the family. University of Western Ontario, Working Paper.

Chatterjee, S., Corbae, D., Nakajima, M. and Ros-Rull, J.-V. (2007), 'A quantitative theory of unsecured consumer credit with risk of default', *Econometrica* **75**(6), 1525–1589.

Chatterjee, S. and Ionescu, F. (2010), Insuring student loans against the risk of college failure, Working Paper No. 10-31, Federal Reserve Bank of Philadelphia.

Cordoba, J. C. and Ripoll, M. (2011), What explains schooling differences across countries? University of Pittsburgh, Working Paper.

Cunha, F. and Heckman, J. (2007), 'The technology of skill formation', *American Economic Review* **97**(2), 31–47.

Cunha, F., Heckman, J. J., Lochner, L. and Masterov, D. V. (2006), Interpreting the evidence on life cycle skill formation, *in* E. Hanushek and F. Welch, eds, 'Handbook of the Economics of Education', Vol. 1, Elsevier, Amsterdam, chapter 12, pp. 697–812.

Cunha, F., Heckman, J. J. and Schennach, S. M. (2010), 'Estimating the technology of cognitive and noncognitive skill formation', *Econometrica* **78**(3), 883–931.

Dahl, G. and Lochner, L. (forthcoming), 'The impact of family income on child achievement: Evidence from the earned income tax credit', *American Economic Review* .

Duncan, G. and Brooks-Gunn, J. (1997), *Consequences of Growing Up Poor*, Russell Sage Foundation, New York.

Duncan, G., Morris, P. and Rodrigues, C. (forthcoming), 'Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments', *Developmental Psychology* .

Duncan, G., Yeung, W.-J., Brooks-Gunn, J. and Smith, J. (1998), 'How much does childhood poverty affect the life chances of children?', *American Sociological Review* **63**(3), 406–423.

Ellwood, D. T. and Kane, T. J. (2000), Who is getting a college education? Family background and the growing gaps in enrollment, *in* S. Danziger and J. Waldfogel, eds, 'Securing the Future: Investing in Children from Birth to College', Russell Sage Foundation.

Fernandez-Villaverde, J. and Krueger, D. (2004), Consumption and saving over the life cycle: How important are consumer durables? University of Pennsylvania, Working Paper.

Filmer, D. and Pritchett, L. (1999), 'The effect of household wealth on educational attainment: Evidence from 35 countries', *Population and Development Review* **25**(1), 85–120.

Galor, O. and Zeira, J. (1993), ‘Income distribution and macroeconomics’, *Review of Economic Studies* **60**(1), 35–52.

Gropp, R., Scholz, J. K. and White, M. J. (1997), ‘Personal bankruptcy and credit supply and demand’, *Quarterly Journal of Economics* **112**(1), 217–251.

Hanushek, E. A., Leung, C. K. Y. and Yilmaz, K. (2003), ‘Redistribution through education and other transfer mechanisms’, *Journal of Monetary Economics* **50**(8), 1719–50.

Heckman, J. and Vytlacil, E. (1998), ‘Instrumental variables methods for the correlated random coefficient model: Estimating the average rate of return to schooling when the return is correlated with schooling’, *Journal of Human Resources* **33**(4), 974–87.

Hopenhayn, H. A. and Nicolini, J. P. (1997), ‘Optimal unemployment insurance’, *Journal of Political Economy* **105**(2), 412–438.

Imbens, G. and Angrist, J. (1994), ‘Identification and estimation of local average treatment effects’, *Econometrica* **62**(2), 467–75.

Ionescu, F. (2008), ‘Consolidation of student loan repayments and default incentives’, *The B.E. Journal of Macroeconomics* **8**(1), 1682–1682.

Ionescu, F. (2009), ‘The federal student loan program: Quantitative implications for college enrollment and default rates’, *Review of Economic Dynamics* **12**(1), 205 – 231.

Ionescu, F. (2011), ‘Risky human capital and alternative bankruptcy regimes for student loans’, *Journal of Human Capital* **5**(2), 153–206.

Johnson, M. (2010), Borrowing constraints, college enrolment, and delayed entry. Mathematica Policy Research, Working Paper.

Kaboski, J. (2007), Explaining returns to schooling and output across countries. University of Notre Dame, Working Paper.

Karoly, L. et al. (1998), *Investing In Our Children: What We Know And Don’t Know About The Costs And Benefits Of Early Childhood Interventions*, The RAND Corporation, Santa Monica.

Keane, M. P. and Wolpin, K. I. (1997), 'The career decisions of young men', *Journal of Political Economy* **105**(3), 473–522.

Keane, M. and Wolpin, K. (2001), 'The effect of parental transfers and borrowing constraints on educational attainment', *International Economic Review* **42**(4), 1051–1103.

Kehoe, T. J. and Levine, D. K. (1993), 'Debt-constrained asset markets', *Review of Economic Studies* **60**(4), 865–888.

Kinsler, J. and Pavan, R. (2010), College quality, educational attainment, and family income. University of Rochester, Working Paper.

Klenow, P. and Rodrguez-Clare, A. (1997), The neoclassical revival in growth economics: Has it gone too far?, *in* B. S. Bernanke and J. J. Rotemberg, eds, 'NBER Macroeconomics Annual 1997', Vol. 12, MIT Press, pp. 73–114.

Kocherlakota, N. R. (1996), 'Implications of efficient risk sharing without commitment', *Review of Economic Studies* **63**(4), 595–609.

Krebs, T. (2003), 'Human capital risk and economic growth', *Quarterly Journal of Economics* **118**(2), 709–744.

Krueger, D. and Perri, F. (2006), 'Does income inequality lead to consumption inequality? Evidence and theory', *Review of Economic Studies* **73**(1), 163–193.

Lang, K. (1993), Ability bias, discount rate bias and the return to education. Unpublished Manuscript, Boston University.

Levy, D. and Duncan, G. (1999), Using sibling samples to assess the effect of childhood family income on completed schooling. Working Paper.

Livshits, I., MacGee, J. and Tertilt, M. (2007), 'Consumer bankruptcy: A fresh start', *American Economic Review* **97**(1), 402–418.

Lochner, L. and Monge-Naranjo, A. (2011*a*), Credit and insurance for human capital investments. University of Western Ontario/Pennsylvania State University, Mimeo.

Lochner, L. and Monge-Naranjo, A. (2011*b*), 'The nature of credit constraints and human capital', *American Economic Review* **101**(6).

Løken, K. (2010), 'Family income and children's education: Using the norwegian oil boom as a natural experiment', *Labour Economics* **17**(1), 118–29.

Løken, K., Mogstad, M. and Wiswall, M. (2010), What linear estimators miss: Re-examining the effects of family income on child outcomes. Working Paper.

Loury, G. C. (1981), 'Intergenerational transfers and the distribution of earnings', *Econometrica* **49**(4), 843–867.

Lovenheim, M. F. (2011), The effect of liquid housing wealth on college enrollment. Working Paper.

Lovenheim, M. F. and Reynolds, C. L. (2011), 'Changes in postsecondary choices by ability and income: Evidence from the national longitudinal surveys of youth', *Journal of Human Capital* **5**(1), 70–109.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992), 'A contribution to the empirics of economic growth', *Quarterly Journal of Economics* **107**(2), 407–437.

Manuelli, R. E. and Seshadri, A. (2010), Human capital and the wealth of nations. Working Paper.

Milligan, K. and Stabile, M. (2011), 'Do child tax benefits affect the well-being of children? Evidence from canadian child benefit expansions', *American Economic Journal: Economic Policy* **3**(3), 175–205.

Monge-Naranjo, A. (2009), 'Entrepreneurship and firm heterogeneity with limited enforcement', *Annals of Finance* **5**(3-4), 465–494.

Navarro, S. (2010), Using observed choices to infer agent's information: Reconsidering the importance of borrowing constraints, uncertainty and preferences in college attendance. University of Wisconsin, Working Paper.

Restuccia, D. and Urrutia, C. (2004), 'Intergenerational persistence of earnings: The role of early and college education', *American Economic Review* **94**(5), 1354–1378.

Shimer, R. and Werning, I. (2008), 'Liquidity and insurance for the unemployed', *American Economic Review* **98**(5), 1922–1942.

Stinebrickner, T. and Stinebrickner, R. (2008), 'The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study', *American Economic Review* **98**(5), 2163–84.

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

Townsend, R. M. (1979), 'Optimal contracts and competitive markets with costly state verification', *Journal of Economic Theory* **21**(2), 265 – 293.

Willis, R. J. and Rosen, S. (1979), 'Education and self-selection', *Journal of Political Economy* **87**(5), S7–S36.