Estimation of Educational Borrowing Constraints Using Returns to Schooling

Stephen V. Cameron

Columbia University and Federal Reserve Bank of New York

Christopher Taber

Northwestern University

This paper measures the importance of borrowing constraints on education decisions. Empirical identification of borrowing constraints is secured by the economic prediction that opportunity costs and direct costs of schooling affect borrowing-constrained and unconstrained persons differently. Direct costs need to be financed during school and impose a larger burden on credit-constrained students. By contrast, gross forgone earnings do not have to be financed. We explore the implications of this idea using four methodologies: schooling attainment models, instrumental variable wage regressions, and two structural economic models that integrate both schooling choices and schooling returns into a unified framework. None of the methods produces evidence that borrowing constraints generate inefficiencies in the market for schooling in the current policy environment. We conclude that, on the margin, additional policies aimed at improving credit access will have little impact on schooling attainment.

I.   Introduction

Does access to credit influence educational outcomes? The answer to this question is fundamental to sensible educational policy and to our understanding of intergenerational transmission of inequality, returns to schooling, economic growth, and many other economic phenomena. Since a direct answer to the question depends on economic variables difficult or impossible to measure, this paper explores alternative approaches to identifying and estimating the influence of credit access on educational choice. The key idea for identification of the role of credit in determining schooling choices is that two types of schooling costs, opportunity costs of schooling (the value of earnings forgone while in school) and direct costs of schooling (monetary cost of tuition, books, transportation, and board and room if necessary), affect schooling choices differently for credit-constrained and unconstrained individuals. Direct costs need financing during periods of school enrollment and present a larger challenge to credit-constrained persons. Gross forgone earnings, by contrast, do not have to be financed during school.[1]

Empirical implementation of our approach requires distinct measures of both direct schooling costs and opportunity costs. Following Card (1995b), we use as a proxy for direct costs an indicator variable for the presence of any college (either a two- or four-year college) in an individual's county of residence. Opportunity costs are proxied by measures of earnings in low-skill industries located in a person's local labor market at the time schooling decisions are made. High prevailing wages in the local labor market imply high forgone earnings of college attendance.

The paper reports empirical findings from four methodologically distinct econometric approaches to estimating the importance of credit access on schooling choice. Each method exploits implied differences in responses of borrowing-disadvantaged students to direct costs and forgone earnings of students to identify the magnitude of credit constraints on schooling choice.

The first approach begins by following Lang (1993) and Card (1995a) and estimating returns to education using an instrumental variable approach.[2] These authors explain that when the causal effect of schooling on earnings is heterogeneous in the population, instrumental variable techniques do not recover the population average payoff to schooling but instead estimate the payoff to schooling for those marginal individuals whose schooling choices are most affected by changes in the value

---

[1] Gross forgone earnings are defined as the sum of all income earned by a person not attending school. For expositional simplicity, our story abstracts from income students earn during school enrollment. Those earnings would enter the model as an offset to direct college costs, and including them affects none of the implications of our model.

[2] We use the term "returns to schooling" to denote the coefficient on schooling in a log wage regression. It is not necessarily the internal rate of return.

of the instrument. They argue that the estimated return to schooling could be higher than the population average return if instrumental variable estimation identifies the return received for the credit-hindered subset of the college-going population. Lang terms this phenomenon "discount rate bias." We take this argument one step further by using it to develop a test for borrowing constraints based on estimated returns to schooling using direct costs and opportunity costs of schooling as alternative instruments. Since schooling decisions of borrowing-constrained students are regulated more by changes in direct schooling costs, the population subset that attends college in response to a fall in direct college costs contains a large concentration of credit-disadvantaged students. As borrowing is costly, credit-disadvantaged persons require a monetary return to college higher than a person of equal ability with access to cheap credit. Thus using measures of direct costs as instruments in wage studies results in high estimated values of schooling returns if credit constraints hamper schooling decisions. In the language of Lang (1993), the "discount rate bias" would be high. By contrast, instrumenting with measures of opportunity costs should lead to smaller estimates of the schooling effect since opportunity costs do not differentially deter credit-disadvantaged persons from college. In other words, this estimator has less "discount rate bias." Therefore, under the assumption that instruments for either direct or opportunity costs are free of ability bias, the presence of borrowing constraints implies that an instrumental variable estimator based on an instrument for direct costs of schooling recovers estimates of schooling returns larger than those obtained when forgone earnings are used as an instrument.

Empirically, this method uncovers no evidence of educational borrowing constraints. In fact, the pattern of point estimates from two-stage least-squares estimates is consistently opposite the pattern predicted if constraints mattered.

Our second approach looks directly at schooling attainment. Exploiting the implication of our behavioral model that credit-hindered persons are more sensitive to direct schooling costs, we study interactions between schooling costs and observable characteristics likely correlated with borrowing constraints. This method, too, fails to detect evidence of differential credit access. The coefficients on the interactions often turn up with the wrong sign and are consistently insignificant.

The third approach formalizes the intuition behind the second approach into a structural econometric model of schooling choice. As this methodology combines information from both schooling returns and schooling choices into a single model, it uses more information and incorporates behavioral restrictions on returns and choices to yield point estimates of the magnitude of borrowing constraints. This method also delivers a framework to make comparisons between a control group of

unconstrained and a comparison group of potentially credit-constrained persons after explicitly taking into account differences in forgone earnings, schooling costs, and other characteristics. This model and the generalization described next are of independent interest and have methodological value beyond this study. In the application here, the return to schooling and the interest rate at which students borrow to finance their education are assumed to vary across individuals. Borrowing rates depend on observed characteristics of individuals.

This method provides the strongest evidence against the importance of borrowing constraints for educational decisions. The rates at which individuals borrow are precisely estimated, and borrowing constraints are very small.

The fourth approach generalizes the third to allow educational borrowing rates and returns to depend on influences unobserved by the econometrician. Like the instrumental variable approach, identification is secured from the use of direct costs and indirect costs as exclusion restrictions. We show formally that these two types of exclusion restrictions enter the model in different ways and nonparametrically identify the magnitude of borrowing constraints. Empirically, estimates of the distribution of borrowing-constrained persons turn out to be degenerate with no individuals borrowing-constrained.

Our results are consistent across all four approaches: none find evidence that borrowing access is an important component of schooling decisions. With this finding in mind, however, it is important to underscore the limits of our study. First, we cannot directly observe credit access for individuals, so identification of the extent of borrowing constraints is indirect. Second, our empirical findings cannot address whether all students in all environments had adequate access to credit for educational investments. During the period covered by our data, large subsidies to school and college were already in place in the United States. Given the policy regime, we find no evidence of inefficiencies in the schooling market resulting from borrowing constraints. This finding suggests that additional subsidies aimed at improving credit access will have little net value and little impact on overall schooling attainment.

The paper unfolds as follows. Section II provides a review of the literature, and Section III presents the economic model and frames the discussion of borrowing constraints in the context of the model. Section IV describes the data. Section V summarizes empirical findings on the question of borrowing constraints from linear regression and instrumental variable estimates of returns to education and analyses of schooling attainment. Section VI lays out the structural models and reports empirical findings. The paper concludes with a summary (Sec. VII).

II.   Previous Work

Evidence favoring the idea that borrowing constraints hinder educational progression, particularly at the college level, is based almost entirely on well-documented correlations between schooling attainment and family income and other family characteristics. The step from correlation to causation is precarious since family income is also strongly correlated with early schooling achievement, where direct costs of schooling have little role. Recent work by Cameron and Heckman (1998, 2001), Shea (2000), and Keane and Wolpin (2001) has attempted to better understand the determinants of schooling choices. Using very different empirical approaches, these researchers turn up no evidence that borrowing constraints hamper college-going or any other schooling decision.

The credit constraint question has also surfaced in the recent literature on returns to education. This literature has aimed at recovering returns to schooling from wage regressions purged of "ability bias." Unobserved ability is thought to bias upward least-squares estimates of returns to schooling. Using instrumental variable methods to correct for the bias, researchers have devised a wide variety of different instruments and typically find instrumental variable estimates anomalously larger than least-squares estimates (Card 1999, 2001). The connection between credit access and measured returns to schooling—a link originally clarified by Becker (1967)—has been investigated by Lang (1993) and Card (1995a) as an explanation for these high instrumental variable estimates. This argument presumes that borrowing constraints are important for schooling decisions.[3]

By and large, the literature studying determinants of schooling and the literature on returns to schooling have evolved in isolation from one another. Both literatures are discussed in the next subsections.[4] Few empirical studies have integrated educational attainment and schooling returns into a unified framework. An important exception is the pioneering work of Willis and Rosen (1979), on which this paper builds and which is discussed at the beginning of Section III.

---

[3] We use the terms "borrowing constraint" and "credit constraint" synonymously and broadly to include not only a hard constraint, which prevents any borrowing outside the family, but also the more standard case in which interest rates for educational borrowing are higher than market interest rates. Definitions are clarified in Sec. III.

[4] Lochner and Monge (2001) extend the literature on borrowing constraints and credit constraints in a different direction. They present a schooling model in which credit constraints are endogenous.

A.   Literature on Determinants of Educational Attainment

A ubiquitous empirical regularity that emerges from the literature on determinants of schooling is the strong correlation between family income and schooling attainment. This correlation has been documented in legions of U.S. data sets covering the entire twentieth century (see, e.g., Mare 1980; Manski and Wise 1983; Hauser 1993; Manski 1993; Kane 1994; Mayer 1997; Cameron and Heckman 1998, 2001; Levy and Duncan 2000) and in data from dozens of other countries in many stages of political and economic development (see, e.g., the studies collected in Shavit and Blossfeld [1993]).[5] Educational financing constraints have been the popular behavioral interpretation of the schooling–family income correlation, particularly in studies of college attendance.

However, credit access is only one of many possible interpretations of this correlation. Family income and other family background measures have been found to be correlated with achievement test performance in elementary and secondary school as well as with schooling continuation choices at all levels of schooling from eighth grade through graduate school. Cameron and Heckman (1998, 2001) adopt a "life cycle" view of the importance of family income and other family factors and argue that family income is a prime determinant of the string of early schooling decisions. They conclude that the measured effect of family income on continuing college is largely a proxy for its influence on early achievement.[6]

Shea's (2000) findings support this interpretation of the data. He isolates the component of family income variation that could arguably be ascribed to "luck," coming from union status, job loss, and other factors, to estimate the causal effect of income on schooling. He finds little or no correlation between this component of income and children's schooling outcomes.[7] Keane and Wolpin (2001) take a different approach. They estimate a rich discrete dynamic programming model of schooling, work, and savings. Model simulations reveal that relaxing borrowing constraints has almost no effect on schooling, but these constraints are important determinants of working during school.

[5] Tomes (1981) and Mulligan (1997) are related but look at the elasticity of schooling to income. They find higher elasticities for families that are more likely to be borrowing-constrained.

[6] Carneiro, Heckman, and Manoli (2002) extend this work using similar methods but get different outcomes. They find evidence that suggests that borrowing constraints may delay entrance and affect college completion and quality.

[7] Shea studies extracts from the Panel Study of Income Dynamics and finds no effects in the full sample. He does find modest evidence of a relationship in the low-income subsample, though such a finding is not inconsistent with Cameron and Heckman's (2001) view that family income effects operate at the earliest stages of schooling.

B.   Returns to Education Literature

A large literature in labor economics has been concerned with estimating the causal effect of schooling on earnings. Ordinary least squares (OLS) regressions of earnings on schooling have long been believed to be biased upward as a result of "ability bias": individuals who attain higher levels of schooling do so in part because they are smart and earn a return on that characteristic as well as on their additional years of education. Empirical evidence for this idea has been found in virtually every data set with pre–labor market measures of scholastic ability, such as standardized test scores. Including test scores in wage regressions leads to a decline in the estimated effect of schooling. Nevertheless, scholastic test scores are imperfect measures of earning ability and leave substantial scope for bias from unobserved components of ability.

A good deal of recent work uses instrumental variables or related techniques to address the problems caused by omitted ability measures.[8] Card (2001) provides an extensive survey of this literature. Contrary to the intuition provided by the ability bias story, Card documents that researchers often find that the estimated coefficient on schooling rises rather than falls when instrumental variable procedures are used. Building on the model in Becker's (1967) Woytinsky lecture, Lang (1993) and Card (1995*a*) explore heterogeneity in borrowing rates as an explanation for this pattern, which Lang terms "discount-rate bias."

In Becker's model, a student invests in schooling until her return is equal to the interest rate she faces. If borrowing-constrained individuals face higher personal interest rates, they will demand higher returns from schooling at the margin. If returns to schooling vary across individuals because of differential credit costs, the pattern of estimates produced by instrumental variable estimators may be explained by the fact that many of these estimators identify the causal effect of schooling from the subset of individuals who are borrowing-constrained and whose return to schooling is higher than the population average return. If returns to schooling were homogeneous, instrumental variables would yield consistent estimates of the causal effect of education for the population. When returns are heterogeneous, instrumental variable estimates must be interpreted with care. Imbens and Angrist (1994) show that instrumental variable estimates measure the treatment effect of schooling (i.e., the causal effect) for groups whose schooling decisions are most sensitive to changes in the instrument used in estimation. For instance, schooling choices of borrowing-constrained individuals may be most sensitive to changes in college tuition. Because of their higher

---

[8] Altonji and Dunn (1996) use a somewhat different strategy that is relevant to our study. They look for interactions between the return to schooling and family background. Their results are mixed, but some specifications point to a positive interaction.

costs of raising funds for schooling, borrowing-constrained individuals also demand the highest returns to continue. Thus the instrumental variable estimate of schooling returns when returns are instrumented with college costs will be an average of returns for the constrained group and will be higher than the population average return.[9] This argument also helps explain patterns of estimates obtained from studies employing "selection" models of schooling. Selection models take into account ability bias, but discount rate bias does not appear. These studies generally report lower estimated returns to education than those obtained from OLS (see, e.g., Willis and Rosen 1979; Taber 2001).

III.   The Model

An economic description of schooling choice is developed in this section. The model illustrates differences in the influence of direct schooling costs and opportunity costs on the schooling choices of borrowing-hindered and unhindered students. The differential influence of each type of cost is essential for identification in all the empirical approaches developed below.

   The closest antecedent of our work is the paper by Willis and Rosen (1979), who integrate future returns to education into their analysis of schooling decisions to account for self-selection into college attendance. Our framework extends their work in two important respects. First, Willis and Rosen study only college-entry decisions among a sample of high school graduates. We do not condition our sample on the high school graduation decision; rather, we model (1) the decision to complete high school, (2) the decision to enter college, and (3) the decision to complete college. Our analysis avoids the sample selection problem that arises by conditioning on high school graduation.[10]

   Second, and more important, Willis and Rosen estimate a first-order linear approximation of their economic model. By doing so, they confound the influences of direct and indirect costs of going to college. Separating these influences underlies our strategy to identify and estimate the quantitative importance of borrowing constraints.

   To avoid inconsistencies between the theoretical and empirical sections of the paper, we keep the specification of the behavioral model developed in this section as close as possible to the specification of the empirical structural models described in Section VI. The main results

derived below hold true in more elegant and general versions of the model.[11]

The model begins with a specification of individual preferences. Individuals derive utility from consumption and tastes for nonpecuniary aspects of schooling. These nonpecuniary tastes represent the utility or disutility from school itself or preferences for the menu of jobs available at each level of schooling.

Assume that lifetime utility for schooling level S is given by

$$V_S = \sum_{t=0}^{\infty} \delta^t \frac{c_t^\gamma}{\gamma} + T(S), \tag{1}$$

where c_t is consumption at time t, $T(S)$ represents nonpecuniary tastes for schooling level S, δ is the subjective rate of time preference, and γ is a parameter of utility curvature with a value in $(-\infty, 1)$. Define $\mathcal{S}$ to be the set of possible schooling choices. Individuals choose S out of this set so that

$$S = \arg\max \{V_S | S \in \mathcal{S}\}. \tag{2}$$

Much of the schooling literature—including Becker (1967), Rosen (1977), Willis and Rosen (1979), Willis (1986), Lang (1993), and Card (1995a)—models heterogeneity of credit access as a person-specific rate of interest, denoted here as r, at which a person can borrow and save throughout life. Credit-constrained persons borrow at a high r, which makes educational financing costly.

This approach to modeling credit cost heterogeneity has the unattractive feature that a high lifetime r implies high returns to savings after labor market entry. This in turn implies that giving assets to borrowing-constrained individuals prior to college will raise their lifetime wealth but not alter their schooling decisions.[12]

Our model departs from the literature by adopting the simple but novel assumption that individuals borrow at their personal rate r while in school but face a common market interest rate for all borrowing and lending after labor market entry. Confining borrowing rate heterogeneity to the schooling years is a natural assumption if one considers the

---

[11] The implications of the model depend on borrowing-constrained individuals' having higher marginal utility of income when in school than when out of school. The main results go through if credit access is modeled as a hard borrowing constraint (no borrowing outside the family), as a higher interest rate for students than for nonstudents, or as an increasing function of the amount borrowed.

[12] This result follows because the assumption of a constant lifetime r gives rise to a separation result that simplifies the model. Given r, individuals choose schooling to maximize the present value of lifetime earnings. Thus, when interest rates and nonpecuniary schooling tastes are held fixed, income transfers have no direct effect on schooling choices. In our setup, an increase in income transfers influences schooling choices by raising the net value of further schooling.

borrowing rate to be determined by the ability to collateralize loans with personal or family assets during school. The specific form of borrowing constraints is not essential to the results in this section. Since our data do not have information on consumption or assets, we focus on a simple type of borrowing constraint that is straightforward to estimate in the structural econometric model described in Section VI.

Define R to be the borrowing rate during school and let R_m denote the market rate, which is normalized for convenience such that $1/R_m = \delta$. Students maximize utility subject to the lifetime budget constraint

$$\sum_{t=0}^{S-1}\left(\frac{1}{R}\right)^t c_t + \left(\frac{1}{R}\right)^S \sum_{t=S}^{\infty} \delta^{t-S} c_t \le I_S, \tag{3}$$

where S is total years of school and I_S is the present value of income net of direct schooling costs. The first-order conditions are

$$c_t = (\delta R)^{t/(1-\gamma)} c_0, \quad t \le S,$$

$$c_t = (\delta R)^{S/(1-\gamma)} c_0, \quad t > S.$$

Plugging these values into the budget constraint yields

$$I_S = \sum_{t=0}^{S-1} R^{t\gamma/(1-\gamma)} \delta^{t/(1-\gamma)} c_0 + (R\delta)^{S\gamma/(1-\gamma)} \sum_{t=S}^{\infty} \delta^t c_0. \tag{4}$$

Finally, solving c_t in terms of I_S and inserting the value into the utility function leaves us with the following expression for lifetime utility of a person choosing S years of school:

$$V_S = \frac{I_S^\gamma \left[\sum_{t=0}^{S-1} R^{t\gamma/(1-\gamma)} \delta^{t/(1-\gamma)} + (R\delta)^{S\gamma/(1-\gamma)} \sum_{t=S}^{\infty} \delta^t\right]^{1-\gamma}}{\gamma} + T(S). \tag{5}$$

Equation (5) represents the indirect lifetime utility function conditional on schooling choice S.

We next solve for the present value of income. To focus on borrowing constraints, we abstract from earnings uncertainty by assuming that earnings streams associated with all levels of S are known with certainty at time 0.[13] Let w_{St} be earnings at time t for an individual with S years of schooling. Individuals have zero earnings while in school and pay direct cost τ_S at time $S - 1$ to attend schooling level S. Abstracting from labor

[13] Uncertainty in future returns to education introduces an option value to further education. For instance, even if predicted returns to college completion were low, individuals may still graduate from college in case realized returns outperform predicted returns (see Taber 2001).

supply decisions, we have the following expression for the present value of income discounted to time $t = 0$:

$$I_S = \left(\frac{1}{R}\right)^S \sum_{t=S}^{\infty} \delta^{t-S} w_{St} - \sum_{t=0}^{S-1} \left(\frac{1}{R}\right)^t \tau_{t+1}$$

$$= \left(\frac{1}{R}\right)^S W_S - \sum_{t=0}^{S-1} \left(\frac{1}{R}\right)^t \tau_{t+1}, \tag{6}$$

where W_S is the present value of earnings associated with schooling level S discounted to time S.

To illustrate the main predictions of the model, consider how changes in direct costs and opportunity costs affect utility in a world with only two schooling levels, $S = 0$ and $S = 1$. Let τ_1 be the direct cost of $S = 1$, and assume that there are no direct costs associated with $S = 0$. Lifetime utility values for $S = 0$ and $S = 1$ are given by

$$V_0 = \frac{W_0^{\gamma}[1/(1 - \delta)]^{1-\gamma}}{\gamma} + T(0) \tag{7}$$

and

$$V_1 = \frac{[(W_1/R) - \tau_1]^{\gamma}\{1 + (R\delta)^{\gamma/(1-\gamma)}[\delta/(1 - \delta)]\}^{1-\gamma}}{\gamma} + T(1). \tag{8}$$

A person chooses $S = 1$ when $V_1 - V_0 > 0$, and $S = 0$ otherwise. Forgone earnings are represented by the wage rate for unskilled ($S = 0$) workers W_0.

Tastes for schooling $T(0)$ and $T(1)$ are not observed by the analyst but distributed randomly in the population. Conditioning on potential wages, direct costs, and the borrowing rate, we get

$$\Pr(S = 1|W_1, W_0, R, \tau_1) = \Pr(V_1 > V_0|W_1, W_0, R, \tau_1) \tag{9}$$

$$= \Pr(D > T(0) - T(1)|W_1, W_0, R, \tau_1), \tag{10}$$

where

$$D = \frac{[(W_1/R) - \tau_1]^{\gamma}\{1 + (R\delta)^{\gamma/(1-\gamma)}[\delta/(1 - \delta)]\}^{1-\gamma}}{\gamma}$$

$$- \frac{W_0^{\gamma}[1/(1 - \delta)]^{1-\gamma}}{\gamma}. \tag{11}$$

Thus, given tastes for education, the larger the D term, the more likely a person completes schooling level $S = 1$. Given this simple relationship between D and schooling attendance, we focus our attention on how direct and opportunity costs of schooling (W_0 and τ_1) influence D.

To explore the relationship between direct and opportunity costs and

the value of D, consider two individuals with identical preferences, one borrowing-constrained and one not, and let $T(1) = T(0) = 0$ for simplicity. Suppose that both persons are indifferent between attending and not attending school, so that $V_0 = V_1$. The unconstrained person borrows at the market rate $R = 1/\delta$; the constrained person borrows at some rate $R > 1/\delta$. Consider each person's reaction at $t = 0$ to a dollar increase in the present discounted value of forgone earnings and alternatively to a dollar increase in direct schooling costs.

For the person borrowing at the market rate, a dollar is a dollar: changes of equal magnitude in direct costs and opportunity costs have the same influence on the schooling decision. To see this, notice from equations (7) and (8) that $V_0 = V_1$ implies that the present values of income for $S = 0$ and $S = 1$ are equal at time 0: $W_0 = W_1\delta - \tau_1$. Hence, a dollar rise in W_0 and a dollar rise in τ_1 have the same effect on the relative value of $S = 1$:

$$\frac{\partial D}{\partial \tau_1} = \frac{-\gamma V_1}{\delta W_1 - \tau_1} = \frac{-\gamma V_0}{W_0} = \frac{\partial D}{\partial W_0}. \tag{12}$$

By contrast, for the credit-constrained person to be indifferent, equations (7) and (8) show it must be that $W_0 > (1/R)W_1 - \tau_1$. Thus a dollar rise in direct costs of school has a larger (in absolute value) influence on the value of schooling than a dollar drop in forgone earnings:

$$\frac{\partial V_1}{\partial \tau_1} = \frac{-\gamma V_1}{(1/R)W_1 - \tau_1} < \frac{-\gamma V_0}{W_0} = \frac{\partial V_0}{\partial W_0}. \tag{13}$$

Put differently, the shadow value of a dollar of income is higher while in school than while out of school.

The three predictions of the model of interest to us are now stated more precisely as propositions. Note first that a rise in R reduces the likelihood that a person chooses $S = 1$ as long as she is not a net saver while in school:[14]

$$\frac{\partial D}{\partial R} = \frac{\partial V_1}{\partial R} = -\frac{\gamma V_1}{RI_1}(c_0 + \tau_1) < 0, \tag{14}$$

where c_0 is optimal time 0 consumption.

PROPOSITION 1. A dollar rise in τ_1 diminishes the value of V_1 more

[14] This is generally true as long as $c_0 > -\tau_1$ and must hold when schooling costs are nonnegative.

for the individual with higher R as long as that person is not a net saver during school:

$$\frac{\partial^2 D}{\partial R \partial \tau_1} = -\frac{\gamma V_1}{R I_1^2}\left[(c_0 + \tau_1)(1 - \gamma) + c_0 (R\delta)^{\gamma/(1-\gamma)}\frac{\delta}{1 - \delta}\right] < 0. \quad (15)$$

PROPOSITION 2. The influence of W_0 on D does not depend on R:

$$\frac{\partial^2 D}{\partial R \partial W_0} = \frac{\partial W_0^{\gamma-1}[1/(1 - \delta)]^{1-\gamma}}{\partial R} = 0. \quad (16)$$

PROPOSITION 3. Consider credit-constrained and unconstrained individuals c and u who are both indifferent about attending college and receive the same utility if they go to college, so that $V_{0c} = V_{1c} = V_{0u} = V_{1u}$. Assume that they have identical nonpecuniary tastes for schooling and that τ_1, γ, and δ are also identical. The "return to education" for the constrained person is higher than it is for the unconstrained person. That is, $R_c > R_u$ implies $W_{1c}/W_{0c} > W_{1u}/W_{0u}$.

The last implication of the model follows directly from the fact that among the conditioning group, a person borrowing at a high R must be compensated with a higher W_1 to remain indifferent. That is, since $\partial V_1/\partial R < 0$ and $\partial V_1/\partial W_1 > 0$ and if R_c is higher than R_u, for c and u to be indifferent, then W_{1c} must be higher than W_{1u}.

The predictions embodied in these propositions are used in all four empirical approaches explored in Sections V and VI below. The first two approaches use the economic intuition derived here to guide empirical specifications and interpretation of estimated parameters. The second two approaches are based on estimated structural implementations of the behavioral model elaborated here. Econometric identification explicitly requires that W_0 and τ_1 enter the model in different ways.

IV.   The Data

Our analysis is based on black, Hispanic, and white males from the 1979–94 waves of the National Longitudinal Survey of Youth (NLSY). Because the NLSY collected detailed information on family background, scholastic achievement, labor market outcomes, county of residence, and school attendance and completion starting at relatively young ages, it is ideal for our study. The NLSY data comprise four distinct samples: a random sample of the population, a random sample of the black and Hispanic populations, a sample drawn from the military, and a sample of the economically disadvantaged, nonblack, non-Hispanic population.

We limit our sample in four ways. First, we exclude from our analysis the military and the nonblack, non-Hispanic disadvantaged subsamples

because they are not drawn according to exogenously determined characteristics.[15] Second, we use only males because their schooling and labor supply decisions are less complicated by fertility and labor market participation considerations. Third, because information about events occurring before January 1978 is retrospective and limited, we confine our sample to males between ages 13 and 16 as of that date in order to have reliable information on schooling attendance, parental income, and county of residence. County of residence is used to construct measures of labor market conditions and measures of college proximity. Finally, 13 percent of the sample was eliminated because respondents did not complete the Armed Services Vocational Aptitude Battery (ASVAB) (see below) or because of missing data in county of residence during high school (measured at age 17 when available), family income, or one of the family background variables. Final sample sizes are reported in table 1.

We construct panel data from NLSY annual observations. Annual observations on each respondent begin no later than age 16 and extend through ages 29–33. For the analysis of wages, we exclude annual wage observations taken before age 22 because most college graduates were still in school before that age.[16]

Summary statistics of the main variables used in the analysis are presented in table 1. Panels A–C show summary statistics of static variables for years of schooling completed, racial-ethnic identity, family background, and geographic location measured at age 17. Details of the variables are provided in the notes to the table.

Panel D shows summary statistics for test score variables. The scores are taken from the 10-part ASVAB. The test was administered to NLSY respondents in the summer of 1980, when respondents in our sample were between ages 15 and 18. The set of four variables word score, math score, science score, and automotive score are raw scores on four of the ASVAB sections. The variable AFQT score is the score on the Armed Forces Qualification Test, a weighted sum of ASVAB components that measure literacy and basic mathematics skills. This variable has been widely used in recent empirical work. The other four test scores are used together in the analysis below as an alternative to AFQT score. Taber (2001) contends that these four scores represent a parsimonious

[15] Thus our data include all observations from the random sample and the random black and Hispanic oversamples. The military and the nonblack, non-Hispanic disadvantaged samples are small relative to the data we include. In addition, the military oversample was discontinued after the 1984 wave of the NLSY.

[16] Starting the panel at an earlier age creates a data set unrepresentative of the population since it would contain more annual observations for high school dropouts and graduates than for college graduates. The estimates presented below are not sensitive to higher age cutoffs.

TABLE 1
DETAILS AND SUMMARY STATISTICS OF THE PRIMARY VARIABLES

Variable	Mean	Standard Deviation	Sample Size
	A. Schooling and Minority Status		
Years of school[a]	12.8	2.5	2,404
Black	.31	.46	2,404
Hispanic	.19	.39	2,404
	B. Family Background Variables[b]		
Highest grade father	10.50	4.10	2,404
Highest grade mother	10.78	3.21	2,404
Number of siblings	3.73	2.63	2,404
Broken home	.24	.42	2,404
Family income[c] (1999 dollars)	$39,601	2,444	2,404
	C. Geographic Controls (Measured at Age 17)		
Urban residence	.79	.40	2,404
Residence in Northeast	.20	.40	2,404
Residence in South	.36	.48	2,404
Residence in West	.19	.39	2,404
Residence in North Central	.25	.43	2,404
	D. Test Score Variables[d]		
AFQT score	61.1	22.3	2,404
Math score	11.9	6.3	2,404
Word score	21.9	8.5	2,404
Science score	14.2	5.4	2,404
Automotive score	13.4	5.5	2,404
	E. Instruments for Schooling		
Local college[e]	.87	.34	2,404
Local earnings at age 17 (1999 dollars)[f]	$22,620	4,209	2,404
Mean local earnings over working life (1999 dollars)[f]	$23,073	4,333	2,404
	F. Time-Varying Labor Market Variables[g]		
Hourly wage (1999 dollars)	$10.01	3.479	13,762
Current local earnings (1999 dollars)	$24,519	6,122	13,762
Work experience (in years)	7.536	3.169	13,762

SOURCE.—The National Longitudinal Survey of Youth. County average earnings is taken from the Bureau of Economic Analysis's Regional Economic Information System data and constructed from annual average earnings in retail, wholesale, and construction industries. Information on college location and college costs is taken from the Department of Education's HEGIS and IPEDS surveys of institutional characteristics.

[a] Years of schooling completed by the 1994 interview date.

[b] Except for family income, these variables are measured at age 14. Broken home is a binary variable indicating whether a person was living with both biological parents.

[c] Parental family income is available annually only for individuals who are dependents of their parents. Our measure takes parental family income as close to age 17 as possible. For about 16 percent of our sample, the measure was unavailable at age 17 and was taken at age 16 or 18.

[d] Test scores are taken from the ASVAB tests; see the text for details.

[e] An indicator variable for presence of a college (two- or four-year) in the county of residence at age 17.

[f] Average earnings in the county in which an individual resides at age 17 among persons employed in industries predominated by unskilled workers. Mean local earnings over working life is the earnings rate averaged from 1980 through 1996 in the county in which an individual resides at age 17. See the text for more details.

[g] These are annual measures. Current local earnings is an annual measure of the same data series as local earnings at age 17 and is taken in the county in which an individual resided at age 17. Hourly wage is hourly earnings at an individual's current or most recent job at the time of each annual interview. Work experience is defined as age $-$ schooling $-$ 6.

set of wage predictors that capture more dimensions of ability than AFQT by itself.

Panel E of the table summarizes the instruments used below for endogenous schooling. Local college is a binary indicator for the presence of any college (either two-year or four-year) in the county of residence at age 17 (or age 16 for the handful who graduate from high school by age 17).[17] College identifiers were merged to NLSY annual county of residence measures from the Department of Education's annual Higher Education General Information Survey (HEGIS) and the Integrated Postsecondary Education Data System's (IPEDS) Institutional Characteristics surveys, which contain annual data on location, type of institution, tuition, and other variables associated with colleges in the United States.[18]

Local labor market conditions were created from annual county-level labor market data on average annual earnings in industries dominated by unskilled workers.[19] Our proxy for forgone earnings is a static variable created from this series taken in a person's county of residence at age 17 (local earnings at age 17 in panel E of the table). The variable mean local earnings over working life is constructed in the following manner. We condition on the county in which a student lived at age 17. We then average the annually varying labor market measure of earnings for this county for the years during which this person works (starting at age 22) and appears in the sample (up to age 33). We also studied the local (county) unemployment rate in our schooling choice analysis, but results using this variable were very imprecise.

Panel F shows means of annually varying variables used in the wage analysis. Hourly wage is the wage at the current or most recently held job as of the interview date in each year. Current local earnings is an annual measure constructed from the local annual earnings series just discussed. It is measured for the county in which a person currently lives and is discussed further below. Work experience is a measure of potential work experience and is constructed as age minus schooling minus six.

---

[17] We also explored a set of three indicator variables for two-year college, four-year college, and both. Our conclusions below were unchanged, though standard errors of estimates improved slightly in some cases when the three indicators were used together.

[18] A number of specialty colleges, generally with enrollments less than 100, and federal institutions, such as the Naval Academy, were excluded.

[19] The data are taken from annual Bureau of Economic Analysis data. Since the bureau's data are reported by industry and not by occupation, we use average earnings in service, agriculture, and the wholesale and retail trade industries. A number of other labor market measures using a variety of industry aggregates were explored, with little difference to the estimates presented below. This was true apparently because, except for government jobs, average wages across a variety of industries in the same local labor market exhibit a high degree of correlation.

V.   Empirical Results from Wage and Schooling Choice Models

A.   *Methodological Issues behind Instrumental Variables*

This section begins with a discussion of methodological issues underlying the application of instrumental variables to the estimation of returns to schooling. Empirical results are reported in Sections V*B*–V*D*.

Consider the following regression equation in which i enumerates individuals and t is an indicator of time:

$$\log (w_{it}) = \beta_0 + S_i \gamma_i + l_{it}\beta_1 + e_{it}\beta_2 + e_{it}^2\beta_3 + \mathbf{X}'_{it}\boldsymbol{\beta}_4 + u_{it}, \qquad (17)$$

where w_{it} represents hourly wage, S_i is years of schooling, l_{it} denotes prevailing wages in the local labor market in which individual i lives at time t, e_{it} is work experience, $\mathbf{X}_{it}$ represents other factors affecting wages, and u_{it} is an error term. To facilitate the discussion below, variables denoting the prevailing wage rate in the local labor market (l_{it}) and work experience (e_{it} and its square) have been separated from other terms in $\mathbf{X}_{it}$.

The returns to schooling literature has sought unbiased estimates of the coefficient on schooling, γ_i. Willis (1986) presents conditions under which this parameter can be interpreted as the internal rate of return to schooling. Two aspects of (17) complicate estimation of schooling returns. The first is the much-studied problem of ability bias: u_{it} may be correlated with S_i as a result of unobserved ability. The second arises because we allow heterogeneous returns to schooling: the coefficient on schooling, γ_i, depends on i and varies across members of the population. If γ_i were homogeneous and u_{it} were orthogonal to S_i, OLS would produce consistent estimates of the causal effect of a year of schooling on wages. We elaborate on these two complications next.

1.   Complications Concerning Instruments for Endogenous Schooling

First, to simplify the discussion, assume that γ_i is a constant parameter γ. It is well understood that a consistent estimate of γ can be recovered by two-stage least squares using instruments correlated with S_i but uncorrelated with u_{it}. The behavioral model of Section III describes two college cost measures that make potential instruments. The first is direct costs of schooling, represented by τ_1 in equation (8). As mentioned above, we follow Card (1995*b*) by using presence of a college in the county of residence at age 17 as a measure of an important component of direct college costs and assume that it is uncorrelated with unobserved ability. For students from families with low and moderate incomes, the opportunity to live at home or have the parental residence close at hand while in college yields a substantial financial advantage. The NLSY data

reveal that the probability of living at home while in college is about 55 percent for students with a college in their county of residence at age 17 and 34 percent for others.

The second cost of college is the opportunity cost, denoted in the term W_0 of equation (7). The prevailing wage rate in the county in which a person lived at age 17 (the variable local earnings at age 17 in table 1) is a candidate instrument since it should be correlated with W_0 but uncorrelated with unobserved ability. The major concern in using this variable as an instrument is that labor market variables at age 17 are almost certainly correlated with local labor market variables later in life and hence correlated with current-period wages. To address this concern, we include directly in the wage regression a time-varying measure of local average earnings in the current county of residence. This variable is called current local earnings in table 1 and is denoted by l_{it} in equation (17). The crucial assumption justifying the instrument is that, conditional on l_{it}, local labor market conditions at age 17 are unrelated to the error term u_{it} in (17).

Including l_{it} directly in the wage function (17) leads to a new concern. Migration is potentially endogenous and related to schooling outcomes. For instance, college graduates may move more readily to better local labor markets. Since l_{it} is measured in the county in which a person resides at time t, it may be endogenous to schooling choice. Consistent estimation of the coefficient β_1 associated with l_{it} in equation (17) requires an additional instrument correlated with l_{it} but not itself endogenous. The natural choice in this case is the prevailing wage rate at time t measured in the county in which person i lived at age 17. Since many individuals do not stray far from their county of residence at age 17, this instrument correlates strongly with prevailing earnings in the current county of residence. In addition, since its value is determined by the county in which a student lived at age 17, it does not depend on residential location decisions made after schooling completion.

A last concern, often mentioned in the returns to schooling literature, is endogeneity of experience. Our experience variable is a measure of potential experience and is equal to age minus schooling minus six. If schooling is endogenous and potential experience depends directly on schooling, potential experience would be endogenous as well. To account for this problem, we follow the literature by instrumenting for experience and experience squared using age and age squared in some specifications below.

2.   Complications Concerning Random Effects

The second complication in our analysis arises because we allow returns to schooling, γ_i, to vary across individuals. If returns were constant and

were given by parameter γ, then either of the instruments based on schooling costs would yield consistent estimates of γ. However, when γ_i enters (17) as a random effect, two-stage least-squares estimates converge to values that depend on the instrument used. Imbens and Angrist (1994) develop limiting values of instrumental variable estimators when a random coefficient is associated with an endogenous variable. In the context here, instrumental variable estimation recovers an average return to schooling for those in the subset of the population induced to change their level of schooling when the cost of schooling changes.[20] Section III establishes that schooling choices of borrowing-constrained persons are more sensitive to changes in direct costs than to changes in opportunity costs. Individuals who are not credit-constrained respond the same to changes of equal magnitude in either cost (propositions 1 and 2). Hence, the composition of those induced to alter their schooling status depends on which cost variable is used as an instrument. When direct cost is used, instrumental variable estimation puts more weight on the returns earned by borrowing-constrained students than when opportunity cost is used. In addition, proposition 3 shows that a credit-restricted student on the margin between going and not going to college requires a higher return to schooling (higher γ_i) than an identical student with access to cheap financing. Together, these arguments imply that instrumental variable estimates of the returns to schooling should be higher when schooling is instrumented with direct costs than estimates recovered when opportunity costs are employed as the instrument. In the language of Lang (1993), the "discount rate bias" is higher when direct costs are the instrument for schooling than when opportunity costs are used.

The instrumental variable analysis reported in Sections VB and VC uses differences in the magnitude of discount rate bias to test for the presence of borrowing constraints in the data. Note that we have said nothing about OLS estimates. If γ_i were a constant as is generally assumed in the literature, OLS estimates would be biased. In our framework, as instrumental variable estimates may contain discount rate bias, our theory predicts nothing about the relationship between instrumental variable and OLS estimates. However, we present OLS results below to facilitate comparison with other estimates reported in the literature.

It is important to point out that while this argument may be intuitively appealing, it is not precise for a number of reasons. First, the actual estimated coefficient on schooling depends on the full joint distribution of schooling variables and regressors for individuals close to the margin.

---

[20] In the language of Imbens and Angrist (1994), instrumental variable estimators converge to the expected treatment effect for those individuals induced to change status by a change in the instrument. In our model the treatment is school and the treatment effect is the return to schooling.

Second, schooling is a choice among more than two options, so the Imbens and Angrist (1994) results do not apply directly. Instead, instrumental variable estimates will be a weighted average of returns for marginal decision makers at each level of schooling. The weights in the average will likely depend on the instrument used. We strongly expect that if borrowing constraints are important, the coefficient on schooling will be higher when direct costs are the instrument, but we have not formally proved this result. A major advantage of the structural econometric model developed below is that identification of borrowing constraints is formally justified.

B.   First-Stage Results

This subsection presents first-stage estimates to demonstrate that the instruments have predictive power in the first stage and that their signs are consistent with the model presented above. Because individuals appear in our data only after leaving school, the variable measuring schooling attainment is constant. This makes for difficult interpretation of a first-stage panel regression of schooling on time-varying local labor market variables and other characteristics. However, to convey the content of the first-stage regressions, we construct the variable mean local earnings over working life, which takes the mean value of l_{it} over the years in which the individual is included in the wage regressions. We regress schooling on this variable, local earnings at age 17, an indicator for the presence of a college in the county (local college), and a number of other control variables.

   The first row in column 1 of table 2 shows that local college has a large and statistically significant effect, implying that individuals with a college in their county complete almost one-half year more of school, on average. The other covariates have estimates with signs and magnitudes consistent with those reported in other work (see, e.g., Cameron and Heckman 2001).

   Column 2 reports estimates when local earnings at age 17 is included in the regression instead of local college (row 2). The estimated coefficient has the expected negative sign but is not statistically significant at conventional levels. This variable apparently reflects both time-series variation in county earnings due to business cycle effects and cross-sectional differences in average earnings and wealth across counties. Adding mean local earnings over working life to control for levels of wealth across counties allows us to sort out these two avenues of influence. The third row of column 3 shows that the coefficient on the new variable is significant and positive, indicating that students from wealthier counties are more likely to attend college—perhaps as a result of superior schools or peer effects. In addition, the coefficient on local

TABLE 2
FIRST STAGE REGRESSION OF YEARS OF SCHOOLING ($N=2{,}404$)

	REGRESSION			
VARIABLE	(1)	(2)	(3)	(4)
Local college	.417			.435
	(.152)			(.148)
Local earnings at 17/1,000		−.023	−.182	−.183
		(.021)	(.074)	(.074)
Mean local earnings over working life/1,000			.130	.123
			(.058)	(.058)
Black	.688	.723	.691	.677
	(.124)	(.124)	(.123)	(.124)
Hispanic	.334	.359	.348	.345
	(.138)	(.143)	(.141)	(.141)
Math score	.167	.168	.167	.166
	(.010)	(.010)	(.010)	(.010)
Word score	.050	.050	.050	.050
	(.010)	(.009)	(.010)	(.010)
Science score	.065	.066	.066	.065
	(.015)	(.015)	(.015)	(.015)
Automotive score	−.065	−.069	−.065	−.066
	(.011)	(.010)	(.011)	(.011)
Highest grade father	.051	.053	.052	.050
	(.016)	(.016)	(.016)	(.016)
Highest grade mother	.038	.039	.039	.039
	(.022)	(.022)	(.022)	(.022)
Number of siblings	−.044	−.047	−.047	−.046
	(.016)	(.016)	(.016)	(.016)
Family income/10,000	.090	.094	.095	.095
	(.022)	(.022)	(.022)	(.022)
Intercept	11.320	11.710	11.993	11.886
	(.273)	(.335)	(.371)	(.321)
Geographic controls[a]	yes	yes	yes	yes
Cohort controls[b]	yes	yes	yes	yes

NOTE.—The dependent variable is years of schooling as defined in table 1. Standard errors are in parentheses. Because local college and the local labor market variables are measured at the county level, the standard errors used in this paper allow for arbitrary correlation among individuals living in the same county at age 17 (i.e., clustering).

[a] Geographic controls include an indicator variable for urban residence at age 17 and indicators for residence in the four standard census regions at age 17. See table 1.

[b] Cohort controls are a set of four indicator variables for age, which ranges from 13 to 16 in 1978.

earnings at age 17, which proxies opportunity costs of school, is negative as expected and statistically significant. Column 4 shows that adding local college to the column 3 specification hardly alters any of the estimates just reported.

Two possible mechanisms explain why the opportunity cost variable reported in the second row has a negative sign. First, students could be more likely to attend school during temporary downturns arising from business cycles. Second, individuals living in counties in which the economy is on a long-run trend toward improvement could be more likely

to attend college. In practice, both avenues seem to have influence.[21] From the point of view of the theory, the distinction is irrelevant. What is important is the comparison of economic conditions during a student's college years and conditions later in life. However, the business cycle influence is more intuitively appealing as a source of identification since it does not depend on long-run features of the county of residence, which may be correlated with individual ability outcomes. Unfortunately, we cannot separate out these different sources without losing substantial statistical power. Evidence reported in Section V*E* suggests that correlation between ability and location may not be an important concern. We show that the time-varying local labor market variable is unrelated to observed measures of ability, so it seems plausible that labor market conditions at age 17 are not related to unobserved ability differences either.

C.   *Instrumental Variables Excluding Forgone Earnings*

Table 3 displays estimated returns to schooling from six different log wage specifications when schooling is instrumented with local earnings at age 17. This instrument, which proxies opportunity costs of school, is not expected to have a differentially large impact on borrowing-constrained individuals, so we expect little "discount rate bias." The specification also includes current local earnings, dummy variables for racial-ethnic identity, experience and experience squared, family background variables, our set of four test score variables, and cohort and geographic controls (see the notes to table 2 for variable definitions). Four test score and four family background variables are included to capture as many dimensions of unobserved ability as the data allow. Current local earnings is instrumented with current earnings in age 17 county as discussed above. As a practical matter, not instrumenting for current local earnings made little difference to the results shown below, so we do not dwell on this point further.

  The estimated schooling return is shown in the first row of each column. Column 1 reports the OLS results. Column 2 shows that the instrumental variable point estimate of the schooling coefficient is about 50 percent larger than the OLS estimate. Instrumenting experience and experience squared with age and age squared leaves the results essentially unchanged (col. 3). In interpreting these findings, one should keep in mind that the standard errors are large enough that we cannot

TABLE 3

Ordinary Least Squares (OLS) and Instrumental Variable (IV) Estimates of Log Hourly Wage Function Using Forgone Earnings as the Instrument for Schooling

	OLS (1)	IV1 (2)	IV2 (3)	OLS (4)	IV1 (5)	IV2 (6)
Schooling	.058	.083	.110	.074	.107	.134
	(.004)	(.042)	(.086)	(.004)	(.034)	(.061)
Current local earnings/1,000	.027	.035	.036	.025	.031	.034
	(.003)	(.006)	(.005)	(.003)	(.005)	(.005)
Work experience	.054	.065	.091	.055	.074	.083
	(.006)	(.018)	(.023)	(.006)	(.019)	(.022)
Experience squared	−.002	−.002	−.004	−.002	−.002	−.004
	(.000)	(.000)	(.001)	(.000)	(.001)	(.001)
Black	−.063	−.085	−.115	−.148	−.162	−.178
	(.024)	(.031)	(.061)	(.022)	(.022)	(.028)
Hispanic	−.020	−.033	−.041	−.061	−.074	−.085
	(.030)	(.029)	(.034)	(.030)	(.029)	(.032)
Highest grade father	−.003	−.005	−.008	−.001	−.005	−.013
	(.004)	(.005)	(.010)	(.004)	(.006)	(.014)
Highest grade mother	−.005	−.008	−.012	−.001	−.007	−.015
	(.005)	(.006)	(.011)	(.005)	(.007)	(.014)
Number of siblings	.002	.003	.005	−.001	.001	.005
	(.003)	(.004)	(.005)	(.005)	(.004)	(.007)
Family income/10,000	.036	.034	.031	.042	.039	.034
	(.005)	(.005)	(.007)	(.005)	(.006)	(.005)
Math score	.010	.007	.002			
	(.002)	(.005)	(.014)			
Word score	.002	.001	−.000			
	(.002)	(.002)	(.004)			
Science score	−.004	−.005	−.006			
	(.003)	(.003)	(.006)			
Automotive score	.012	.014	.016			
	(.002)	(.003)	(.007)			
Geographic controls	yes	yes	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes	yes	yes
Instrument for schooling: local earnings at 17	NA	yes	yes	NA	yes	yes
Instrument for current local earnings: current earnings in age 17 county	NA	yes	yes	NA	yes	yes
Instruments for work experience and work experience squared: age and age squared	NA	no	yes	NA	no	yes
Number of individuals	2,225	2,225	2,225	2,225	2,225	2,225
Total observations	13,762	13,762	13,762	13,762	13,762	13,762

Note.—The dependent variable is log hourly wage. The IV1 specification treats experience and experience squared as exogenous; IV2 instruments them with age and age squared. Current local earnings is instrumented with current local earnings in age 17 county in all instrumental variable specifications. Standard errors (in parentheses) are robust with respect to arbitrary correlation across time for observations on the same person and across persons who live in the same county at age 17 (i.e., clustering). See the note to table 2. NA means not applicable. Geographic and cohort controls are defined in the note to table 2.

reject the hypothesis that OLS estimates and instrumental variable estimates are the same.

Columns 4–6 exhibit estimates of the same specifications except that test scores and family income are excluded. The OLS and instrumental variable point estimates of the schooling coefficient are both higher, but the instrumental variable estimate remains larger than its OLS counterpart. Other specifications not reported here, including a more standard one with AFQT score instead of the set of four test scores, all yield similar patterns. These results are also similar to those of Arkes (1998), who uses state unemployment rates in a similar design and finds instrumental variable estimates higher than OLS estimates.

A potential problem with this specification is that it does not account for influences of economic downturns on schooling choices that operate through family income or through the decision to work while in college. Borrowing-constrained families may find raising funds for college more difficult during recessions. This possibility can reverse the direction of labor market effects: schooling may increase during booms for children whose parents are borrowing-constrained. Thus the influence of local labor market conditions on schooling attendance is no longer monotonic. Intuitively, the income and substitution effects of changes in local labor market conditions go in opposite directions for borrowing-constrained families, and it is not clear which effect dominates. However, if families are not borrowing-constrained, there is no income effect on schooling (as in a standard human capital model). Thus it is possible that borrowing-constrained families send their children to college at higher rates during a boom, whereas non-borrowing-constrained families send their children at higher rates during a bust when forgone earnings are low.

The negative association between county earnings and schooling reported in table 2 supports the dominance of forgone earnings on schooling decisions. However, this makes the instrumental variable results in table 3 even more surprising. If children of borrowing-constrained families have higher marginal returns to schooling and decrease schooling during recessions, the discount rate bias should be negative. Our empirical findings are counterintuitive if credit constraints are important. The mechanics of this argument are formalized in section A of our technical appendix (available at http://www.econ.northwestern.edu/faculty/taber).

The case in which students work while in school is analogous. Students who are borrowing-constrained would presumably be more likely to work in college. Thus the effect of local labor market conditions would likely have a larger effect on individuals who are not borrowing-constrained since they depend less on opportunities to work during college in order

to attend. Once again, this argument makes the empirical results more surprising.

D.   *Instrumental Variables Excluding Direct Costs of Schooling*

Estimates of analogous specifications using presence of a local college as the instrument for schooling are presented in panel A of table 4. Columns 1–3 present results with no controls for current local earnings in the regression. The OLS results are shown in column 1, instrumental variable results in column 2, and instrumental variable results from a specification that includes instruments for work experience and its square appear in column 3. The estimated schooling effect is reported in the first row of the table. The instrumental variable specifications in columns 2 and 3 reveal a large, causal effect of schooling: over 300 percent larger than the OLS point estimate.

The use of this instrument leads to a concern that colleges are not randomly assigned to counties in the United States. Thus it makes sense to control for county-level earnings directly in the specification and instrument for this variable with current earnings in age 17 county, following our discussion above regarding possible endogeneity of this variable. Essentially, this addition to the specification controls for differences in wealth among counties. To our knowledge, this strategy has not been tried in other studies that instrument schooling with a variable for local college access (see the survey in Card [1999]). This control turns out to be important. Adding current local earnings yields a strikingly different pattern: instrumental variable and OLS estimates are nearly identical (cols. 4–6).

Panel B of table 4 confirms these findings using the same specifications reported in panel A with the exception that the AFQT score replaces the four test scores as a control for ability in columns 1–3 and controls for ability are dropped altogether for the results shown in columns 4–6. These specifications are included since they correspond closely to those commonly found in the literature. The OLS results are shown with AFQT score in column 1 and with no ability measure in column 3. When schooling and current local earnings are instrumented together (cols. 2 and 5) and when schooling, current local earnings, experience, and experience squared are all instrumented (cols. 3 and 6), the main finding is stronger than suggested by panel A: Instrumental variable estimates of the schooling coefficient are well below their OLS counterparts.

This subsection contrasts instrumental variable estimates using direct costs (table 3) with instrumental variable estimates using opportunity costs (table 4). Theory predicts that if borrowing constraints are important for schooling decisions, instrumental variable estimates using

direct costs will be higher than instrumental variable estimates using opportunity costs. The results in these tables do not bear out that prediction. We find no evidence for discount rate bias using this approach.

E.   Validity of the Instruments

In general, without a maintained assumption that one of the instruments is valid, it is impossible to test the validity of any of them. In addition, since the causal effect of schooling is assumed to vary across individuals, standard overidentification tests do not apply. Nevertheless, it is still worthwhile to investigate concerns about correlation between instruments and the omitted ability component of wage equation residuals. Finding no relationship between observable measures of ability and the instruments cannot prove the validity of the instruments since correlation with unobserved ability components may remain. Nevertheless, such a finding still lends credence to their use.

   Column 1 of table 5 shows the estimated relationship between the indicator variable for presence of a local college and a standard set of regressors: local earnings at 17, mean local earnings over the working life, racial-ethnic indicators, the set of four test scores, and four family background variables. The pattern of estimates is not encouraging for the use of the local college variable as an instrument: math score and automotive core are significantly related to presence of a college in the county. If local college were correlated with unobserved ability in the same way as these observed measures, then instrumental variable schooling estimates would be biased upward.[22] Nevertheless, this correlation makes the instrumental variable results presented above even more surprising since it implies that the instrumental variable estimates should be lower than they actually are.

   By contrast, the fact that test scores do not predict local earnings at age 17 (col. 2) supports using local labor market variation as an instrument. Column 2 also shows that the only variable other than mean local earnings over the working life with any significance is family income. Given that family income is measured when students were close to age 17, this result is not particularly astounding.[23]

   The discovery that test scores predict presence of a local college is not necessarily a problem since test score variables are included in the specifications. Table 6 displays results that are more favorable to the legitimacy of using local college as an instrument. The table displays

---

[22] Altonji, Elder, and Taber (2002) provide a model that formally justifies this type of argument. We do not have enough control variables in this case to use the type of formal analysis they suggest.

[23] Along the same lines, one may question whether family income is a valid control. The empirical results here are robust to inclusion of family income.

TABLE 4
OLS and IV Estimates of the Log Hourly Wage Function with Presence of
Local College as an Instrument

	OLS (1)	IV1 (2)	IV2 (3)	OLS (4)	IV1 (5)	IV2 (6)
	A. Test Scores as Ability Controls					
Schooling	.062	.228	.193	.058	.057	.061
	(.004)	(.109)	(.084)	(.004)	(.115)	(.076)
Current local earnings/1,000				.027	.035	.035
				(.002)	(.006)	(.005)
Experience	.054	.124	.075	.054	.054	.102
	(.006)	(.047)	(.020)	(.006)	(.049)	(.019)
Experience squared	−.002	−.003	−.004	−.002	−.002	−.005
	(.000)	(.001)	(.001)	(.000)	(.001)	(.001)
Black	−.029	−.111	−.119	−.063	−.073	−.082
	(.026)	(.066)	(.067)	(.023)	(.063)	(.059)
Hispanic	.009	−.023	−.020	−.020	−.028	−.029
	(.035)	(.045)	(.044)	(.026)	(.037)	(.036)
Highest grade father	−.003	−.014	−.015	−.003	−.003	−.003
	(.004)	(.009)	(.010)	(.004)	(.008)	(.008)
Highest grade mother	−.002	−.016	−.017	−.005	−.006	−.006
	(.005)	(.010)	(.011)	(.005)	(.010)	(.010)
Number of siblings	−.001	.003	.005	.002	.003	.003
	(.004)	(.005)	(.005)	(.003)	(.005)	(.005)
Family income	.039	.033	.032	.036	.035	.034
	(.005)	(.007)	(.007)	(.005)	(.006)	(.007)
Math score	.010	−.008	−.010	.010	.010	.010
	(.002)	(.013)	(.014)	(.002)	(.013)	(.013)
Word score	.003	−.001	−.002	.002	.002	.002
	(.002)	(.004)	(.004)	(.002)	(.003)	(.003)
Science score	−.005	−.012	−.013	−.004	−.003	−.003
	(.003)	(.006)	(.006)	(.003)	(.006)	(.005)
Automotive Score	.009	.017	.018	.012	.012	.012
	(.002)	(.005)	(.006)	(.002)	(.005)	(.005)
Geographic controls	yes	yes	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes	yes	yes
Instruments for schooling: college in county	NA	yes	yes	NA	yes	yes
Instruments for current local earnings: current earnings in age 17 county	NA	no	no	NA	yes	yes
Instruments for work experience and work experience squared: age and age squared	NA	no	yes	NA	no	yes
	B. Alternative Ability Controls					
Schooling	.054	.027	.042	.074	.052	.065
	(.004)	(.095)	(.060)	(.004)	(.078)	(.046)
Current local earnings/1,000	.026	.035	.034	.025	.031	.031
	(.002)	(.006)	(.006)	(.002)	(.006)	(.006)
Experience	.054	.042	.106	.055	.043	.104
	(.006)	(.046)	(.017)	(.006)	(.044)	(.016)
Experience squared	−.002	−.002	−.005	−.002	−.002	−.005
	(.000)	(.001)	(.001)	(.000)	(.001)	(.001)
Black	−.081	−.077	−.087	−.147	−.154	−.160
	(.021)	(.066)	(.062)	(.021)	(.026)	(.026)

TABLE 4
(*Continued*)

	OLS (1)	IV1 (2)	IV2 (3)	OLS (4)	IV1 (5)	IV2 (6)
Hispanic	−.040	−.043	−.043	−.061	−.065	−.064
	(.026)	(.040)	(.040)	(.026)	(.032)	(.033)
Highest grade father	−.004	−.002	−.001	−.001	.001	.002
	(.004)	(.009)	(.009)	(.004)	(.010)	(.010)
Highest grade mother	−.006	−.005	−.005	−.001	.001	.000
	(.005)	(.008)	(.008)	(.005)	(.011)	(.010)
Number of siblings	.002	.003	.003	−.001	−.001	−.001
	(.003)	(.004)	(.004)	(.003)	(.006)	(.006)
Family income/10,000	.036	.036	.036	.042	.044	.043
	(.005)	(.006)	(.006)	(.005)	(.009)	(.009)
AFQT score	.005	.006	.006			
	(.001)	(.004)	(.003)			
Geographic controls	yes	yes	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes	yes	yes
Instrument for schooling: local college	NA	yes	yes	NA	yes	yes
Instrument for current local earnings: current earnings in age 17 county	NA	yes	yes	NA	yes	yes
Instruments for work experience and work experience squared: age and age squared	NA	no	yes	NA	no	yes
Number of individuals	2,225	2,225	2,225	2,225	2,225	2,225
Total wage observations	13,762	13,762	13,762	13,762	13,762	13,762

NOTE.—The dependent variable is log hourly wage. The IV1 specification treats experience and experience squared as exogenous; IV2 instruments them with age and age squared. Current local earnings is instrumented with current local earnings in age 17 county in cols. 4–6. Standard errors (in parentheses) are robust with respect to arbitrary correlation across time for observations on the same person and across persons who live in the same county at age 17 (i.e., clustering). Geographic and cohort controls are defined in the notes to table 2. NA means not applicable.

estimates of a probit analysis of indicators for college attendance and high school dropping out on racial-ethnic identifiers, the four test scores, the four family background variables, cohort and geographic controls, and the instruments. To have any faith in the exogeneity of local college, the variable should have a stronger effect on college attendance than high school dropping out.[24] However, if local college picks up a community's "pro-schooling" character rather than the reduction in the total cost of college, the variable should predict high school graduation as well. Indeed, as table 6 shows, local college has a large and statistically significant influence on college attendance but a smaller and statistically insignificant effect on high school graduation, lending support to the validity of this variable as an instrument.

[24] The variable may have an indirect effect on dropping out of high school. If it increases the value of college, it must increase the value of high school graduation because high school graduates have the option of attending college.

TABLE 5
REGRESSIONS OF DIRECT AND INDIRECT COST INSTRUMENTS ON SCHOOLING AND WAGE
DETERMINANTS ($N=2{,}404$)

	DEPENDENT VARIABLE	
COVARIATE	Local College (1)	Local Earnings at Age 17 (2)
College in county		.038
		(.091)
Local earnings at 17/1,000	.006	
	(.014)	
Mean local earnings over working life/	.013	.748
10,000	(.011)	(.017)
Black	.032	−.094
	(.035)	(.070)
Hispanic	.005	−.007
	(.037)	(.078)
Math score/10	.030	−.052
	(.016)	(.033)
Word score/10	.014	.022
	(.014)	(.054)
Science score/10	.023	.022
	(.024)	(.054)
Automotive score	−.053	−.006
	(.017)	(.043)
Highest grade father	.006	−.004
	(.002)	(.007)
Highest grade mother	−.001	−.001
	(.003)	(.007)
Number of siblings	−.002	.007
	(.003)	(.007)
Family income/10,000	−.001	.035
	(.003)	(.012)
Geographic controls	yes	yes
Cohort controls	yes	yes

NOTE.—See table 2 for variable definitions. Standard errors (in parentheses) are robust with respect to arbitrary correlation across persons who live in the same county at age 17 (i.e., clustering). See the note to table 2. Geographic and cohort controls are defined in the notes to table 2.

F. Interactions between Observables and Direct Costs

The identification argument sketched in Section III emphasizes the excess sensitivity of credit-constrained individuals to direct schooling costs. Following the logic of that argument, we could quantify the differential effects of college costs and college access if credit constraints were observable. Since credit constraints are not observable, we estimate interactions between local college and variables that we might expect to be related to credit availability and schooling attainment.

Table 7 exhibits estimates of years of schooling regressed on our standard set of covariates as well as five sets of variables interacted with the presence of a local college at age 17. In each column, the interactions are displayed after the main effect estimates. Column 1 shows

TABLE 6
DETERMINANTS OF COLLEGE ATTENDANCE AND HIGH SCHOOL DROPPING OUT: AVERAGE
DERIVATIVES FROM PROBIT MODELS (N=2,404)

	DEPENDENT VARIABLE	
COVARIATE	College Attendance (1)	High School Dropping Out (2)
College in county	.166	−.035
	(.036)	(.030)
Local earnings at 17	−.043	.028
	(.022)	(.016)
Mean local earnings over working life	.035	−.009
	(.017)	(.013)
Black	.170	−.109
	(.037)	(.023)
Hispanic	.144	−.034
	(.041)	(.024)
Math score	.032	−.020
	(.003)	(.002)
Word score	.017	−.007
	(.003)	(.002)
Science score	.010	−.006
	(.004)	(.003)
Automotive score	−.013	.002
	(.003)	(.002)
Highest grade father	.013	−.003
	(.004)	(.003)
Highest grade mother	.010	−.004
	(.007)	(.004)
Number of siblings	−.016	.004
	(.005)	(.003)
Family income	.011	−.023
	(.007)	(.005)
Geographic controls	yes	yes
Cohort controls	yes	yes

NOTE.—Average derivatives are formed by calculating the derivative for each individual in the sample with respect to each right-hand-side variable and then averaging. Standard errors (in parentheses) of the probit estimates are robust with respect to arbitrary clustering correlation across persons who live in the same county at age 17. Geographic and cohort controls are defined in the notes to table 2.

interactions with black and Hispanic indicator variables. For a given income level, it is well documented that minorities have less wealth than whites and thus may be credit-constrained relative to whites. We find the opposite: the interactions are negative, showing that the presence of a college seems to be more important for whites than for minorities (although the estimates are statistically insignificant).

Column 2 shows that the interaction with father's education is positive but the interaction with mother's education is negative. Again, this pattern is hard to interpret as evidence of credit constraints. Columns 3 and 4 display the interaction with family income and number of siblings. Again, the interactions are insignificant and have the wrong sign. For

TABLE 7

REGRESSION OF YEARS OF SCHOOLING ON PRESENCE OF LOCAL COLLEGE INTERACTED
WITH ALTERNATIVE COVARIATES ($N=2,404$)

	REGRESSION				
	(1)	(2)	(3)	(4)	(5)
Main effects:					
College in county	.596	3.149	.103	.592	.491
	(.166)	(.072)	(.236)	(.229)	(.15)
Local earnings at 17	−.102	−.074	−.107	−.104	−.109
	(.059)	(.038)	(.059)	(.058)	(.058)
Mean local earnings over working	.067	.038	.068	.065	.070
life	(.046)	(.031)	(.046)	(.046)	(.045)
Black	.968	.333	.690	.689	.709
	(.211)	(.086)	(.125)	(.125)	(.125)
Hispanic	.914	.018	.347	.344	.361
	(.379)	(.105)	(.142)	(.142)	(.142)
Math score	.165	.077	.166	.166	.104
	(.010)	(.006)	(.010)	(.010)	(.03)
Word score	.050	.013	.050	.050	.057
	(.010)	(.007)	(.010)	(.010)	(.022)
Science score	.067	.042	.065	.066	.065
	(.015)	(.012)	(.015)	(.015)	(.015)
Automotive score	−.067	−.031	−.066	−.066	−.066
	(.011)	(.007)	(.011)	(.011)	(.011)
Highest grade father	.051	−.040	.051	.051	.052
	(.016)	(.028)	(.016)	(.016)	(.016)
Highest grade mother	.038	.074	.038	.038	.038
	(.022)	(.053)	(.022)	(.022)	(.022)
Number of siblings	−.046	−.013	−.046	−.012	−.046
	(.016)	(.011)	(.016)	(.038)	(.016)
Family income	.092	.068	−.008	.095	.093
	(.022)	(.017)	(.068)	(.022)	(.022)
Interactions:					
Black × local college	−.330				
	(.220)				
Hispanic × local college	−.635				
	(.390)				
Highest grade father × local college		.071			
		(.030)			
Highest grade mother × local college		−.065			
		(.055)			
Family income × local college			.111		
			(.067)		
Number of siblings × local college				−.039	
				(.041)	
Math score × local college					.069
					(.031)
Word score × local college					−.007
					(.021)
Geographic controls	yes	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes	yes

NOTE.—Standard errors (in parentheses) are robust with respect to arbitrary correlation across persons who live in the same county at age 17. See the note to table 2. Geographic and cohort controls are defined in the notes to table 2.

completeness, column 5 presents interactions with math score and word score.[25] This exercise produces no evidence of credit constraints.

In related work, both Card (1995*b*) and Kling (2001) examine similar sets of interactions in the National Longitudinal Survey of Young Men (NLS Young Men). They find evidence that individuals from low family backgrounds are affected more by the presence of a college. Kling also compares estimates of interactions on NLSY data and finds the NLSY estimates to be smaller in magnitude and not statistically significant compared to estimates from NLS Young Men. He reaches essentially the same conclusion that we do.

A number of factors may explain why the NLSY results differ from those using the NLS Young Men. Chiefly, the NLSY data are taken from more recent cohorts. Individuals in Card's data made their high school graduation and college entry decisions in the decade between the early 1960s and early 1970s. During this time, few states had significant student aid programs, and the major federal student grant and guaranteed loan programs that were passed into legislation in the mid-1960s were not funded at significant levels until after 1970.[26] Thus it is quite possible that earlier generations of young people faced credit access problems that later cohorts did not. Of course, the data sets differ in other ways, so further exploration of this hypothesis is an important avenue for future research.

VI.   Evidence from the Structural Model

This section presents estimates of a structural version of the model developed in Section III. A structural approach is valuable for at least three reasons. First, extending the empirical specifications to an explicit structural model contributes a robustness check to the empirical findings. While many empirical economists debate the relative merits of structural versus reduced-form or instrumental variable methods, few dispute that the methods typically depend on different identifying assumptions. It is reassuring when one obtains the same results from different approaches.

A second reason for the structural approach is to obtain the benefits of formal analysis. The model that guided the regression and instrumental variable analysis could at best suggest patterns of estimates con-

---

[25] The col. 5 interactions are somewhat difficult to interpret. While word score is inconsequential, math score is positive and significant, indicating that the presence of a local college induces more schooling from higher-ability students. It is unclear how to interpret this finding. Given the strength of test scores in predicting schooling, it is not surprising that the interaction is significant, perhaps as a result of functional form restrictions in the regression.

[26] An exception was the National Direct Student Loan program, which originated in 1958. However, the program was small and targeted science and engineering students.

sistent with the presence of borrowing constraints. The framework developed in this section provides an explicit definition of borrowing constraints and shows how they can be nonparametrically identified.[27] The clarity of this approach will aid future analysts in further explorations of borrowing constraints. For example, the manner in which unobserved heterogeneity enters the model is explicitly described. In addition, this approach advances our understanding of identification since the model exemplifies how the same economic idea (the difference between the effects of direct costs and forgone income in our case) can be used for identification in regression, instrumental variable, and structural models.

A third advantage of a structural model is that it provides estimated magnitudes of the parameters of interest. In our case, the model yields explicit estimates of the extent of borrowing constraints rather than heuristic tests of existence. Thus the estimated model allows us to perform and interpret policy simulations, such as predicting the impact of an expansion of programs giving below-market interest rate loans to the children of poor families.[28]

A.   The Econometric Model

This subsection extends the economic framework of Section III into two formal econometric models. The first mirrors the regression approach presented in Section V*F* by assuming that credit cost heterogeneity is governed by characteristics observable to the economist. The second approach assumes that borrowing constraints enter the model as unobserved heterogeneity. The notation below follows that established in Section III, with the exception that the subscript i is introduced to enumerate individuals.

If person i chooses schooling level S, his wages at time t are given by

$$\log(w_{Sit}) = \gamma_S + \mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{Lit}\boldsymbol{\beta}_L + \mathbf{X}'_{Eit}\boldsymbol{\beta}_E + u_{Sit}, \tag{18}$$

where u_{Sit} represents a random error term that is person-specific and specific to each schooling level. For convenience, observed right-hand-side exogenous variables are parceled into three vectors: $\mathbf{X}_{Wi}$, a vector of characteristics that do not vary over time; $\mathbf{X}_{Lit}$, time-varying local labor market measures; and $\mathbf{X}_{Eit}$, time-varying measures of work experience (experience and experience squared in our empirical implementation below). We assume that the u_{Sit} are orthogonal to all right-hand-side

---

[27] The technical appendix formally proves nonparametric identification.

[28] We do not perform such counterfactuals in this paper because we find no evidence of borrowing constraints.

variables for all S and t. Schooling effects are captured through the level shifter, γ_S, and the shocks, u_{Sit}.

If individual i were to choose schooling level S, his present value of earnings discounted to the time he leaves school would be

$$\sum_{t=S}^{\infty} \delta^{t-S} \exp\left(\gamma_S + \mathbf{X}'_{Wit}\boldsymbol{\beta}_W + \mathbf{X}'_{Lit}\boldsymbol{\beta}_L + \mathbf{X}'_{Eit}\boldsymbol{\beta}_E + u_{Sit}\right). \tag{19}$$

As in Section III, we abstract from uncertainty about the shocks $\{u_{Sit}\}$ and assume that students choose schooling to solve a certainty equivalence problem, so random variables are replaced with expected values. Let E_{0i} denote future wage expectations for individual i at time 0. Then the certainty equivalent present value of earnings net of direct costs dated at the initial period is defined for schooling level S as

$$I_{Si} = \left(\frac{1}{R_i}\right)^S \exp\left(\gamma_S + \mathbf{X}'_{Wi}\boldsymbol{\beta}_W\right) \left\{ \sum_{t=S}^{\infty} \delta^{t-S} \exp\left[\mathbf{X}'_{Eit}\boldsymbol{\beta}_E + E_{0i}(\mathbf{X}'_{Lit}\boldsymbol{\beta}_L)\right.\right.$$

$$\left.\left. + E_{0i}(u_{Sit})\right] \right\} - \sum_{t=0}^{S-1} \left(\frac{1}{R_i}\right)^t \mathbf{X}'_{Ci}\boldsymbol{\beta}_C, \tag{20}$$

where R_i represents the individual borrowing rate. Costs of schooling (τ_t in Sec. III) have been parameterized to depend on observables through the index $\mathbf{X}'_{Ci}\boldsymbol{\beta}_C$.[29]

For empirical implementation, we make three simplifying assumptions. First, we assume that wage equation residuals take the form

$$u_{Sit} = \theta_{Si} + \omega_{Sit}, \tag{21}$$

where θ_{Si} is a time-invariant random effect known to the individual when schooling decisions are made but not observed by the econometrician. The component ω_{Sit} is assumed orthogonal to all time 0 information for all t, so $E_{0i}(u_{Sit}) = \theta_{Si}$. Both components of the error terms are maintained to be strictly orthogonal to all right-hand-side observables for all S and t, although θ_{Si} may be arbitrarily correlated with $\theta_{S'i}$ for any S and S'.

Second, we approximate time 0 expectations of future local labor market conditions, $E_{0i}(\mathbf{X}'_{Lit}\boldsymbol{\beta}_L)$, by a linear function of local labor market variables known to individuals when schooling decisions are made. We let these variables depend on S and denote them by $\mathbf{X}_{LSi}$ with associated

---

[29] For computational convenience, we have assumed that costs depend only on observables. We see no reason why this should bias the findings in either direction.

coefficient vector $\tilde{\boldsymbol{\beta}}_{LS}$. To reduce notation, we incorporate the wage equation intercept γ_S into $\tilde{\boldsymbol{\beta}}_{LS}$ so that[30]

$$\mathbf{X}'_{LSi}\tilde{\boldsymbol{\beta}}_{LS} \equiv E_{0i}(\mathbf{X}'_{Lit}\boldsymbol{\beta}_l) + \gamma_S + \log\left[\sum_{t=S}^{\infty} \delta^{t-S} \exp(\mathbf{X}'_{Eit}\boldsymbol{\beta}_E)\right]. \qquad (22)$$

Third, we assume log utility for computational convenience. We see no reason why this simplification should bias the results in any particular direction.

Solving for the value function conditional on S and plugging in the present discounted value of earnings for individual i associated with schooling level S yields

$$V_{Si} = \left(\frac{1}{1-\delta}\right)[\log(I_{Si}) + \log(1-\delta)]$$

$$+ \left[\sum_{t=0}^{S-1} \delta^t t + \left(\frac{\delta^S}{1-\delta}\right)S\right]\log(\delta R_i) + \mathbf{X}'_{Ti}\boldsymbol{\beta}_{TS} + \nu_{Si}$$

$$\equiv \alpha_1 \log\left[\left(\frac{1}{R_i}\right)^S \exp(\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{LSi}\tilde{\boldsymbol{\beta}}_{LS} + \theta_{Si})\right.$$

$$\left. - \sum_{t=0}^{S-1}\left(\frac{1}{R_i}\right)^t \mathbf{X}'_{Ci}\boldsymbol{\beta}_C\right] + \alpha_{2S} + \alpha_{3S}\log(R_i) + \mathbf{X}'_{Ti}\boldsymbol{\beta}_{TS} + \nu_{Si}, \qquad (23)$$

where the error term and observable index $\mathbf{X}'_{Ti}\boldsymbol{\beta}_{TS} + \nu_{Si}$ are introduced to parameterize nonpecuniary benefits or costs associated with schooling level S (this is the $T(S)$ function defined in eq. [5] of Sec. III). The terms α_1, α_{2S}, and α_{3S} are parameters that are implicitly defined in (23) and do not vary across individuals.

B.   Sketch of Identification

To convey the basic mechanism through which identification is achieved, again assume two levels of schooling and no direct cost for $S = 0$. The value functions of schooling for person i as derived in equation (23) are given by

$$V_{0i} = \alpha_1(\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{L0i}\tilde{\boldsymbol{\beta}}_{L0} + \theta_{0i}) + a_1 + \mathbf{X}'_{Ti}\boldsymbol{\beta}_{T0} + \nu_{0i} \qquad (24)$$

---

[30] The term $\log\left[\sum_{t=S}^{\infty}\delta^{t-S}\exp(\mathbf{X}'_{Eit}\boldsymbol{\beta}_E)\right]$ is a constant that depends on S but not on i, so it can be incorporated into the intercept.

and

$$V_{1i} = \alpha_1 \log \{\exp\left[-\log(R_i) + \mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{L1i}\tilde{\boldsymbol{\beta}}_{L1} + \theta_{1i}\right] - \mathbf{X}'_{Ci}\boldsymbol{\beta}_C\}$$

$$+ a_2 + a_3 \log(R_i) + \mathbf{X}'_{Ti}\boldsymbol{\beta}_{T1} + \nu_{1i}, \tag{25}$$

where a_1, a_2, and a_3 are scalars (as defined in [23]) that do not vary across individuals.

Though (24) and (25) are complicated, these expressions are close to a familiar linear index model. A nonlinearity arises inside the logarithm in the first term of (25) and captures the interaction between the borrowing rate R_i and schooling costs $\mathbf{X}'_{Ci}\boldsymbol{\beta}_C$. No such interaction exists between interest rates and forgone earnings, which are represented by $\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{L0i}\tilde{\boldsymbol{\beta}}_{L0} + \theta_{0i}$. This feature of the model delivers identification of the parameters of interest.

We estimate two versions of the model. The first restricts variation in R_i to be determined through particular sets of observables, such as race or family income. The second version treats R_i as a variable known to the individual but not observed by the econometrician. Identification of each case is discussed separately.

1. Case I: R_i Determined by Observed Characteristics

The first approach assumes the absence of persistent, unobserved individual influences on wages ($\theta_{Si} = 0$) and assumes that R_i varies only with observables. The empirical content of this version of the model is closely related to the regression approach presented in Section V*F*. To see why, suppose that local labor market and wage variables are not determinants of schooling value ($\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{LSi}\tilde{\boldsymbol{\beta}}_{SL} = 0$ for both $S = 0$ and $S = 1$). Let R_i be parameterized by the index $\mathbf{X}'_{Ri}\boldsymbol{\beta}_R$, and assume that the schooling costs index, $\mathbf{X}'_{Ci}\boldsymbol{\beta}_C$, varies only with the presence of a local college. Equation (25) shows that $\log(R_i)$ enters the value function through the terms $\log\{\exp\left[-\log(R_i)\right] - \mathbf{X}'_{Ci}\boldsymbol{\beta}_C\}$ and $a_3 \log(R_i)$. Since the observable variables that determine R_i are also included in the set of observables governing tastes, $\mathbf{X}_{Ti}$, and since they are all dummy variables (such as race), $a_3 \log(R_i)$ is not separately identified from $\mathbf{X}'_{Ti}\boldsymbol{\beta}_{T1}$. Therefore, identification of $\log(R_i)$ comes completely from the term $\log\{\exp\left[-\log(R_i)\right] - \mathbf{X}'_{Ci}\boldsymbol{\beta}_C\}$. In addition, since $\mathbf{X}_{Ci}$ contains only a constant and an indicator variable, testing for interest rate heterogeneity (i.e., $\boldsymbol{\beta}_R = 0$) is identical to testing for interactions between the presence of a local college and the variables determining the borrowing rate (such as race) and almost identical to tests for the ad hoc interactions estimated in the schooling regressions of table 7. The main difference is that the schooling outcome is modeled here as a discrete variable.

However, the precisely specified econometric model set forth in equations (24) and (25) makes clear that schooling choice will generally be influenced by determinants of forgone wages, $\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{LSi}\tilde{\boldsymbol{\beta}}_{SL}$. Thus, when testing for borrowing constraints by looking for schooling cost interactions, one should account for the influence of forgone earnings by incorporating $\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{LSi}\tilde{\boldsymbol{\beta}}_{SL}$ into the model.

Because $\theta_{Si} = 0$ in this version of the model, there is no selection bias from schooling choices, so wage equation parameters $\boldsymbol{\beta}_W$ and $\tilde{\boldsymbol{\beta}}_{L1}$ can be consistently estimated from the wage equation alone. Thus the existence of heterogeneity of borrowing rates in the structural model is identified by the interaction between $\mathbf{X}_{Ri}$ and the presence of a college once we control for wage effects through $\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{LSi}\tilde{\boldsymbol{\beta}}_{LS}$.

A major advantage of specifying this model is that we can identify the magnitude of the borrowing constraint. The level of $\log(R_i)$ is not identified separately from the intercept in $\tilde{\boldsymbol{\beta}}_{LS}$. However, the relative values of $\log(R_i)$ among groups with different values of $\mathbf{X}'_{Ri}$ can be compared to a control group that borrows at the market rate. The magnitude of the difference in $\log(R_i)$ across groups can be pinned down by comparing the size of the interaction between $\log(R_i)$ and $\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{LSi}\tilde{\boldsymbol{\beta}}_{LS}$ (forgone earnings) relative to the interaction between $\log(R_i)$ and $\mathbf{X}'_{Ci}\boldsymbol{\beta}_C$ (direct costs). For instance, if there were just two groups a and b who face interest rates R^a and R^b, the level effect of $\log(R^a)$ is not identified, though the difference $\log(R^b) - \log(R^a)$ can be identified. In the empirical work below we choose a control group and assume that they borrow at the market rate. We then estimate the difference in borrowing rates between the control group and the constrained groups.

2. Case II: R_i Determined by Unobserved Characteristics

Next consider the more complicated case in which R_i is an unobserved random variable. To simplify the exposition, define $\mathbf{X}_i = (\mathbf{X}_{Wi}, \mathbf{X}_{L0i}, \mathbf{X}_{L1i})$. Taking the difference in schooling values from equations (24) and (25) yields the latent variable representation

$$V_{1i} - V_{0i} = \alpha_1 \log\left[\exp\left(\mathbf{X}'_i\Gamma_1 + \epsilon_{1i}\right) - \mathbf{X}'_{Ci}\boldsymbol{\beta}_C\right] + \mathbf{X}'_i\Gamma_2 + \epsilon_{2i}, \qquad (26)$$

where

$$\mathbf{X}'_i\Gamma_1 \equiv \mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{L1i}\tilde{\boldsymbol{\beta}}_{L1}, \qquad (27)$$

$$\epsilon_{1i} \equiv \theta_{1i} - \log(R_i), \qquad (28)$$

$$\mathbf{X}_i'\Gamma_2 \equiv \mathbf{X}_{Ti}'\boldsymbol{\beta}_{T1} - \alpha_1(\mathbf{X}_{Wi}'\boldsymbol{\beta}_W + \mathbf{X}_{L0i}'\tilde{\boldsymbol{\beta}}_{L0}) + a_2 - a_1 - \mathbf{X}_{Ti}'\boldsymbol{\beta}_{T0}, \quad (29)$$

and

$$\epsilon_{2i} \equiv a_3 \log(R_i) + \nu_{1i} - \alpha_1\theta_{0i} - \nu_{0i}, \quad (30)$$

so that Γ_1 and Γ_2 are defined appropriately. Individuals choose schooling option 1 if $V_{1i} - V_{0i} > 0$, and option 0 otherwise. Note that ϵ_{1i} enters the expression inside $\log(\cdot)$ in equation (26) and is correlated with the residual ϵ_{2i} both through the correlation between θ_{0i} and θ_{1i} and because $\log(R_i)$ enters both residuals.

The vector $(\mathbf{X}_i, \mathbf{X}_{Ci})$ is observable to the analyst but $(\theta_{1i}, \theta_{0i}, \log(R_i), \nu_{1i}, \nu_{0i})$ is not. We assume that $(\mathbf{X}_i, \mathbf{X}_{Ci})$ is independent of $(\theta_{1i}, \theta_{0i}, \log(R_i), \nu_{1i}, \nu_{0i})$. The selection equation (26) has three indices $(\mathbf{X}_i'\Gamma_1, \mathbf{X}_i'\Gamma_2, \mathbf{X}_{Ci}'\boldsymbol{\beta}_C)$ and two error terms $(\epsilon_{1i}, \epsilon_{2i})$. The three degrees of freedom represented by these index functions are sufficient to trace out the joint distribution of $(\epsilon_{1i}, \epsilon_{2i})$.

Equations (28) and (30) show that these indices cannot also identify the distribution of $\log(R_i)$ from the distribution of θ_{1i}. However, since wages depend on θ_{1i} but not on R_i, the wage equation given by (18) provides sufficient variation to trace out the distribution of $\log(R_i)$ up to a location normalization.

Theorem 1 contained in the appendix (available at http://www .econ.northwestern.edu/faculty/taber) proves nonparametric identification of the joint distribution of the error terms from the schooling choice and wage equations, $(\epsilon_{1i}, \epsilon_{2i}, u_{1it})$, using two exclusion restrictions, one for opportunity costs and one for direct costs.[31] Once this joint distribution is identified, we can identify the conditional expectation

$$E(u_{1it}|\epsilon_{1i}) = E(\theta_{1i}|\theta_{1i} - \log(R_i)).$$

The appendix further demonstrates that this conditional expectation can be used to estimate the distribution of R_i when R_i is independent of θ_{1i}. Thus the entire distribution of interest rates faced by students is identified up to a normalization.

Intuition for this result follows from the fact that the form of the selection bias

$$E(\theta_{1i}|\alpha_1 \log\{\exp[\mathbf{X}_i'\Gamma_1 + \theta_{1i} - \log(R_i)] - \mathbf{X}_{Ci}'\boldsymbol{\beta}_C\} + \mathbf{X}_i'\Gamma_2 + \epsilon_{2i} > 0,$$

$$\mathbf{X}_i'\Gamma_1, \mathbf{X}_{Ci}'\boldsymbol{\beta}_C, \mathbf{X}_i'\Gamma_2) \quad (31)$$

depends on the realized value of college costs. This expression is complicated and depends on the full distribution of $(\theta_{1i}, R_i, \epsilon_{2i})$. Under the same logic outlined in Section III, the term $\theta_{1i} - \log(R_i)$ is more im-

[31] With the notation used above, the first type of exclusion restriction affects $\mathbf{X}_i'\Gamma_2$, but not $\mathbf{X}_i'\Gamma_1$ or $\mathbf{X}_{Ci}'\boldsymbol{\beta}_C$; the second enters $\mathbf{X}_{Ci}'\boldsymbol{\beta}_C$ but not $\mathbf{X}_i'\Gamma_2$ or $\mathbf{X}_i'\Gamma_1$.

portant when costs of schooling are high because the $\log(\cdot)$ term in equation (31) is nonlinear. Suppose that borrowing constraints were not important. There would be more selection bias in counties that do not have a college since θ_{1i} is relatively more important when college is costly. If borrowing constraints are binding and schooling costs are high, schooling decisions are dictated by borrowing costs. Thus we would observe less selection bias when no college is present.

C. Evidence When R_i Is Determined by Observed Characteristics

We parameterize interest rates by a function $R(\cdot)$ and an index so that $R(\mathbf{X}'_{Ri}\boldsymbol{\beta}_R) \geq 1$ and $R(0) = 1/\delta$, where $\mathbf{X}_{Ri}$ are observables that govern the borrowing rate.[32] Thus R_i cannot be less than one but can be arbitrarily large. This form facilitates comparison between borrowing-constrained individuals and a control group that is not constrained. Since we do not include an intercept in $\mathbf{X}_{Ri}$, any population group with $\mathbf{X}_{Ri}$ equal to zero borrows at the market rate. For example, let $\mathbf{X}_{Ri}$ consist of dummy variables for black and Hispanic status. For whites, $\mathbf{X}'_{Ri}\boldsymbol{\beta}_R = 0$, so $R_i = 1/\delta$. A positive coefficient on the black or Hispanic dummy implies that this subgroup borrows at an interest rate greater than the market rate.

In this subsection, we report parameter estimates for the model defined in equations (18) and (23). To begin, we ignore selection by setting $\theta_{Si} = 0$, which implies that the schooling choice residuals (ν_{Si}) are independent of the wage residuals (u_{Sit}). This restriction is relaxed in the next subsection. The $\boldsymbol{\beta}_W$ parameters are estimated by OLS in a standard Mincer-style log wage regression, and the estimated values are inserted into the value function (the standard errors of the value function parameters are adjusted appropriately). The cost index, $\mathbf{X}'_{Ci}\boldsymbol{\beta}_C$, consists of only the variable local college and a constant. The vector of taste variables, $\mathbf{X}_{Ti}$, contains family background variables, AFQT score, minority dummies, and dummies for urban residence and region of residence.[33] Local labor market variables enter through the index $\mathbf{X}'_{LSi}\tilde{\boldsymbol{\beta}}_{LS}$. We incorporate two measures of local labor market conditions. The first is the local earnings variable used above. It is restricted to affect only the wages of high school graduates and high school dropouts. The second variable is the long-run average income in the county of residence at age 17 (mean earnings over working life) and affects wages at all schooling levels.

We allow for four levels of schooling: high school dropout ($S = 0$),

---

[32] The exact form of $R(\mathbf{X}'_{Ri}\boldsymbol{\beta}_R)$ is $\exp(\exp\{\mathbf{X}'_{Ri}\boldsymbol{\beta}_R + \log[-\log(\delta)]\})$.

[33] Most of these variables entered the schooling regressions in table 2 except the college in county indicator and the local labor market variables. Also, AFQT is used instead of the set of four test scores.

high school graduate ($S = 2$), some college ($S = 4$), and college graduate or more ($S = 6$).[34] During high school, students incur no direct costs of schooling but must pay some direct costs to attend college and to graduate from college.

We assume that ν_{Si} has a generalized extreme value (GEV) distribution, so the model can be estimated as a nested logit. We use two levels of nesting. The high school decisions are nested together, and college decisions are nested. Observable components of the utilities take the form

$$\mu_{0i} = \alpha_1(\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{L0i}\tilde{\boldsymbol{\beta}}_{L0}) + \alpha_{20},$$

$$\mu_{2i} = \alpha_1[-2\log(R_i) + \mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{L2i}\tilde{\boldsymbol{\beta}}_{L2}] + \alpha_{22} + \alpha_{32}\log(R_i) + \mathbf{X}'_{Ti}\boldsymbol{\beta}_{T2},$$

$$\mu_{4i} = \alpha_1 \log\left[\left(\frac{1}{R_i}\right)^4 \exp(\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{L4i}\tilde{\boldsymbol{\beta}}_{L4}) - \sum_{t=2}^{3}\left(\frac{1}{R_i}\right)^t \mathbf{X}'_{Ci}\boldsymbol{\beta}_C\right] + \alpha_{24}$$
$$+ \alpha_{34}\log(R_i) + \mathbf{X}'_{Ti}\boldsymbol{\beta}_{T4},$$

$$\mu_{6i} = \alpha_1 \log\left[\left(\frac{1}{R_i}\right)^6 \exp(\mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{L6i}\tilde{\boldsymbol{\beta}}_{L6}) - \sum_{t=2}^{5}\left(\frac{1}{R_i}\right)^t \mathbf{X}'_{Ci}\boldsymbol{\beta}_C\right] + \alpha_{26}$$
$$+ \alpha_{36}\log(R_i) + \mathbf{X}'_{Ti}\boldsymbol{\beta}_{T6},$$

where the α_{jS} are defined in equation (23).

Nesting yields the following schooling probabilities:

$$\Pr(S = 6 | S > 2, \mu_{0i}, \ldots, \mu_{6i}) = \frac{\exp(\mu_{6i}/\rho_c)}{\exp(\mu_{6i}/\rho_c) + \exp(\mu_{4i}/\rho_c)},$$

$$\Pr(S = 2 | S > 0, \mu_{0i}, \ldots, \mu_{6i}) =$$
$$\frac{\exp(\mu_{2i}/\rho_h)}{\exp(\mu_{2i}/\rho_h) + [\exp(\mu_{4i}/\rho_c) + \exp(\mu_{6i}/\rho_c)]^{\rho_c/\rho_h}},$$

$$\Pr(S = 0 | \mu_{0i}, \ldots, \mu_{6i}) =$$
$$\frac{\exp(\mu_{0i})}{\exp(\mu_{0i}) + \{\exp(\mu_{2i}/\rho_h) + [\exp(\mu_{4i}/\rho_c) + \exp(\mu_{6i}/\rho_c)]^{\rho_c/\rho_h}\}^{\rho_h}},$$

where ρ_h and ρ_c are parameters for the nesting of high school graduates and college attenders, respectively.[35] The model is estimated by maximum likelihood, restricting ρ_h and ρ_c to lie between zero and one. Taber

---

[34] Recipients of the general equivalency diploma who attend no college are counted as dropouts (see Cameron and Heckman 1993).
[35] The probability for $S = 4$ can be constructed from these three probabilities.

(2000) shows that this nested logit model can be interpreted as a forward-looking, discrete-choice dynamic programming model.

A number of normalizations are necessary for identification. We assume that $\boldsymbol{\beta}_{TS} = 0$ for dropouts, which normalizes location as in a standard polychotomous choice model. In the high school equations, we cannot separately identify the intercepts in $\boldsymbol{\beta}_{TS}$ from the intercept in $\boldsymbol{\beta}_{LS}$, so the latter are set to zero. We normalize $\boldsymbol{\beta}_C$, the coefficient on local college, to one in the equations for college. This normalization is needed because the scale of the indices cannot be separately distinguished from the intercept in the tastes index.[36]

Numerically, we found the intercept in the cost equation to be difficult to nail down, so we fixed it to a value of two in the simulations.[37] Extensive experimentation with other values for this parameter showed that the results were insensitive to this choice. We also had numerical trouble estimating a precise value for the coefficient on the local earnings variable in the college states. A reduced-form regression yielded a value of approximately 0.3, so we fixed the coefficient to this value in the two college equations. Again, the results were not sensitive to this restriction. Finally, we assumed $\delta = 0.97$, which implies a market interest rate of $R = 1/\delta \approx 1.031$.

Table 8 exhibits parameter estimates for the base model, in which we do not allow for interest rate heterogeneity ($\mathbf{X}'_{Ri}\boldsymbol{\beta}_R = 0$). Panel A shows estimates of taste parameters, panel B shows estimates of local labor market variables, and panel C shows estimates of the GEV scale and nesting parameters. For the sake of brevity, we do not provide a general discussion of the parameters, but they appear reasonable to us and broadly consistent with the results in the educational attainment literature discussed in Section II*A*.

Column 1 of table 9 exhibits estimated educational borrowing rates using four alternative sets of observed characteristics. Column 2 presents negative log likelihood values for each specification. Panel B reports implied borrowing rates when two indicator variables for black and Hispanic racial-ethnic status determine the borrowing rate. Unexpectedly, point estimates of the borrowing rate come in at the boundary value of 1.0, which implies that blacks and Hispanics borrow for education below the assumed market interest rate.

Panel C exhibits the case in which borrowing rates depend on parental education through four dummy variables for interactions between father and mother with a college education or a high school education. The

---

[36] To see this, notice in eq. (23) that one can multiply $\boldsymbol{\beta}_C$ by any positive number γ, increase the intercept in $\tilde{\boldsymbol{\beta}}_{LS}$ by $\log(\gamma)$, and decrease the intercept in $\boldsymbol{\beta}_{TS}$ by $\alpha_1 \log(\gamma)$ without changing the value of V_{Si}. Normalizing the scale of $\boldsymbol{\beta}_C$ in this way performed better computationally than alternative normalizations.

[37] This amounts to assuming that living at home cuts the cost of schooling in half.

TABLE 8
ESTIMATED STRUCTURAL SCHOOLING MODEL I: HETEROGENEITY IN BORROWING RATES
DEPENDS ON OBSERVED CHARACTERISTICS WITHOUT BORROWING CONSTRAINT
RESTRICTIONS

	Attend High School	Graduate High School	Attend College	Graduate College
A. Taste Parameters ($\boldsymbol{\beta}_T$)				
Black		.855	1.215	1.570
		(.169)	(.202)	(.252)
Hispanic		.136	.488	.539
		(.187)	(.212)	(.279)
AFQT score		.045	.054	.101
		(.005)	(.008)	(.019)
Father dropout		−.055	−.123	−.028
		(.144)	(.158)	(.198)
Father some college		.076	.233	.603
		(.258)	(.273)	(.293)
Father college graduate		.445	.618	1.322
		(.349)	(.359)	(.357)
Mother dropout		−.142	−.227	−.392
		(.140)	(.157)	(.209)
Mother some college		.077	−.162	.376
		(.295)	(.318)	(.319)
Mother college graduate		.200	.332	.907
		(.396)	(.406)	(.417)
Number of siblings		−.029	−.092	−.124
		(.042)	(.047)	(.055)
Family income		.162	.121	.136
		(.043)	(.045)	(.054)
Geographic controls		yes	yes	yes
Cohort controls		yes	yes	yes
B. Local Labor Market Parameters ($\boldsymbol{\beta}_L$)				
Intercept			5.689	5.236
			(.417)	(.338)
Local earnings at age 17	.044	.009		
	(.047)	(.022)		
Mean local earnings over working life	.302	.293	.300	.300
	(.017)	(.008)		
C. Additional Parameters				
Scale	1.281			
	(.140)			
High school nesting (ρ_H)	.452			
	(.146)			
College nesting (ρ_C)	1.000			

NOTE.—Standard errors are in parentheses. Since the model is overparameterized, the standard errors are likely to be misleading. In the estimates shown here, estimates of borrowing constraint parameters are not estimated. This is the base model. See panel A of table 9. The equations for schooling choice and wages are specified as

$$V_s = \alpha_1 \log\left[\left(\frac{1}{R_i}\right)^S \exp\left(\tilde{\gamma}_S + \mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{L0i}\tilde{\boldsymbol{\beta}}_{LS}\right) - \sum_{t=0}^{S}\left(\frac{1}{R_i}\right)^t \mathbf{X}'_{Ci}\boldsymbol{\beta}_C\right] + \alpha_2(S) + \alpha_3(S)\log(R_i) + \mathbf{X}'_{Ti}\boldsymbol{\beta}_{TS} + u_{TSi};$$

$$\log(w_{Sit}) = \gamma_S + \mathbf{X}'_{Wi}\boldsymbol{\beta}_W + \mathbf{X}'_{Lit}\boldsymbol{\beta}_L + E_{it}\boldsymbol{\beta}_E + \alpha_S\theta + \nu_{Sit}.$$

In this set of estimates the market interest rate is fixed such that $1/(1 + r) = \delta = .97$. Geographic and cohort controls are defined in the notes to table 2.

TABLE 9
Tests for the Presence of Borrowing Constraints Calculated from Structural
Model I (Estimates Reported in Table 8)

Specification	Borrowing Rate[a] (1)	Negative Log Likelihood (2)
A. Unrestricted		2,547.69
Everyone	1.031	
B. Racial groups		2,547.34
Whites	1.031	
Blacks	1.000	
Hispanics	1.000	
C. Parents' education		2,547.42
Both college educated	1.031	
Father 12, mother 12	1.000	
Father 12, mother college	1.000	
Father college, mother 12	1.033	
D. Family income		2,546.91
Top third	1.031	
Middle third	1.032	
	(.006)	
Bottom third	1.038	
	(.008)	
E. Number of siblings[b]		2,547.69
Zero	1.031	
Two	1.031	
	(.003)	
Four	1.031	
	(.006)	

Note.—The standard errors (in parentheses) are presented only for the family income variables and number of siblings because the model converged to a corner solution in the other cases.

[a] The market rate of return is fixed to be 1.031.

[b] In the results shown in panels B–D, the borrowing rate is interacted with the dummy variables shown. Here it is interacted with the number of siblings, so the comparison group is the set of individuals with no siblings.

control group comprises individuals with two college-educated parents who are assumed to borrow at the market rate. Estimated borrowing rates show little deviation from the market rate.

In panel D, borrowing rates depend on whether parental family income falls into the top, middle, or bottom third of the family income distribution for families with children of college age. The top third is the control group. Estimated interest rates for the middle and bottom thirds are barely distinguishable from the market rate, and again statistical tests fail to reject the hypothesis of no credit limitations. The standard errors show that the interest rate parameters are precisely estimated, so failure to reject the null hypothesis is a consequence of the small differences in parameter values and not of lack of statistical power.

Panel E shows the case in which borrowing rates depend on family size. One might expect to find the strongest evidence for heterogeneity in credit access in this case. However, borrowing rates turn out to be

insensitive to the number of children in the family (number of siblings). Once again, confidence intervals are tight, indicating that the null is rejected because of small coefficients, not lack of power.

D.   Evidence When R_i Is Not Observed

Failing to find evidence of heterogeneity in borrowing rates in observables, we extend the structural model to capture borrowing rate heterogeneity in unobservables. We assume that individuals come in one of two types: those borrowing at the market rate $1/\delta$ and those who are "constrained" and borrow at rate $R_c > 1/\delta$. We estimate the borrowing rate R_c and the fraction of the population that borrows at the higher rate, denoted as P_c. The distribution of borrowing rates is restricted to be independent of the observables and other error terms.

Unobserved ability enters the model through a single standard normal factor θ_i, which is known by the agent at the time schooling choices are made and is independent of observables. The error terms in the wage equations are defined as

$$u_{Sit} = \phi_{WS}\theta_i + \omega_{Sit}, \tag{32}$$

where ϕ_{WS} is the wage equation factor loading term, and ω_{Sit} is independently and identically distributed over time and is orthogonal to θ_i for all t. Ability may also be correlated with unobserved tastes for schooling. The taste residual for school level S is now defined as

$$\nu_{Si} = \phi_{TS}\theta_i + \tilde{\nu}_{Si}, \tag{33}$$

where ϕ_{TS} is the tastes equation factor loading for school level S, and $\tilde{\nu}_{Si}$ has a GEV distribution, which yields the nested logit model from the previous section.

A major goal of the approach in this subsection is to identify the extent of borrowing constraints from economic influences regulating the behavioral interaction between interest rates and schooling costs and not from functional form assumptions on the distribution of the error terms. This goal led to two nonstandard estimation strategies. First, because we have no presumption about the form of the selection bias, we retain flexibility in its modeling. Ideally, we would place no restrictions on the joint distribution of the taste and wage errors (ν_{2i}, ν_{4i}, ν_{6i}, u_{0it}, u_{2it}, u_{4it}, u_{6it}). However, full nonparametric estimation of this distribution is not computationally feasible. Even after we restrict the distribution of the error terms to specific functional forms with a one-factor representation, there are still seven factor loading terms on θ_i that determine the form of the selection bias. We purposely overparameterize the selection bias function in order to force identification from

interactions between the selection terms and the college costs index. As a result, we could not obtain standard errors of the parameters because the estimated Hessian matrix turned out to be approximately singular in all cases. This posed little difficulty since consistent estimation of the R_c distribution is still possible, and likelihood ratio type statistics were employed to test for borrowing constraints.

We initially attempted to estimate the model using maximum likelihood, but estimates appeared sensitive to restrictions on the factor loading terms. In particular, we could not be sufficiently flexible with the distribution of ω_{Si} to distinguish the distribution of R_c from a flexible functional form for the school choice equation. Identifying the model through stronger functional form restrictions is not in the spirit of our nonparametric identification results. This difficulty led us to a second nonstandard estimation strategy. We used an iterative procedure to force identification of the distribution of R_c and other parameters to come from appropriate variation in the data.

Partition model parameters into those identified primarily from the selection equation ($\mathbf{\Psi}_1$) and those that require the wage equation ($\mathbf{\Psi}_2$) for identification. Redefine $\mathbf{X}_i$ to denote all observables, and let $\mathcal{L}_i(\mathbf{\Psi}_1, \mathbf{\Psi}_2; \mathbf{X}_i)$ be the log likelihood function for the schooling decision only (derived as a nested logit with unobserved heterogeneity). The model is estimated using the following iterative procedure: (1) Fix $\mathbf{\Psi}_2$ and solve for the value of $\mathbf{\Psi}_1$ that maximizes the log likelihood function. (2) Fix $\mathbf{\Psi}_1$ and solve for the value of $\mathbf{\Psi}_2$ that minimizes nonlinear least squares,

$$\sum_{i=1}^{N} \sum_{t=1}^{T_i} \{\log (w_{S_i it}) - E[\log (w_{S_i it}) | \mathbf{X}_i, S_i; \mathbf{\Psi}_1, \mathbf{\Psi}_2]\}^2, \tag{34}$$

where S_i represents the realized value of S for person i.

We iterate until convergence.[38] Locally, this procedure is equivalent to generalized method of moments with moment conditions

$$E\left[\frac{\partial \mathcal{L}_i(\mathbf{\Psi}_1, \mathbf{\Psi}_2; \mathbf{X}_i)}{\partial \mathbf{\Psi}_1}\right] = 0,$$

$$E\left[\frac{\partial E[\log (w_{S_i it}) | \mathbf{X}_i, S_i; \mathbf{\Psi}_1, \mathbf{\Psi}_2]}{\partial \mathbf{\Psi}_2} \{\log (w_{S_i it})\right.$$

$$\left. - E[\log (w_{S_i it}) | \mathbf{X}_i, S_i; \mathbf{\Psi}_1, \mathbf{\Psi}_2]\}\right] = 0. \tag{35}$$

[38] In practice we use a damping parameter to aid convergence.

The distribution of R_c was estimated with the parameters in $\mathbf{\Psi}_2$. To keep the form of the selection bias flexible, we estimate the factor loading terms on θ_i in the wage equation also. We use the following partition:

$$\mathbf{\Psi}_1 = \{\alpha_1, \rho_h, \rho_c, (\boldsymbol{\beta}_{TS}, \tilde{\boldsymbol{\beta}}_{LS} : S \in \{0, 2, 4, 6\})\},$$

$$\mathbf{\Psi}_2 = \{\boldsymbol{\beta}_W, \boldsymbol{\beta}_L, \boldsymbol{\beta}_E, P_c, R^c, (\gamma_S, \phi_{LS}, \phi_{WS} : S \in \{0, 2, 4, 6\})\}.$$

Panel A of table 10 displays parameter estimates from this approach.[39] The model produced an estimated borrowing rate of 1.071 for the borrowing-constrained group (compared to 1.030 for the unconstrained group). However, the estimate of the fraction of the population borrowing at the higher rate hit the boundary constraint so that the estimated fraction is essentially zero. Both because of the reason mentioned above and because the estimate is on the boundary, standard errors of the parameters could not be calculated. We experimented with large numbers of alternative starting values and specifications, none of which produced evidence of heterogeneity in borrowing rates. Thus this final approach, like the previous ones, provides no evidence that borrowing constraints are important for schooling decisions.

VII. Summary and Conclusions

This paper develops economic reasoning that shows that direct costs of schooling affect borrowing-constrained individuals in a fundamentally different way than opportunity costs. Using variation in direct costs and opportunity costs, we test for the importance of educational borrowing constraints in four different ways: (1) instrumental variable wage regressions, (2) years of schooling regressions with interactions between college costs and various observables likely to be correlated with borrowing constraints, (3) a structural econometric model in which borrowing rates depend on observables, and (4) a structural model that allows for unobservable heterogeneity in borrowing rates. We find no evidence that borrowing constraints impede schooling progression using any of the methods.

A cardinal issue that arises with this type of finding is statistical power. Have we found no evidence of borrowing constraints because they are

---

[39] Experimenting with a wide variety of specifications and starting values showed that the model did not typically converge when no restrictions were placed on the ϕ_{ST} parameters (factor loading on heterogeneity in tastes for schooling). Consequently, we restricted the factor loadings in the high school dropout and high school graduation equations to be equal ($\phi_{T0} = \phi_{T2}$) and the some college and graduate college factor loadings to be equal also ($\phi_{T4} = \phi_{T6}$). These restrictions did not appear to affect estimates but aided convergence substantially.

TABLE 10

Estimates of Structural Model II Allowing for Unobserved Heterogeneity in
Borrowing Rates: Borrowing Rate Distribution and Wage Equation

A. Parameters of Borrowing Rate (R_i) Distribution

Borrowing rate[a]	1.031	1.071
Probability	1.00	.00

B. Wage Equation

		Factor load-ing:	
Black	−.095		
Hispanic	−.047	Dropout (α_0)	.176
Experience	.064	High school graduate (α_1)	.155
Experience squared	−.002	Some college (α_2)	.153
AFQT score	.007	College graduate (α_3)	.069
Current local earnings	.027	Family background controls	yes
Intercept:		Geographic controls	yes
Dropout	1.046	Cohort controls	yes
High school graduate	1.139		
Some college	1.031		
College graduate	1.324		

C. Taste Parameters (β_T)

	Graduate High School	Attend College	Graduate College
Black	.632	1.725	1.873
Hispanic	.057	.838	.793
AFQT score	.043	.088	.128
Father dropout	−.144	−.235	−.108
Father some college	.051	.678	.900
Father college graduate	.317	1.243	1.683
Mother dropout	−.166	−.482	−.662
Mother some college	.144	−.271	.336
Mother college graduate	.198	.650	1.244
Number of siblings	−.007	−.113	−.099
Factor loading (δ_T)	.000	2.130	2.130
Geographic controls	yes	yes	yes

D. Local Labor Market Parameters

	Attend High School	Graduate High School	Attend College	Graduate College
Intercept			2.746	3.225
Local earnings at age 17	.408	.258		
Mean local earnings over life	.245	.172	.300	.300

E. Additional Parameters

Scale	.231
High school nesting (ρ_H)	.452
College nesting (ρ_C)	1.000

Note.—See the note to table 8 for a description of the model equations. Cohort and geographic controls are defined in the notes to table 2. Family background controls are the same as those used in the wage regressions reported in table 3.

[a] The market rate of return is fixed to be 1.031.

not problematic or because we cannot measure the relevant parameters with sufficient precision? It is never possible to completely separate these two claims. In the end, the answer chosen depends on prior beliefs about the size of the effect and the empirical content of the data. Our view is that power is less of a concern in the schooling choice regressions and the structural model that restricts borrowing rate heterogeneity to vary with observables. The presence of a local college variable is powerful in these models; if interactions with it were important, we would expect to detect them. The small standard errors on the interest rates in table 9 support this view. Estimates are precise and show little heterogeneity in borrowing rates across families of different sizes and income levels.

By contrast, it is less clear how much power we have with the imprecisely determined instrumental variable estimates of schooling returns. Nevertheless, in interpreting the instrumental variable results, one should keep several points in mind. First, evidence supporting discount rate bias is weak. A number of studies report instrumental variable estimates of schooling returns larger than OLS estimates, but instrumental variable standard errors in these studies are consistently found to be large. Second, a few recent studies instrument schooling with a variable such as presence of a local college. Our results show that these estimates are sensitive to inclusion of measures for local labor market conditions. Proper inclusion of local labor market controls weakens or eliminates the inflated instrumental variable estimates of schooling returns and yields instrumental variable estimates near or below OLS estimates. It seems plausible that had other researchers included local labor market controls, they may have found instrumental variable estimates of schooling returns near or below OLS estimates. If so, Card's (1999) case that instrumental variable estimates are consistently higher than OLS estimates would weaken considerably. Third, our conclusions are not based on standard errors alone. Point estimates consistently turn up with the wrong sign to be consistent with credit constraints.

Fourth, theoretical bias arising from problems with instruments in the instrumental variable wage specifications generally goes in the wrong direction compared to actual estimates. Table 5 suggests that return to schooling instrumented with presence of a local college may be biased upward because of potential correlation with unobserved ability. Moreover, if family contributions to borrowing-constrained students diminish during economic recessions, instrumental variable estimates will be biased downward and will come in below OLS estimates. The instrumental variable results per se may not be convincing evidence that borrowing constraints do not exist, but at the very least, they cast doubt on the discount rate bias story that predicts that instrumental variable estimates of schooling returns are biased upward.[40]

[40] Ashenfelter, Harmon, and Oosterbeek (1999) have an alternative argument for the

When standard errors could be calculated in the structural selection model, confidence intervals were found to be small. When the parameters of the heterogeneity distribution converged to boundary values, standard errors could not be calculated. In these cases, though, the boundary values of the estimates imply no heterogeneity in credit access. We believe that this approach is useful and hope that it leads to additional studies.

All told, our results are consistent with the conclusions of Shea (2000), Cameron and Heckman (2001), and Keane and Wolpin (2001). All four methods we employ produce no evidence of borrowing constraints. When interpreting our findings, we reiterate two caveats. First, none of our evidence or the evidence in any of these other studies is direct since the actual rates at which individuals borrow for school cannot be observed. Second, our results do not imply that credit market constraints would not exist in the absence of the assortment of private and government programs currently available. Instead, our findings show that given the large range of subsidies to education that currently exist, there is no evidence of underinvestment in schooling resulting from credit access.

References

Altonji, Joseph G., and Thomas A. Dunn. 1996. "The Effects of Family Characteristics on the Return to Education." *Rev. Econ. and Statis.* 78 (November): 692–704.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2002. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." Manuscript. Evanston, Ill.: Northwestern Univ., Dept. Econ.

Angrist, Joshua D., and Alan B. Krueger. 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*, vol. 3A, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier.

Arkes, Jeremy. 1998. "Using State Unemployment Rates during Teenage Years as an Instrument to Estimate the Returns to Schooling." Manuscript. Santa Monica, Calif.: Rand Corp.

Ashenfelter, Orley, Colm Harmon, and Hessel Oosterbeek. 1999. "A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias." *Labour Econ.* 6 (November): 453–70.

Becker, Gary S. 1967. *Human Capital and the Personal Distribution of Income: An Analytical Approach.* Woytinsky Lecture no. 1. Ann Arbor, Mich.: Inst. Public Admin.

Cameron, Stephen V., and James J. Heckman. 1993. "The Nonequivalence of High School Equivalents." *J. Labor Econ.* 11, no. 1, pt. 1 (January): 1–47.

———. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males." *J.P.E.* 106 (April): 262–333.

instrumental variable results in the literature. They argue that the pattern may be due to a troubling bias in published specifications. Since instrumental variable standard errors are often high, if researchers prefer specifications in which the *t*-statistic is greater than two, there will be a bias toward publishing studies with high instrumental variable estimates.

———. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *J.P.E.* 109 (June): 455–99.

Card, David. 1995*a*. "Earnings, Schooling, and Ability Revisited." In *Research in Labor Economics*, vol. 14, edited by Solomon W. Polachek. Greenwich, Conn.: JAI.

———. 1995*b*. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*, edited by Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky. Toronto: Univ. Toronto Press.

———. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, vol. 3A, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier.

———. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69 (September): 1127–60.

Carneiro, Pedro, James J. Heckman, and Dayanand Manoli. 2002. "Human Capital Policy." Manuscript. Chicago: Univ. Chicago, Dept. Econ.

Hauser, Robert M. 1993. "Trends in College Entry among Blacks, Whites, and Hispanics." In *Studies of Supply and Demand in Higher Education*, edited by Charles Clotfelter and Michael Rothschild. Chicago: Univ. Chicago Press (for NBER).

Heckman, James J., and Edward Vytlacil. 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." *J. Human Resources* 33 (Fall): 974–87.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (March): 467–75.

Kane, Thomas J. 1994. "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education." *J.P.E.* 102 (October): 878–911.

Keane, Michael P., and Kenneth I. Wolpin. 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *Internat. Econ. Rev.* 42 (November): 1051–1103.

Kling, Jeffrey R. 2001. "Interpreting Instrumental Variables Estimates of the Returns to Schooling." *J. Bus. and Econ. Statis.* 19 (July): 358–64.

Lang, Kevin. 1993. "Ability Bias, Discount Rate Bias, and the Return to Education." Manuscript. Boston: Boston Univ., Dept. Econ.

Levy, Dan, and Greg Duncan. 2000. "Using Sibling Samples to Assess the Effect of Family Childhood Income on Completed Schooling." Working Paper no. 168 (April). Evanston, Ill.: Northwestern Univ., Joint Center Policy Res.

Lochner, Lance, and Alexander Monge. 2001. "Human Capital Formation with Endogenous Credit Constraints." Manuscript. Evanston, Ill.: Northwestern Univ., Dept. Econ.

Manski, Charles F. 1993. "Income and Higher Education." *Focus* 14 (Winter): 14–19.

Manski, Charles F., and David A. Wise. 1983. *College Choice in America*. Cambridge, Mass.: Harvard Univ. Press.

Mare, Robert D. 1980. "Social Background and School Continuation Decisions." *J. American Statis. Assoc.* 75 (June): 295–305.

Mayer, Susan E. 1997. *What Money Can't Buy: Family Income and Children's Life Chances*. Cambridge, Mass.: Harvard Univ. Press.

Mulligan, Casey B. 1997. *Parental Priorities and Economic Inequality*. Chicago: Univ. Chicago Press.

Rosen, Sherwin. 1977. "Human Capital: A Survey of Empirical Research." In *Research in Labor Economics*, vol. 1, edited by Ronald G. Ehrenberg. Greenwich, Conn.: JAI.

Shavit, Yossi, and Hans-Peter Blossfeld. 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries.* Boulder, Colo.: Westview.

Shea, John. 2000. "Does Parents' Money Matter?" *J. Public Econ.* 77 (August): 155–84.

Taber, Christopher R. 2000. "Semiparametric Identification and Heterogeneity in Discrete Choice Dynamic Programming Models." *J. Econometrics* 96 (June): 201–29.

———. 2001. "The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?" *Rev. Econ. Studies* 68 (July): 665–91.

Tomes, Nigel. 1981. "The Family, Inheritance, and the Intergenerational Transmission of Inequality." *J.P.E.* 89 (October): 928–58.

Willis, Robert J. 1986. "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions." In *Handbook of Labor Economics*, vol. 1, edited by Orley Ashenfelter and Richard Layard. Amsterdam: North-Holland.

Willis, Robert J., and Sherwin Rosen. 1979. "Education and Self-Selection." *J.P.E.* 87, no. 5, pt. 2 (October): S7–S36.