Borrowing Constraints and the Returns to Schooling

Stephen Cameron [1]
Department of Economics
Columbia University and
Federal Reserve Bank of New York

Christopher Taber
Department of Economics and
Institute for Policy Research
Northwestern University

May 1, 2000

[1]For helpful comments we thank Joe Altonji, Tim Conley, Lance Lochner, Larry Kenny, Craig Olson, and seminar participants at the Federal Reserve Bank of New York, MIT, Northwestern, Princeton, University of Florida, University of North Carolina. We thank Tricia Gladden for superior research assistance and thoughtful comments. We also than Jeff Kling for providing us with his code. For financial support, Taber acknowledges NSF Grant SBR-97-09-873, and Cameron acknowledges NSF Grant SBR-97-30-657.

Abstract

To a large degree, the expansion of student aid programs to potential college students over the past 25 years in the United States has been based on the presumption that borrowing constraints present an obstacle to obtaining a college education. Economists and sociologists studying schooling choices have found empirical support for college subsidies in the well-documented, large positive correlation between family income and schooling attainment. This correlation has been widely interpreted as evidence of credit constraints. Recently, however, Cameron and Heckman (1998, 2000), Keane and Wolpin (1999), and Shea (1999) have examined more closely family income's influence on schooling and question whether borrowing constraints plays any role on college choices.

Over the last 20 years, a separate literature in economics has aimed at estimating the return to schooling purged of "ability bias." Unobserved ability upwardly biases least-squares estimates of returns to schooling. However, using instrumental variables methods to correct for the bias , researchers have employed a variety of different instruments and have generally found that instrumental variables produces estimated schooling returns that are larger, not smaller, than least-squares estimates. The connection between access to credit and measured returns to schooling–a connection first noted by Becker (1972)–has been explored recently by Lang (1993) and Card (1995a,2000), who argue that "discount rate bias" can explain the anomolously high instrumental variables estimates. This argument implicitely suggests that borrowing constraints are important for schooling decisions.

Our paper attempts to integrate and reconcile these two literatures. Building on the seminal work of Willis and Rosen (1979), we develop a framework that allows us to study schooling determinants and returns together. Identification of the effect of borrowing constraints from the fact that foregone earnings–the indirect costs of school–and the the direct costs of schooling each affect borrowing constrained persons differently from unconstrained individuals. We apply this idea in a number of ways using least-squares, instrumental variables regression, and a structural economic model to measure the extent of borrowing constraints on schooling choices. None of these methods produces evidence of borrowing constraints.

1 Introduction

Do borrowing constraints influence education outcomes of children? The importance of the answer to this question for the design of education policy and a number of economic phenomena goes without saying. Numerous studies by economists and sociologist have attacked this question almost exclusively through studies of the correlation between family income and other family characteristics and schooling attainment or specific levels of attainment, college entry in particular. The positive correlation between family income and schooling attainment has been widely interpreted as evidence supporting the idea that borrowing constraints hinder educational choices. However, the step from correlation to causation is a precarious one as family income is also strongly correlated with secondary school achievement and other characteristics associated with college entry.

Recent work by Cameron and Heckman (1998, 2000), Keane and Wolpin (1999), and Shea (1998) has attempted to better understand the determinants of schooling choices. Using very different methods, these researchers have found little evidence that favors the idea that borrowing constraints hinder college-going or any other schooling choice.

In the returns to schooling literature, Lang (1993) and Card (1995a, 2000) have argued that borrowing constraints produces a "discount rate bias" in estimated returns to education and can reconcile anomalously large instrumental variable estimates of the returns to schooling. This argument implicitely suggests that borrowing constraints are important for schooling decisions.[1]

Our paper attempts to integrate and reconcile these two literatures. Identification of borrowing constraints in all of our empirical approaches build on the following implication of the model. When educational borrowing constraints are operating, opportunity costs of schooling (forgone earnings) and direct costs (such as tuition) influence schooling choices differently for credit-constrained and unconstrained individuals. As direct costs need to be financed during school, they impose a larger burden on credit constrained students. Forgone earnings generally do not need to be financed during school and do not place a higher burden on credit-restricted students.

[1] We use the term "borrowing constraint" broadly. It does not necessarilly include only hard constraint. We interpret a borrowing rate higher than the market interest rate as a borrowing constraint.

We make three contributions in this paper. First, we present an empirically-tractable economic model of human capital accumulation under credit constraints. The model is novel in several ways made clear below. Second, we develop the econometric methodology to estimate the model. By articulating our behavioral assumptions with simple economic theory, we clarify the economic assumptions needed for empirical identification of credit constraints. We then formally demonstrate the manner in which these exclusion restrictions deliver identification in our structural model. We then use this theory to guide our econometric approach to estimating the model. The econometric methodology nests recent empirical work that studies heterogenous returns to education as a local average treatment effect problem (Angrist and Imbens, 1993) and extends these ideas into a more structural specification..

The third contribution of this paper is empirical. Working from the economic model, we present evidence on the importance of borrowing constraints from five different estimation techniques–linear regression estimates of returns to schooling, instrumental variable estimates of returns to schooling, a simple discrete-choice analysis of schooling determinants, and two structural econometric models. We discuss the instruments we use to identify both forgone earnings and the direct costs of schooling and show they are powerful predictors. Our findings are consistent across the approaches: we find no evidence of borrowing constraints.

It is important to highlight that we cannot address whether students have perfect access to credit. During the period covered by our data, there are large subsidies to school already in place in the United States. Given the policy regime, we find no evidence that credit constraints restricts investments in secondary school and college. Our evidence suggests that expansion of these subsidies will have little impact on overall schooling attainment.

This paper unfolds as follows. Section two provides a review of the literature. Section three presents the economic model on which we base our empirical strategy and discusses identification of borrowing constraints. Section four overviews the data, and section five presents empirical evidence on the question of borrowing constraints from linear regression and instrumental variable estimates of the returns to education and from regression and discrete-choice analyses of schooling attainment. Section six discusses estimation of the structural model. The model jointly imposes the implications of borrowing constraints on the schooling returns and schooling attainment choices. Empirical findings follow. The paper concludes with a summary.

2 Background and Significance

2.1 Literature on Educational Attainment Returns to Schooling

Few empirical studies have integrated educational attainment and returns to schooling in a coherent framework. An important exception is Willis and Rosen (1979). Our work extends Willis and Rosen's work in a number of ways. However, as most of the literature relevant to the question of educational financing constraints comes from the literature on determinants of educational attainment or from recent research on the returns to education.

A ubiquitous empirical regularity that emerges from the literature on determinants of schooling is that family income is strongly correlated with schooling attainment. This correlation has been found in legions of U.S. data sets covering the entire 20th century (see Hauser 1993, Kane 1994, Manski 1993, Manski and Wise 1983, Mare 1980, Cameron and Heckman 1998, and Cameron and Heckman, 2000, to name a few) and in data from dozens of other countries in all stages of political and economic development (see Blossfeld and Shavit, 1991, for instance). Educational financing constraints has been the most popular behavioral interpretation of the schooling-family income correlation.

However, limited credit access is only one of many possible interpretations of this correlation. Family income and family background measures has 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 elementary school through graduate school. Cameron and Heckman (2000 and 1998) adopt a 'life-cycle' view of the importance of family income and other family factors and find that family income is a prime determinant of the string of schooling-continuation decisions that lead to high school graduation and college entry. They conclude that the measured influence of the correlation between family income and college continuation is largely a proxy for its influence on earlier achievement. Shea's findings (1996) support this interpretation of the data. Shea examines unpredictable components of family income and finds little or no correlation with children's schooling and future earnings.[2]

Keane and Wolpin (1999) take a different approach. The estimate a rich discrete dynamic programming model of schooling, work, and savings. When they simulate the model, they find that relaxing borrowing constraints has almost no effect on schooling. However, borrowing constraints have important and significantly impacts work during school.

2.2 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 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 because they are smarter and earn a return on their higher ability as well as on their additional years education. Omitting measures of ability in a regression study of wages or earnings upward biaes estimated returns to schooling. Empirically, 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 both test scores and schooling levels in the regression leads to a decline in the measured effect of schooling. Nevertheless, scholastic test scores are imperfect measures of earning ability, and these measures leave substantial scope for bias from other unobserved components of ability.

To correct for ability bias, researchers have used instrumental variables techniques to recover returns to schooling purged of bias. Anomalously, instrumental variable estimates of the return to schooling have typically been found to be larger than their OLS counterparts. Lang (1993) and Card (1995a) explore borrowing constraints as one of several possible explanations for this pattern, which Lang terms "discount-rate bias." If returns to schooling vary across individuals because of differential access to credit for educational investments, the pattern of estimates produced by instrumental variables estimators may be explained by the fact that many of these estimators identify the causal effect of schooling for borrowing-constrained individuals who receive returns to schooling at the margin that are higher than

[2]Shea studies extracts from the Panel Study of Income Dynamics, and finds no effects in the full sample. He finds modest evidence of a relationship in the low-income subsample. This finding is not inconsistent with Cameron and Heckman's (2000) interpretation that family income effects operate at the earliest stages of schooling.

the population average.

Much recent work uses instrumental variables or related techniques to address the problems caused by omitted ability measures. Card (2000) provides an extensive survey of this literature. Contrary to the intuition of ability bias, he documents that researchers often find the coefficient on schooling rises rather than falls when instrumental procedures are used. Building on Becker's Woytinsky model (1972), Lang (1993) and Card (1995a) argue that "discount rate bias" is one potential explanation for this counter-intuitive result. In the standard Becker model, a student invests in schooling until her return from schooling human capital is equal to the interest rate she faces. If borrowing constrained individuals face a higher personal interest rate, the model implies they will have higher returns from schooling at the margin.

If there is no heterogeneity in the returns to schooling, instrumental variables provides a consistent estimate of the return. However, if there is heterogeneity in returns, then the IV estimate of the causal effect of schooling must be interpreted with care. Essentially, this story follows the logic of "local average treatment effects" exposited by Imbens and Angrist (1992). They show that IV measures the treatment effect of schooling (that is, the causal effect) for groups whose schooling decisions are most sensitive to changes in the instrument used in estimation. Suppose that borrowing constrained individuals are most sensitive in their schooling choices to changes in, say, college tuition. Because of their higher costs of borrowing funds for schooling costs, borrowing constrained individuals also demand the highest returns to continue. Thus, the IV estimate of schooling returns will be an average of returns for this group and will be higher than the the true population average of returns. Heckman and Vytlacil (1998) present a more complete description of the econometrics behind Card's (1995a) model.[3]

This argument may also explain estimates obtained from studies of "selection" models of schooling, which have focused on taking account of ability bias and not discount rate bias. These studies often find lower estimated returns to education in selection-corrected models (see, for instance, Willis and Rosen, 1979, or Taber, 1999). The argument for discount rate bias would not necessarily apply in these cases.

[3] Angrist and Krueger (2000) also embody the idea of discount rate bias into their econometric framework.

2.3 Basic Methodology

This paper takes both the returns to education literature and schooling determinants literature a step further. The hypothesis that borrowing constraints affect educational choice cannot be separated from other mechanisms in a study of returns alone or a study of schooling attainment that does not explicitly account for returns. This paper develops a structural model of schooling choice that nests both returns and school choices and allows us to estimate the importance of borrowing constraints.

Returns to education are assumed to be heterogenous. We introduce two types of exclusion restrictions for college costs–one for forgone earnings and one for direct college costs. A rise in the direct cost of college is more costly for borrowing constrained individuals, so when we relax these costs the individuals who are induced to attend college are more likely to be borrowing constrained. Thus we may suspect that when we use this type of exclusion restriction the measured returns to schooling will increase. On the other hand if we examine the opportunity cost of school (i.e. forgone earnings), there is no particular reason to expect these to be more important for borrowing constrained individuals than others. We use both instrumental variables methods and structural estimates to measure the extent of borrowing constraints.

We use these different sources of variation in costs to examine the effects of borrowing constraints on the returns to schooling using four econometric methods. (1) We use each type of cost as an instrumental variable and examine the differential impact on the returns to schooling. Since direct costs of schooling affect borrowing constrained individuals more strongly, we may expect that instrumenting for schooling using these costs will lead to IV estimates of the returns to schooling that are higher than OLS. On the other hand, there is no particular reason to expect much discount rate bias when we use forgone income as an instrumental variable. If the ability bias dominates in this case, the coefficient on IV would actually be lower than OLS. While one cannot tell with certainty which direction the bias in each case will go, IV models based on the direct costs of schooling should lead to a higher estimate of the return to schooling than with forgone earnings. (2) The story above relies on interactions between access to credit and costs of schooling. We can look for borrowing constraints directly by examining interactions between costs and observables that are likely to be related to borrowing constraints. (3) We then develop a structural version of the model and test more formally for evidence of borrowing constraints. We show that the identification of the difference interest rates comes through the type of interaction between costs and observables as in the previous method. (4) Finally we combine the methods using the structural model to look more formally for unobservable credit constraints. The basic intuition for identification is that as a result of the interaction, direct costs of schooling are a relatively more important determinant of schooling for individuals who face credit constraints. Thus the form of selection bias will be different for individuals who face different direct costs.

We follow Card (1995b) by using an indicator for whether there is a two- or four-year college in the individual's county as a measure of the direct costs of schooling. For students from families with low and moderate incomes, the ability to live at home while in college may lead to substantial savings. The data reveal that the probability of living at home while in college is about 55% for students with a college in their county, and only 34% for others. Further, even for students living on campus the transportation costs and the convenience of having parents close at hand may lead to a substantial financial advantage. As a measure of the foregone earnings of schooling, we use measures of income in low skill industries in the county of residence.

3 The Model

The model begins with a specification of individual utility. Individuals derive utility from both consumption and tastes for nonpecuniary aspects of schooling. Tastes represent the utility or disutility from school itself as well as preferences for the types of jobs available to high school graduates versus college graduates. Assuming agents have power utility over consumption in each period, lifetime utility for a given level of schooling 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 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 $(-\infty, 1)$. Defining the set of possible schooling choices by $\mathcal{S}$, individuals choose S out of this set so that

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

Much work on schooling including Becker (1972), Rosen (1977), Willis and Rosen (1979), Willis (1986), Lang(1994) and Card (1995), captures heterogeneity in credit access by differences in the rate of interest r at which a person can borrow and save. A credit constrained person faces a high r, which means educational financing is more costly.

An unattractive feature of this approach is that individuals facing a high r face high costs of educational financing but also receive high returns to savings after labor market entry. Thus, r influences demand for schooling directly by changing the price of educational financing and indirectly through a wealth effect that operates through the post-schooling return to savings. A higher return on savings raises demand for schooling, offsetting to some degree the negative effects of higher financing costs.[4] In this paper, we adopt the simple but novel assumption that individuals face a common market interest rate for all borrowing and lending once they enter the labor market, normalized for convenience such that $(1 + r)^{-1} = \delta$. Confining borrowing-rate heterogeneity to the schooling years is a natural assumption if one considers the borrowing rate to be determined by the ability to collateralize loans with personal or family assets during school.

Letting $R = 1 + r$, students maximize utility subject to the budget constraint,

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

where S is total years of school and I_s is the present value of income net of direct schooling costs.

[4]In the Becker-Woytinsky framework, this complication does not arise. Assuming a single r prevails throughout a persons lifetime gives rise to a separation result. Given r, individuals choose schooling to maximize the present value of lifetime earnings. Thus, in the Becker-Woytinsky model, as in our model, S enters directly as a determinant of the present value of income (see equation [5] below). Unlike the Becker-Woytinsky model, our assumption that borrowing and lending after labor market entry is transacted at the market interest rate means that S also enters as a determinant of the present value of consumption expenditures (equation [3] below) because consumption is discounted at the market rate after period S and at rate r before.

The first-order conditions of this problem are,

$$\begin{aligned} c_t &= (\delta R)^{\frac{t}{1-\gamma}} c_0 \qquad t \leq S, \\ c_t &= (\delta R)^{\frac{S}{1-\gamma}} c_0 \qquad t > S. \end{aligned}$$

Plugging these levels into the budget constraint yields

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

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

$$V_S = \frac{I_S^{\gamma} \left(\sum_{t=0}^{S-1} R^{\frac{t\gamma}{1-\gamma}} \delta^{\frac{t}{1-\gamma}} + (R\delta)^{\frac{S\gamma}{1-\gamma}} \sum_{t=S}^{\infty} \delta^t \right)^{1-\gamma}}{\gamma} + T(S). \tag{4}$$

Equation (4) is a conditional indirect lifetime utility function as it depends on the choice variable S.[5]

We next solve for the present value of income. To focus on borrowing constraints and abstract away from uncertain earnings we assume full certainty of earnings streams.[6] Let w_{tS} be earnings at time t for an individual with S years of schooling. Individuals have zero earnings while in school and pay tuition τ_t at time $t-1$ to attend schooling level t. Abstracting from labor supply, we have the following expression for the present value of income at $t = 0$:

$$\begin{aligned} I_S &= \left(\frac{1}{R}\right)^S \sum_{t=S}^{T} \delta^{t-S} w_{ts} - \sum_{t=0}^{S-1} \left(\frac{1}{R}\right)^t \tau_{t+1} \\ &\equiv \left(\frac{1}{R}\right)^S W_S - \sum_{t=0}^{S-1} \left(\frac{1}{R}\right)^t \tau_{t+1}, \end{aligned} \tag{5}$$

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

[5] It is no surprise at this point that a rise in R raises V_S. Holding S fixed, a rise in R lowers both the present value of earnings, I_S, and the present value of consumption expenditures. In the derivation above, I_S is fixed, so a rise in R means consumption expenditures must rise so its present value stays on par with I_S.

[6] Uncertainty in the returns to education introduces an option value to going to school even when predicted returns are low.

To illustrate the implications of the model, consider how changes in direct costs and opportunity costs affect schooling choices and their returns in a simple version of the model. Suppose there are only two levels of schooling, $S = 0$ and $S = 1$, and let τ_1 be the direct cost of $S = 1$. There are no nonpecuniary tastes for education and no direct cost of $S = 0$. The values of $S = 0$ and $S = 1$ are given by

$$\begin{aligned} V_0 &= \frac{W_0^{\gamma} \left(\frac{1}{1-\delta}\right)^{1-\gamma}}{\gamma} \\ V_1 &= \frac{(W_1/R - \tau_1)^{\gamma} \left(1 + (R\delta)^{\frac{\gamma}{1-\gamma}} \sum_{t=1}^{\infty} \delta^t\right)^{1-\gamma}}{\gamma} \end{aligned}$$

A person chooses $S = 1$ when $D_1 \equiv V_1 - V_0 > 0$ and $S = 0$ otherwise.

To clarify the intuition and explore the implications of heterogeneity in R, consider two individuals both indifferent between attending and not attending school; that is, $(V_0 = V_1)$. One person borrows at the market rate $(R = 1/\delta)$. The other is credit constrained and borrows at $R > 1/\delta$. Consider the each person's reaction at time zero to a dollar rise in foregone earnings, W_0, and alternatively to a dollar increase τ_1. For the person borrowing at the market rate, a dollar is a dollar: $V_0 = V_1$ implies $\delta W_0 = W_1 - \tau_1$. That is, a dollar rise in W_0 and a dollar rise in τ_1 have the same effect on the decision to pursue $S = 1$:

$$-\frac{\partial D_1}{\partial \tau_1} = \frac{\gamma V_1}{(\delta W_1 - \tau_1)} = \frac{\gamma V_0}{W_0} = \frac{\partial D_1}{\partial W_0}. \tag{6}$$

In order for the credit constrained student to be indifferent between schooling choices it must be that $W_0 > \frac{1}{R} W_1 - \tau_1$. This implies that the shadow value of a dollar during school is higher than it is after labor force entry:

$$-\frac{\partial D_1}{\partial \tau_1} = \frac{\gamma V_1}{\left(\frac{1}{R} W_1 - \tau_1\right)} > \frac{\gamma V_0}{W_0} = \frac{\partial D_1}{\partial W_0}. \tag{7}$$

If the constrained and unconstrained students have identical values of W_0 and the same utility functions (that is, γ is the same), then (6) and (7) imply the effect on D_1 of a dollar rise in foregone earnings is the same for both persons, but a dollar rise in τ_1 reduces D_1 more for the credit-constrained student:

$$-\frac{\partial D_1}{\partial \tau_1}\bigg|_{R>1/\delta} > -\frac{\partial D_1}{\partial \tau_1}\bigg|_{R=1/\delta}. \tag{8}$$

Because the shadow value of a dollar borrowed during school is higher for credit constrained students, it follows in turn that credit-constrained individuals choose $S = 1$ only when they receive higher returns to schooling than the marginal person borrowing at the market rate.

The main testable implications of the model can be stated more generally. Note first that a rise in R reduces the likelihood a person chooses $S = 1$ as long as she is not a net saver while in school:[7]

$$\frac{\partial D_1}{\partial R} = -\frac{\gamma V_1}{R I_1}(c_0 + \tau_1) < 0. \tag{9}$$

First, the borrowing rate, R, influences the schooling decisions through its interaction with direct costs τ_1 in V_1. A dollar rise in τ_1 diminishes the value of V_1 more for individuals with higher R. Thus, given that a person is not a net saver during school,

$$\frac{\partial^2 D_1}{\partial R \partial \tau_1} = -\frac{\gamma V_1}{R I_1^2}[(c_0 + \tau_1)(1-\gamma) + c_0 (R\delta)^{\frac{\gamma}{1-\gamma}} \frac{\delta}{1-\delta}] < 0. \tag{10}$$

Second, opportunity costs, W_0, operate only through V_0, so there is no interaction with R. Hence, a change in W_0 has the same influence on the schooling decisions of the constrained and unconstrained student:

$$\frac{\partial^2 D_1}{\partial R \partial W_0} = 0. \tag{11}$$

Third, the return to education for borrowing constrained persons is higher than it is for unconstrained and otherwise identical individuals. For studying returns to schooling using an instrumental variables (IV) framework, both the schooling choice S and the returns to S a person receives depend on R. In the population, returns to schooling is a random variable rather than a unidimensional parameter. Imbens and Angrist (1992) show that in this situation, IV recovers the average return on schooling for the subset of the population that is induced to change status by the change in the instrument.[8] More concretely, when one instruments for school choice using college costs, IV estimates recover the average return to college for the group of individuals who go to college when college costs fall. The model

[7] This is always the case when schooling costs are positive. If schooling is subsidized, a person is a net borrower while in school as long as $c_0 > -\tau_1$.

[8] In the language of Imbens and Angrist (1992), instrumental variables estimators converge to the expected treatment effect for those individuals induced to change status by the change in the instrument. In our model the treatment is school and the treatment effect is the returns to schooling.

above implies that the group that is induced to change will depend on whether it is the opportunity cost or direct costs of schooling that fall. Since direct costs have a relatively larger effect on borrowing constrained students, the group that changes with direct costs should have a higher concentration of borrowing constrained students. Among students who are close to indifferent about attending college, individuals with higher interest rates will also tend to have higher returns to college. This means that the IV estimate of the returns to schooling should be higher when direct costs are used as an instrument than when the opportunity cost is used. Using the language of Lang(1993) and Card (1996) the "discount rate bias" will be higher when direct costs are used as an instrument relateive to when opportunity costs are used. Below, we explore each of these implications of the model using OLS, IV, and a structural estimation method.

4 The Data

Our analysis is based on Black, Hispanic, and White males from the 1979-1994 waves of the NLSY (National Longitudinal Survey of Youth). We include only cases drawn as part of the random sample and exclude the military subsample and the non-Black non-Hispanic disadvantaged sample. We use only male samples because their schooling decisions and labor supply are less complicated by fertility and labor market participation considerations.

Because the NLSY collected detailed information about school attendance and completion starting from January of 1978, it is an ideal data set for our study. Schooling observations begin at age 15 and extend through age 24 for all individuals included in our sample. Because information before 1978 is retrospective and limited, we confine our extract to males between age 13 and 17 in January of 1978 in order to have reliable information on parental income and county of residence, which is used to construct a number of measures of labor market conditions and college proximity.

The NLSY Geocode data has annual measures of unemployment rates in county of residence. This variables showed little impact on schooling choices. Consequently, we merged with the NLSY a supplementary data set from the Bureau of Economic Analysis (BEA) containing detailed annual measures of regional labor market conditions (wages and employment) by industry. The BEA data are collected from state unemployment insurance records. From these data, we measured the opportunity cost of schooling as average local income ("local" is defined as the county of residence) in industries overrepresented by unskilled workers. Since the BEA data are reported by industry and not occupation, we use average earnings per job in service, agriculture, and the wholesale and retail trade industries. One large component of these industries is retail food establishments. For transitions into college, the set of industries for which average wages is constructed includes the aforementioned set together with manufacturing, construction, mining and extraction, and transportation and public utilities.

A panel of annual records on location of all public two-and four-year colleges (including Universities) in the Unites States were constructed from the Department of Education's annual HEGIS and IPEDS "Institutional Characteristics" surveys. By matching location with a person's county of residence, we were able to determine the presence of a two-year college, a four-year college, or either in the county of residence. A number of specialty colleges, generally with small enrollments less than 100, were excluded. Federal institutions, such as the Naval Academy, were also excluded. This variables can be measured at any age. For our analysis, however, they are measured at age 17 to avoid the obvious problem that people who attend college generally reside in the same county the college is located.

We present summary statistics of the main variables used in the analysis in Table 1.

5 Regression and Instrumental Variable Results

5.1 Methodological Issues

Our first empirical goal is to use the instrumental variables approach discussed above using the two costs of schooling as exclusion restrictions. There are a few problems that arise in implementation of our approach. The first potential problem is that since high school graduates leave school earlier, our data contain more observations on high school graduates than college graduates. We solve this problem by focusing on men of age 22 and older.[9]

The second problem that arises is the use of the local labor market variables. Our goal is to use the opportunity cost of college as an instrumental variable. We measure this by

[9]We have experimented with higher age cutoffs and find that the basic results are not sensitive to this choice.

using income in the county that the student lived at age 17. A problem with this variable is that labor market variables at age 17 are likely to be correlated with local labor market variables later in life. To deal with this potential problem we include a measure of the local income in the wage regression itself,

$$\log(w_{it}) = \beta_0 + S_i\beta_1 + \ell_{it}\beta_2 + X_{it}\beta_3 + u_{it},$$

where w_{it} is the hourly wage, S_i is schooling, and ℓ_{it} is a measure of the local income in the county in which individual i lives at time t. This seems to potentially solve the problem since it is reasonable to believe that conditional on current labor market conditions, the local labor market conditions at age 17 are unrelated to earnings directly. However the use of this variable leads to an additional potential problem. We have every reason to believe that moving is endogenous and is related to schooling outcomes itself. If this is the case, ℓ_{it} is also endogenous. A natural instrument in this case is the local income rate at time t in the county in which i lived at age 17. Since many individuals do not switch counties, or do not move far when they do, this instrument is strongly correlated with local income at t. Secondly, since it is determined by the county in which the student lives at age 17, it does not depend on moving after schooling completion so is not endogenous.

A third potential problem is the endogeneity of experience. We assume that experience is equal to age minus education minus six. However, as other authors have pointed out, if education is endogenous, then experience is endogenous as well. We instrument for experience and experience squared using age and age squared in some of our specifications.

5.2 First Stage Results

In this model the first stage regression is not particularly meaningful. Since we have more than one observation per individual in the wage data, we also have more than one observation per individual in the first stage. Secondly, since the the dependent variable, schooling, does not vary, but the local income measure does the regression is hard to interpret. In an attempt to convey the content of the first stage regressions we construct the average of the local income variable in the county at age 17 across the years in which the respondent works. We then regress schooling on our standard variables as well as this average, the local income at age 17, and the indicator for whether there is a college in the county. The results of these regressions are reported in Table 2.

In Column (1) we present the regression of schooling on a number of covariates including the dummy variable for a college in the county. This effect of this variable is large implying that individuals with a college in their county complete almost one half year more of school on average. The other covariates in the regression have signs and magnitudes similar to other work (see e.g. Cameron and Heckman, 2000). In column (2) we present results where we include the local income variable at age 17 in the regression. It has the expected sign, but is not significant. However, in column (3) we control for the mean local income over working life. The results here are as expected. The coefficient on this new variable is positive indicating that richer counties are likely to have more students attend college.[10] This leads the coefficient on the opportunity cost of school to be negative as expected. There are two potential explanations of this result, both of which seem to be important for the result. First, individuals could be more likely to attend school during temporary downturns. Second, individuals in counties in which the economy is improving could be more likely to attend college. In terms of the model above, we would expect both of these effects to influence schooling, and it is not important which. Intuitively, the first possibility is perhaps more appealing as a source of identification since it does not seem to embed essential features of the county itself. However, we will present some evidence below that we do not see a pattern when we look at observable ability measures, so it seems plausible that it isn't related to unobservable ability differences. In the final column we combine the two variables and show that both remain strong predictors.

5.3 IV Excluding Forgone Earnings

We next examine the effects of the returns to schooling when we use local labor market variable as an instrument for the returns to schooling in Table 3. Given that this variable does not necessarily have a larger impact on borrowing constrained individuals, we would not particularly expect the coefficient on schooling to increase when we use it as an instrumental variable. However, we see in the first three columns that the IV point estimate is

[10]Perhaps as a result of superior schools or peer effects.

substantially above the OLS estimate. One should keep in mind that the standard errors are large enough so that we cannot reject the hypothesis that the OLS estimates and the IV estimates are the same. In columns 4-6 we perform the same exercise but do not include test scores and family income in the model. The point estimates are higher in this case, but the pattern is quite similar. In experimenting with other specifications we see similar patterns. These results are similar to Arkes (1998) who uses state unemployment rates in a similar design and also finds IV estimates that are higher than the OLS estimates.

This result that the coefficients increase when excluding forgone earnings is not necessarily inconsistent with the notion of discount rate bias above. It may happen to be that the people at the margin tend to be borrowing constrained and that their marginal return exceeds the OLS estimate. If this were the case, then we would expect the coefficient to increase even more when we use the presence of a college as an instrument.

One potential problem with the model above in addressing schooling decisions during recessions is that there is no role for the family. In this case it may be that when local labor market conditions are poor, borrowing constrained families have more trouble sending their children to college. This would reverse the direction of the effect; that is schooling would increase during a boom particularly for children whose parents are borrowing constrained. If we incorporated this possibility into our model, the effects of local labor market conditions on schooling attendance would no longer be monotonic. We might expect borrowing constrained families to be more likely to send their children to college during a boom, while non-borrowing constrained families would be less likely (as a result of the foregone earnings). Given that we find a negative association between county income and schooling, this second effect appears to dominate. This makes these results perhaps more surprising. Suppose borrowing constrained families have higher marginal returns to schooling and that the students from these families are led to decrease schooling during a recession. This effect should lead to smaller values of the IV estimate. The fact that the IV estimate is higher than the OLS estimate appears even more counterintuitive if credit constraints were important. We formalize this argument in the appendix.

5.4 IV Excluding Direct Costs of Schooling

The results using presence of a local college are presented in Tables 5a and 5b. The first three columns present results that do not control for the local income variable. These estimates yield very large effects of the causal effect of schooling. However, controlling for the local income variable yields striking results. The estimates of the returns to schooling decline enormously so that they are very similar in magnitude to the OLS results. In Table 5b we present additional specifications. In the first set of results we control for the AFQT score from the NLSY instead of the four test scores. These give IV estimates that are substantially lower than OLS. When we include no controls for test scores or family income we get results similar to the last three columns of Table 5a. The fact that the results decline so much when we control for the local wages indicate that presence of a college seems to be positively correlated with wealth in the county. Most importantly, the IV point estimates are not higher than the OLS estimates.

Combining the results from the two tables we seem to see the opposite of what the borrowing constrained model would predict. Instrumenting with measures of direct costs of college rather than the opportunity cost does not lead to higher estimates of the returns to college. Thus we find no evidence that discount rate bias is important in this exercise.

5.5 Validity of the Instruments

In general, without a maintained assumption that one of our instruments is valid, it is impossible to test them. In addition, since we allow the treatment effect to be random, a standard over-identification test will not work. However, it is often informative to examine the relationship between the excluded variables and the observables in the wage equation. While the lack of a relationship between observable measures of ability and the instruments does not prove that there is no relationship between unobservable measures of ability and them, it does lend some credence to their use.

In Table 6 we present regressions of the two instruments on various observables. The dummy variable for presence of a college in the county is the dependent variable in the first column. The results in this regression are somewhat discouraging. The math score from the ASVAB test does seem to be positively related to whether there is a college in the county. Given the importance of this variable in both the schooling equation and the wage equation, this seems potentially problematic. Fathers education and the automotive knowledge score also are related to this instrument. This result makes our finding even more surprising. Presence of a college seems to be positively correlated with observable ability. If it were positively correlated with unobservable ability as well then the schooling estimate is biased upwards.[11] The good news however is that conditional on the average wage during working life, test scores do not help predict the local labor market variable at age 17. These results are shown in the second column of Table 6. While certainly not conclusive, this result is favorable to the use of the local labor market variation as an instrument.

The fact that test scores help predict the presence of a college in the county is not necessarily problematic, since we control for these variables in our regression. Table 7 presents results that seem more promising for the use of the college in county as an instrumental variable. In this table we present results from probit models of our standard set of regressors (from e.g. Table 2) on college attendance and high school dropout status. If the college in county variable is exogenous it should have a much stronger effect on college attendance than dropping out of high school.[12] However, if college in county is picking up "pro-schooling" aspects of the community rather than just the reduction in the cost of college, then we would expect it to predict high school graduation as well. Our results in Table 6 show that presence of a college in the county has a large effect on college attendance but a statistically insignificant effect on high school graduation. While the point estimate of the dropout effect is not zero, these results seem to suggest that the instrument is affecting school choice as we would have expected. Again while this is not conclusive, it does suggest that the instrument is working in the manner we would expect if it were a valid instrument.

[11] Altonji, Elder, and Taber (2000) provide a model that justifies this type of argument.

[12] It still may have some effect on high school dropout status by changing the option value of high school graduation.

5.6 Interactions between Observables and Direct Costs

The argument about identification above relied on the relationship between access to credit and the direct costs of schooling. We used the model to show that the direct costs of schooling place a larger burden for individuals who are credit constrained. If we could observe individuals in the data who were credit constrained we could test this directly since we expect the presence of a college in their county to have a larger impact on their schooling decisions. While we cannot observe credit constraints directly, we can observe some variables that we would expect to be related. In particular we can look for interactions between the presence of a college and race, family background, and family income variables. Card (1995b) and Kling (1999) also look for these interactions in the National Longitudinal Survey of Young Men and show that individuals from low family background are affected more by the presence of a college. Our results are different.

In Table 7 we present the results of schooling regressed on our standard set of covariates and interactions between the presence of a college and different variable. The first column interacts the dummy variable with racial dummies for black and Hispanic. Since it is well known that blacks tend to have less wealth for the same amount of income, it might make sense to expect blacks to be credit constrained. In fact we find that the interaction is negative showing that the presence of a college seems to be more important for whites than for blacks (although it is statistically insignificant). The sign on the Hispanic interaction is also the opposite of what we would expect if they had less access to credit. In the second column we interact the dummy variable with parents education. In this case the interaction with father's education is positive and the interaction with mother's education is negative. It is hard to interpret this as evidence of credit constraints. In the next two columns we present the results for family income and number of siblings. Once again the interactions are insignificant and of the wrong sign. We see essentially no evidence of credit constraints in this table.

There are a number of reasons why our results may differ from Card (1995b). Perhaps the largest difference is that our cohorts are much younger than theirs. Our results are consistent with the idea that credit constraints may have been important when the NLS Young Men were making their schooling decisions in the 1960's, but much less important when the NLS Youth were making their decisions in the early 1980s. Of course, the data sets differ in a number of other ways so this may not be the explanation. Discovering whether this is actually the case is an important avenue for future research.[13]

6 Evidence from Structural Model

6.1 The Econometric Model

In this section we extend the model above into a specific econometric framework which we take directly to the data. Let wages take the form,

$$\log(w_{sit}) = \gamma_S + X_{Wi}^{'}\beta_W + X_{\ell it}^{'}\beta_\ell + E_{it}\beta_E + E_{it}^2\beta_{E^2} + u_{sit},$$

where for individual i, E_{it} is experience, $X_{\ell it}$ consists of local labor market variables at time t, X_{Wi} is a vector of non-time varying variables that influence wages, and u_{sit} represents an error term that is orthogonal to the other covariates. If indidual i chooses shooling level s, his present value of earnings dated at the time he leaves school is,

$$\sum_{t=S}^{\infty} \delta^{t-S} e^{\gamma_S + X_{Wi}^{'}\beta_W + X_{\ell it}\beta_\ell + E_{it}\beta_E + E_{it}^2\beta_{E^2} + u_{sit}}.$$

We abstract from uncertainty by assuming that students choose schooling to solve the certainty equivalence problem. Let E_S denote expectation conditional on information known at time S. Decisions are made as above conditioning on,

$$I_{si} = \left(\frac{1}{R_i}\right)^S e^{\gamma_S + X_{Wi}^{\prime}\beta_W} \left(\sum_{t=S}^{\infty} \delta^{t-S} e^{E_{it}\beta_E + E_{it}^2\beta_{E^2} + E_S\left(X_{\ell it}^{'}\beta_\ell\right) + E_s(u_{it})}\right) - \sum_{t=0}^{S-1} \left(\frac{1}{R_i}\right)^t X_{ic}^{\prime}\beta_c$$

[13]In results not reported in the version of his paper that we cite, Kling also ran some regressions using the NLSY data. In contrast to us he found results similar to Card in a relationship between poor family background and schooling. Kling kindly provided us with the code he used to generate his data and we attempted to uncover the difference. Even though the variables we used were very similar, the difference in the results seem to come from both the differences in the definition of college in county and the fact that we are using a younger sample. Both Card (1995b) and Kling(1997) focus on parents education in these interactions. When we used Kling's variable

and interact it with race, number of siblings, and family income as in our specification we get found no interactions (except for race which goes in the unexpected direction as in our results). We also repeated our IV estimates described above using Kling's measure of presence of a college. It yields point estimates that are very similar, but substantially higher standard errors.

where R_i represents the individual borrowing rate (which we will discuss below) and we have parameterized the costs of schooling (τ_t in the notation above) to $X_{ic}^{'}\beta_c$.[14]

We further simplify the model by assuming that we can write $u_{sit} = \theta_{si} + \omega_{it}$ where θ_{si} is known to the student during schooling, but ω_{it} is orthogonal to information during school so that $\theta_{si} = E_S(u_{sit})$. We approximate $E_S(X_{\ell it}^{\prime}\beta_\ell)$ by a linear function of local labor market variables known to the students when schooling decisions are made. We denote these variables by $X_{\ell si}$ with coefficient vector $\widetilde{\beta}_{\ell S}$. To simplify the notation we incorporate an intercept into $\widetilde{\beta}_{\ell S}$ so that,

$$X_{\ell si}^{\prime}\widetilde{\beta}_{\ell S} = E_S\left(X_{\ell it}^{\prime}\beta_\ell\right) + \gamma_S + \log\left(\sum_{t=S}^{\infty}\delta^{t-S}e^{E_{it}\beta_E + E_{it}^2\beta_{E^2}}\right).$$

We also simplify the model by using log utility.[15] Solving for the value of schooling under log utility, and plugging in our solution for I_{si}one can show that,

$$\begin{aligned} V_{si} &= \left(\frac{1}{1-\delta}\right)\log\left(I_{si}\right) + \left(\frac{1}{1-\delta}\right)\log\left(1-\delta\right) + \left[\sum_{t=0}^{S-1}\delta^t t + \left(\frac{\delta^{S+1}}{1-\delta}\right)S\right]\log\left(\delta R_i\right) + X_{Ti}^{'}\beta_{TS} + \nu_{si} \\ &= \alpha_1 \log\left(\left(\frac{1}{R_i}\right)^S e^{X_{Wi}^{'}\beta_W + X_{\ell si}^{'}\widetilde{\beta}_{\ell S} + \theta_{si}} - \sum_{t=0}^{S}\left(\frac{1}{R_i}\right)^t X_{ci}^{'}\beta_c\right) \\ &\quad + \alpha_2(S) + \alpha_3\left(S\right)\log(R_i) + X_{iT}^{'}\beta_{TS} + \nu_{si}, \end{aligned}$$

where $X_{Ti}^{'}\beta_{TS} + \nu_{si}$ represents nonpecuniary benefits/costst associated with schooling level s. Notice that this expression is very close to a standard linear index model. The only nonlinearity arises from the term inside the logarithm. From this expression one can see that there is in an interaction between interest rates R_i and costs $X_{ic}^{'}\beta_c$, but that there is no such interaction between interest rates and foregone earnings. This aspect of the model delivers identification of the parameters of interest as we demonstrate in the next section.

6.2 Sketch of Identification

To get a general feel for identification, once again assume that there are two schooling choices, levels 0 and 1, with no direct costs for schooling level 0. In the implementation of

[14]Notice that we have assumed that costs depend only on observables. This was chosen for computational convenience, but we see no reason why this should bias the findings in either direction.

[15]We see no reason why this should bias the results in any particular direction.

the model we actually use four levels of schooling, but this only aids identification.

$$
\begin{aligned}
V_{0i} &= \alpha_1 \left(X_{Wi}^{'}\beta_W + X_{\ell 0i}^{'}\widetilde{\beta}_{\ell 0} + \theta_{0i} \right) + a_1 \\
&\quad + X_{Ti}^{'}\beta_{T0} + \nu_{T0i} \\
V_{1i} &= \alpha_1 \log(e^{-\log(R_i) + X_{Wi}^{'}\beta_W + X_{\ell 1i}^{'}\widetilde{\beta}_{\ell 1} + \theta_{1i}} - X_{ci}^{'}\beta_c) + a_2 \\
&\quad + a_3 \log(R_i) + X_{Ti}^{'}\beta_{T1} + \nu_{T1i}.
\end{aligned}
$$

We estimate two versions of the model. The first allows for heterogeneity in R_i by letting it depend on particular observables (such as race, family income, etc.) as in Section 5.5. The second treates R_i as an unobservable and estimates its distribution.

To understand the source of identification in the first case, suppose that there was no unobserved heterogeneity ($\theta_{si} = 0$), that we did not condition on local labor market variables or the wage variables $\left(i.e. \text{ set } X_{Wi}^{'}\beta_W + X_{\ell 1i}^{'}\widetilde{\beta}_{\ell 1} = 0 \right)$, and that the proxy for R_i is included as one of the taste covariates. We could parameterize $-\log(R_i) = X_R'\beta_R$. Since $X_{ci}^{'}\beta_c$ varies only with the presence of a college in the county, testing $\beta_R = 0$ would be equivalent to testing for an interaction between the presence of a college and X_R in a binary choice model. This is virtually identical to our approach in Section 5.5, but one can see from the model that our analysis in the previous section was incorrect. This specification demonstrates that when looking for interactions, we should control for the wage variables and the local labor market variables in the manner shown above. Since there is no selection in this model, β_w and $\widetilde{\beta}_{\ell 1}$ can be estimated from the wage equation (with some assumption about how agents predict future local labor market variables).

Now consider the more complicated case in which R_i is unobserved. Taking the difference and combining the parameters, we obtain the latent variable representation,

$$
\begin{aligned}
Y_i^* &= V_{1i} - V_{0i} \\
&= \alpha_1 \log(e^{X_i \Gamma_1 + \varepsilon_{1i}} - X_{ci}^{'}\beta_c) + X_{i,\ 2} + \varepsilon_{2i},
\end{aligned}
$$

where,

$$\begin{aligned}
X_{i,\ 1} &= X_{Wi}'\beta_W + X_{\ell 1i}'\widetilde{\beta}_{\ell 1} \\
\varepsilon_{1i} &= \theta_{1i} - \log(R_i) \\
X_{i,\ 2} &= X_{Ti}'\beta_{T1} - \alpha_1 \left(X_{Wi}'\beta_W + X_{\ell 0i}'\widetilde{\beta}_{\ell 0} \right) + a_2 - a_1 - X_{Ti}'\beta_{T0} \\
\varepsilon_{2i} &= a_3 \log(R_i) + \nu_{T1i} - \alpha_1 \theta_{0i} - \nu_{T0i}.
\end{aligned}$$

Individuals choose schooling option 1 if $Y_i^* > 0$, and choose option 0 otherwise.

We assume that (X_i, X_{ci}) is observable while $(\theta_{1i}, \theta_{0i} \log(R_i), u_{T1i}, u_{T0i})$ is unobservable and independent of the observables. Looking at the selection equation we have essentially three indices,

$$\left(X_{i,\ 1}, X_{i,\ 2}, X_{ic}'\beta_c \right),$$

and two error terms $(\varepsilon_{1i}, \varepsilon_{2i})$. These three degrees of freedom allow us to trace out this joint distribution.

However, the schooling equation alone is not sufficient for identification of the distribution of R_i as one can not separate it from θ_{1i}. Additional data on wage earnings is needed to distinguish between these error terms. In addition we have data on wages where,

$$\log(w_{i1t}) = \gamma_S + X_{iW}'\beta_W + X_{i\ell t}'\beta_\ell + E_{it}\beta_E + E_{it}^2\beta_{E^2} + u_{1i}.$$

Theorem 1 in the Appendix demonstrates that we can identify the joint distribution of $(\varepsilon_{1i}, \varepsilon_{2i}, u_{1i})$ using the two types of exclusion restrictions. Given this joint distribution we can identify the conditional expectation of $E\left(u_{1i} \mid \varepsilon_{1i}\right) = E\left(\theta_{1i} \mid \theta_{1i} - \log(R_i)\right)$. We also show in the appendix that we can use this conditional expectation to estimate the distribution of R_i when R_i is independent of θ_{1i}. Thus we can identify the distribution of interest rates faced by students up to a normalization that we discuss in the appendix.

The identification result crucially depends on the fact that we have two different types of exclusion restrictions: local labor market variables that affect students school choices but not w_{i1t}, and the presence of college which influences only costs. Using the notation above, the first type of exclusion restriction affects $X_{i,\ 2}$, but not $X_{i,\ 1}$or $X_{ic}'\beta_c$and the second enters $X_{ci}'\beta_c$ but not $X_{i,\ 2}$ or $X_{i,\ 1}$. This gives us essentially two degrees of freedom which allow us to trace out the two dimensional distribution $(\varepsilon_{1i}, \varepsilon_{2i})$.[16] Combining this information with the wage data allows us to separate R_i from $\theta_{1i} - \log(R_i)$ since only θ_{1i} influences wages.

The basic intuition for identification in practice comes from examining the dependence of the form of selection bias on whether there is a college in one's county or not,

$$E\left(\theta_{1i} \mid \alpha_1 \log(e^{X_i\Gamma_1+\theta_{1i}-\log(R_i)} - X_{ci}^{'}\beta_c) + X_{i,\ 2} + \varepsilon_{2i}, X_{i,\ 1}, X_{ci}^{'}\beta_c, X_{i,\ 2}\right).$$

In general this expression is complicated and will depend on the full joint distribution of $(\theta_{1i}, R_i, \varepsilon_{2i})$. As an example, suppose that interest rates are independent of the other error terms. Following exactly the same logic as in section 3 above, because the $\log(\cdot)$ expression is nonlinear the term $\theta_{1i} - \log(R_i)$ will be relatively more important when costs of schooling are high which corresponds to no local college. If borrowing constraints were not important this means that since θ_{1i} is relatively more important when a college is not present, there will be motr selection bias in these counties. However, if R_i is relatevly more important than θ_{1i},then when costs are high schooling decisions will be dictated by the extent of borrowing constraints. This would imply that there is less selection bias when no college is present. This basic intuition is extremely close to that for identification in the Instrumental Variables model.

6.3 Evidence from observables

We parameterize interest rates so that,

$$R_i = \exp\left[\exp\left(X_{Ri}'\beta_R + \log\left[-\log(\delta)\right]\right)\right].$$

where X_{Ri} are variables that may determine the borrowing rate. This functional form was chosen for two reasons. First, R_i is restricted to reasonable ranges, it will be strictly greater than one (but can be arbitrarily large). Second, it facilitates comparison betweem borrowing constrained individuals with others that are not constrained. In this specification when $X_{Ri}'\beta = 0$, R_i is equal to the market interest rate and these individuals are not

[16] The appendix requires full support of $X_{ci}\beta_c$. This is clearly not true in our data, but our goal is simply to test for the presence of borrowing rate heterogeneity not to identify the full joint distribution. The goal of the formal identification section is to demonstrate the manner in which multiple types of exclusion restrictions can facilitate identification.

constrained. We do not include an intercept in X'_{Ri},so we can restrict a certain subgroup of the population to be able to borrow and lend freely by setting the value of X_{Ri} to zero for those individuals. For example we will test for racial inequality in access to credit by using dummy variables. In this case X_{Ri} consists of only a dummy variable for black and a dummy variable for Hispanics. Thus for whites both dummy variables are zero, which imposes that for them $R_i = \frac{1}{\delta}$.

In this section we ignore the selection problem so that in the notation above $\theta_{si} = 0$ and ν_{Tsi} is independent of u_{it}. We relax this restriction in the next subsection.

We estimate the β_W parameters by OLS and use the estimated value of this parameter in estimation (the standard errors are adjusted appropriately). The cost index, $X'_{ci}\beta_c$, consists of only the distance from college and a constant. The vector of taste variables, X_{Ti} contains most of the variables entering the schooling decisions as in Table 2 above.[17] The local labor market variables enter the model through the index $X'_{\ell si}\widetilde{\beta}_{\ell S}$. We incorporate two measures of local labor markets into $X_{\ell si}$. The first is the income at age seventeen and eighteen of the county in which the individual lived at age 17. This is restricted to affect only the earnings 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. This variable enters all four equations.

We assume that there are four levels of schooling, high school dropouts(S=0), high school graduates(S=2), some college(S=4), and college graduates(S=6). During high school there are no direct pecuniary costs of schooling, but they must be incurred for students who attend some college and for college graduates. We also assume that ν_{Tsi} has a Generalized Extreme Value distribution so that the model can be estimated as a nested logit. We use two levels of nesting first nesting all high school graduates together and then nesting college attenders together. The observable components of the utilities take the

[17] The college in county indicator and the local labor market variables are not included in X_{Ti}.

form,

$$\begin{aligned}
\mu_{i0} &= \alpha_1\left(X'_{Wi}\beta_W + X'_{\ell 0i}\widetilde{\beta}_{\ell 0}\right) + \alpha_2(0) \\
\mu_{i2} &= \alpha_1\left(-2\log(R_i) + X'_{Wi}\beta_W + X'_{\ell 2i}\widetilde{\beta}_{\ell 2}\right) + \alpha_2(2) + \alpha_3(2)\log(R_i) + X'_{Ti}\beta_{T2} \\
\mu_{i4} &= \alpha_1\log\left(\left(\frac{1}{R_i}\right)^4 e^{X'_{iW}\beta_W + X'_{\ell 4i}\widetilde{\beta}_{\ell 4}} - \sum_{t=2}^{3}\left(\frac{1}{R_i}\right)^t X'_{ic}\beta_c\right) + \alpha_2(4) + \alpha_3(4)\log(R_i) + X'_{Ti}\beta_{T4} \\
\mu_{i6} &= \alpha_1\log\left(\left(\frac{1}{R_i}\right)^6 e^{X'_{Wi}\beta_W + X'_{\ell 6i}\widetilde{\beta}_{\ell 6}} - \sum_{t=2}^{5}\left(\frac{1}{R_i}\right)^t X'_{ci}\beta_c\right) + \alpha_2(6) + \alpha_3(6)\log(R_i) + X'_{Ti}\beta_{T6}
\end{aligned}$$

The nesting gives the following schooling probabilities,

$$\begin{aligned}
\Pr(S &= 6 \mid S > 2, \mu_{i0,\ldots,}\mu_{i6}) = \frac{\exp\left(\frac{\mu_{i6}}{\rho_c}\right)}{\exp\left(\frac{\mu_{i6}}{\rho_c}\right) + \exp\left(\frac{\mu_{i4}}{\rho_c}\right)} \\
\Pr(S &= 2 \mid S > 0, \mu_{i0,\ldots,}\mu_{i6}) = \frac{\exp\left(\frac{\mu_{i2}}{\rho_H}\right)}{\exp\left(\frac{\mu_{i2}}{\rho_H}\right) + \left[\exp\left(\frac{\mu_{i4}}{\rho_C}\right) + \exp\left(\frac{\mu_{i6}}{\rho_c}\right)\right]^{\frac{\rho_C}{\rho_H}}} \\
\Pr(S &= 0 \mid \mu_{i0,\ldots,}\mu_{i6}) = \frac{\exp(\mu_0)}{\exp(\mu_{i0}) + \left(\exp\left(\frac{\mu_{i2}}{\rho_H}\right) + \left[\exp\left(\frac{\mu_{i4}}{\rho_C}\right) + \exp\left(\frac{\mu_{i6}}{\rho_c}\right)\right]^{\frac{\rho_C}{\rho_H}}\right)^{\rho_H}}.
\end{aligned}$$

The model is estimated by maximum likelihood restricting ρ_H and ρ_C betweem zero and one. This nested logit can also be interpreted as incorporating uncertainty in tastes for schooling. In this case it has the form of a forward looking discrete choice dynamic programming model (see Taber 2000).

There are a number of normalizations necessary for identification. We assume that $\beta_{T,S}$ is zero for dropouts which normalizes the location as in a standard in polychotomous choice models. In the high school equations we can not separately identify the intercepts in β_{Ts} from the intercept in $\beta_{\ell s}$ so we set the latter to zero. In a feature that is more unique to our model we normalize the β_C coefficient on distance from college to one in the equations for college. This normalization is needed because we cannot separately identify the scale of the indices from the intercept in the tastes for schooling index.[18]

It turned out that in practice the intercept in the cost equation was very difficult to identify empiricallyso we fixed it to two in the simulations we present.[19] We experimented

[18]Normalizing this coefficient seems to work better computationally than the alternatives.

[19]The idea being that by living at home one can cut the cost of schooling in half.

extensively with other values for this parameter and the results are not sensitive. In practice, we also had trouble identifying the coefficient on the local earnings variable in the two college states. A reduced form regression yields a value of approximately 0.3 for it, so we fixed it to this value for the last two schooling equations. We again explored the sensitivity of our results to this assumption by trying alternative values and found our results to be very robust. We have also assumed that $\delta = 0.97$ which yields $R = \frac{1}{\delta} \approx 1.031$.

The log-likelihood function for the model estimated without heterogeneity in access to credit is -2547.69. The parameters are presented in Table 8. For the sake of brevity we do not provide a general discussion of the parameters, but they seem reasonable to us.

The first case we examine is heterogeneity in interest rates across racial groups. As mentioned above we include two dummy variables, one for African Americans and one for Hispanics. The results of this model are presented as the second group of results in Table 9. The point estimates go to the boundary in the unexpected direction and in the pseudo-likelihood ratio test[20] we see no evidence that these are significant.

We next examine parental education. In this case we include variables for father's education and mother's education. These are presented as the third set of results in the table. Once again, with a likelihood ratio test we cannot reject the null hypothesis of no heterogeneity in interest rates. In this case the point estimates go in the expected direction for mothers education, but in the unexpected direction for fathers. The parameter for fathers seems to dominate in the simulations. Our next set of results interacts family income with access to credit. In this case we divide the data into three groups on the basis of their family income (based on their parent's income). We restrict the top third to not be borrowing constrained and estimate different rates for the other two groups. In this case the middle third has a slightly higher rate, the lower group has a somewhat higher rate, but the results once again are statistically insignificant. Perhaps the case in which one would expect the strongest evidence we have for heterogeneity in access to credit comes from the interaction with the number of siblings in the family. In this case, we find that the borrowing rate is extremely insensitive to the number of children in the family.

[20]We call it a "pseudo likelihood ratio test" because we have included estimated parameters in the procedure it is not valid. Correcting the standard errors for the first stage estimates made very little difference, so we suspect that to be the case for the likelihood ratio test as well.

6.4 Evidence of Heterogeneity in Unobservable Borrowing Rates

Our next approach is to search for evidence of heterogeneity in unobserved borrowing rates. We assume there are two types of individuals, those that are unconstrained and borrow at the market rate $R^1 = 1/\delta$,and those that are constrained and borrow at a higher rate $R^2 > 1/\delta$. We estimate both this higher rate (R^2) and the fraction of the population of each type (P_R). We assume the distribution of borrowing rates is independent of the other error terms and the observables.

We allow for heterogeneity in ability which enters through a single normally distributed factor θ_i which is assumed to be known during school. The error term in the wage equation is defined as,

$$u_{sit} = \phi_{ws}\theta_i + \omega_{sit},$$

where ω_{sit} is orthogonal to θ_i. Ability may also be correlated with unobserved tastes for schooling which are now defined as,

$$\nu_{si} = \phi_{Ts}\theta_i + \widetilde{\nu}_{si},$$

where $\widetilde{\nu}_{si}$ has a GEV distribution yielding a nested logit model as in the previous section.

A major goal of our work is to use the argument for identification to guide the implementation of the model. This suggests that identification of the distribution of R_i should come from the interaction between interest not from the functional form assumptions about the distributions of the error terms. This leads us to two nonstandard strategies. First we have been very flexible about the form of selection bias. Ideally we would leave the joint distribution of $(v_{2i}, v_{4i},\ v_{6i}, \omega_{0i}, \omega_{2i}, \omega_{0i}, \omega_{2i})$ completely unrestricted. This full joint distribution would be difficult to identify and would oftern lead to a singular hessian so that standard errors can not be calculated. However, since our goal is to estimate the distribution of R_i, we can still perform likelihood ratio type tests to test for borrowing constraints so constructing standard errors is not necessary. Full nonparametric estimation of the seven dimensional distribution was not compuationally feasible. Even though we have restricted the error terms to be normal with a one factor representation, there are still 7 parameters that determine the form of the selection bias: the corresponding seven factor loading terms on θ_i, even after restricting those further we could not obtain standard errors. We have purposely chosen to overparameterize the form of the selection bias in order to force identification to come from the interaction between the bias and the college in county variable.

Our second nonstandard strategy came from our desire to force identification of the distribution of R_i to come from the form of selection bias as was suggested in the identification section above. We originally attempted to estimate the model by maximum likelihood. However, after extensive experimentation we have concluded that this does not work well. If one could be sufficiently flexible in the distribution of all of the error terms we believe this approach may work, but we found that allowing for this type of flexibility was not feasible and results were very sensitive to the precise functional forms that we used. In particular, for computational reasons we could not relax the distribution of ω_{si} sufficiently so that we could really distinguish the distribution of R_i from a more flexible functional from for the school choice equation. This is not at all in the spirit of our identification result above in which we showed that nonparametric identification of the distribution of θ_i comes from the form of the selection bias in the wage equation. We use an iterative procedure to force identification of the distribution of R_i and other parameters to come from the variation in data that we prefer to use. In particular, we partition the parameters of our model to those we want to identify from the selection equation (Ψ_1) and those that we want to identify in the wage equation (Ψ_2). Denote all observables by X_i and let $\pounds_i\left(\Psi_1, \Psi_2, X_i\right)$ be the log likelihood from the schooling choice and $E(\log(w_{it}) \mid X_i, \Psi_1, \Psi_2)$. We esimate the model using the following iterative procedure,

1. Fix Ψ_2 and solve for the value of Ψ_1 that maximizes the likelihood of the school choice model.

2. Fix Ψ_1 and solve for the value of Ψ_2 that minimizes the nonlinear least squares,

$$\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left(\log(w_{it}) - E(\log(w_{it}) \mid X_i, \Psi_1, \Psi_2)\right).$$

We iterate this procedure until convergence.[21]This procedure is equivalent to a GMM

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

model where we use the moment conditions

$$E\left[\begin{array}{c}\frac{\partial \pounds_i(\Psi,X_i)}{\partial\Psi_1} \\ \frac{\partial E(\log(w_{it})|X_i,\Psi_1,\Psi_2)}{\partial\Psi_2}\left(\log(w_{it})-E(\log(w_{it})\mid X_i,\Psi_1,\Psi_2)\right)\end{array}\right]=0.$$

The decision of which parameters to use in which equation was straight forward. The parameters of the wage equation we estimated from the regression model and the parameters of the tastes for schooling we estimated using the schooling choice model. Following the logic of the identification result we want the distribution of the borrowing rates to come from the wage equation, so the distribution of R_i was estimated from the regression model. Following the first strategy above, the goal is to keep the form of the selection bias flexible so we estimate the factor loading terms on $\theta_i$in the wage equation as well. That led us to the following partition,

$$\begin{aligned}\Psi_1 &= \left(\alpha_1,\rho_H,\rho_c,(\beta_{Ts,}\widetilde{\beta}_{\ell s}),s\in\{0,2,4,6\}\right)\\ \Psi_2 &= \left(\beta_W,\beta_\ell,\beta_E,\beta_{E^2},P_R,R^2,(\gamma_s,\phi_{Ts},\phi_{ws}),s\in\{0,2,4,6\}\right).\end{aligned}$$

We experimented with a wide variety of specifications and starting values. We found that the model did not typically converge when we put no restrictions on the ϕ_{Ts} parameters (factor loading on heterogeneity in taste for schooling). We restrict the model so that the taste for dropout and high school are the same ($\phi_{T0}=\phi_{T2}$) and the taste for the two schooling types are the same ($\phi_{T4}=\phi_{T6}$). Again this does not appear to affect the basic result but does aid convergence substantially.[22]

The main result can be seen in the first two rows of Table 10. The model becomes degenerate finding that no individuals are borrowing constrained. We estimated the model using a large number of starting values and a number of different specifications and consistently found no evidence of heterogeneity in borrowing rates. Even our simplified parametric modelisflexible enough so that the Hessian that is approximately singular. As a result the standard errors were not at all accurate so we do not present them in this table. Even if we could obtain accurate standard errors of the other parameters, we could not obtain them for the key parameters involved in interest rate heterogeneity since we are at a corner so we are not particularly concerned about this problem. While one could in principle use

[22]We also explored a linear specification for this parameter that also gave the same basic result.

the GMM criteria function to formally test whether there is evidence of heterogeneity in borrowing rates, since we are at a corner we know this test would not reject. Once again we find no evidence of borrowing constraints.

The wage parameters yield a strange pattern in that "some college" workers tend to have low wages although the schooling effects are monotonic otherwise. It is somewhat hard to interpret the results on the factor loading terms. The four factor loading terms are positive as one would expect, although the point estimate of the college graduate factor loading term is substantially smaller than the others which is somewhat interesting. To get an idea of the precision to which these parameters are estimated we ran a simple OLS regression on the selection terms and the estimates of $E(\theta_i \mid S_i, X_i; \Psi)$ ignoring the fact that we simultaneously estimated the parameters Ψ. By design this yields exactly the same point estimates as in Table 10 and yields standard errors on the factor loading terms (ϕ_{Ws}) of approximately 0.12. The rest of the parameters of the model are similar to our estimates without heterogeneity.

7 Discussion and Conclusions

This paper develops a model showing that direct costs of schooling affect borrowing constrained individuals in a fundamentally different way than opportunity costs of schooling. It then uses these two sources of variation to test for borrowing constraints in four different ways, 1) instrumental variables, 2) regression models allowing for interactions between observables and presence of a college, 3) structural estimation allowing borrowing rates to depend on observables, 4) structural estimation allowing for unobservable heterogeneity in borrowing rates. We find no evidence of borrowing constraints using any of the methods.

One issue that inevitably arises with this type of finding is power. Have we found no evidence of borrowing constraints because there are not borrowing constraints or because we do not have enough precision to measure them? As is virtually always the case, this is hard to address because it depends on ones prior beliefs about the size of the effect. However, we will speculate a bit in an attempt to address this possibility. In our view this is much less of a concern in the regression framework and in the first specification of the structural model. The college in county variable is very strong, so if the interaction with it was important one would expect to have power. Card (1995b) rejected this hypothesis so it seems as if we should have power as well.

In contrast the instrumental variable estimates of the returns to schooling are imprecise. It is much less clear how much power we have in this case. In interpreting our results, there are a number of things to keep in mind. First, the evidence in support of discount rate bias is also weak. There are a number of studies that find larger IV estimates than OLS estimates, but each one is also imprecise. Second, a number of these studies are based on a variable like presence of a college. Our results show the sensitivity of these results to inclusion of the local income variable. If inclusion of this variable had a similar effect on those studies, the pattern of consistently finding IV estimates that are higher than OLS estimates would weaken considerably. Third, our result is not just that we can't reject the null hypothesis. In fact the point estimates goes the wrong direction. Fourth, problems with the instruments seem to bias things in the wrong direction. The regression results in Table 4 suggest that the estimate using presence of a college may be biased upwards, while the argument about family contributions during recessions suggest that the IV estimates here may be biased downwards. While these IV results in themselves may not be convincing that borrowing constraints don't exist, at the very least they cast some doubt on the discount rate bias story for higher IV estimates particularly when combined with the other evidence on borrowing constraints in schooling decision (e.g. Cameron and Heckman (2000), Keane and Wolpin (1990, and Shea (2000).[23]

Given the problem in calculating standard errors in the structural selection model it is even harder to obtain a good sense for the level of power. Once again it is important to keep in mind that it is not just that we fail to reject the null, but the point estimates show no evidence of heterogeneity in access to credit. We believe this basic approach is useful for examining this issue and hope it leads to additional studies on this topic which also test for these types of borrowing constraints. While one study may not be conclusive, a large number with similar results could be.

In general our results are consistent with Cameron and Heckman (2000), Keane and

[23] Ashenfelter and Harmon (1999) have an alternative argument. They argue that the pattern may be due to specification/publication bias. Since the IV standard errors are often high, if researchers prefer specifications in which the t-statistic is greater than two there will be a bias towards higher IV estimates reported in published papers.

Wolpin (1999), and Shea (2000) in that in all four methods we find no evidence of borrowing constraints. Once again it is important to keep in mind that this does not necessarily mean that credit market constraints would not exist in the absence of the programs currently available. It implies instead that given the large range of subsidies to education that currently exist, there is no evidence of large inefficiencies in the schooling market resulting from borrowing constraints.

References

Altonji, J., Elder, T., and Taber, C., "Selection on Observed and Unobserved Variables: Assesing the Effectiveness of Catholic Schools," unpublished manuscript, 2000.

Angrist, J., and Krueger, A., "Empirical Strategies in Labor Economics," in Ashenfelter and Card eds. *Handbook of Labor Economics,* 2000.

Arkes, J., "Using State Unemployment Rates During Teenage Years as an Instrument to Estimate the Returns to Schooling," unpublished manuscript, Center for Naval Analyses, 1998.

Ashenfelter, O., and Harmon, C., "Editors Introduction," *Labour Economics* (special issue on education), forthcoming 1999.

Becker, Gary S., *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education,* (Chicago: University of Chicago Press), 1975.

Blossfeld, H. and Shavit, Y., 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*, Boulder, Co: Westview Press, 1993.

Cameron, S., and Heckman, J., "The Nonequivalence of High School Equivalents," *The Journal of Labor Economics*, 11(1993),1-47.

_____1998a. "Life Cycle Schooling and Educational Selectivity: Models and Choice," , *Journal of Political Economy*, April, 1998.

_____1998b."The Dynamics of Education Attainment for Blacks, Whites, and Hispanics," Columbia University manuscript, 1998. Presented at NBER, April 1997.

Card, David, "Earnings, Schooling, and Ability Revisited," *Research in Labor Economics 14,* 1995a, 23-48.

Card, David, "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," in Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky, eds. *Aspects of Labor Market Behavior: Essays in Honour of John Vanderkamp,* (Toronto:University of Toronto Press), 1995b, 201-222.

Card, David, "Esimating the Return to Schooling: Progress on Some Persistent Econometric Problems," unpublished manuscript, 1999.

Card, David, "The Causal Effect of Education on Earnings," in Ashenfelter and Card eds. *Handbook of Labor Economics,* 2000.

Hauser, R., 1993. "Trends in College Attendance Among Blacks, Whites, and Hispanics,." in *Studies of Supply and Demand in Higher Education.* C. Clotfelter and M. Rothschild, eds. University of Chicago Press.

Heckman, J., and Singer, B., "A Method for Minimizing the Impact of Distributional Assumptions in Economic Models for Duration Data," *Econometrica*, 52(1984), 271-320.

Heckman, J. and Edward Vytlacil, "Instrumental Variables for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling when the Return is Correlated with Schooling," forthcoming, *Journal of Human Resources,* 1998.

Imbens, G., and Angrist, J., "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(1994).

Kane, T. 1994. "College Entry by Blacks since 1970: The Role of College Costs, Family Background and the Returns to Education," *Journal of Political Economy*,102, 878-911.

Keane, M., and Wolpin, K., "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," unpublished manuscript, 1999.

Kling, Jeffrey, "Interpreting Instrumental Variables Estimates of the Returns to Schooling," Princeton Industrial Sections Working paper no. 415, 1999.

Lang, Kevin, "Ability Bias, Discount Rate Bias, and the Return to Education," unpublished manuscript, Boston University, 1994.

Manski, C. and D. Wise, 1983. *College Choice in America*, Harvard University Press.

Manski, C., 1993. "Income and Higher Education," *Focus* (University of Wisconsin-Madison, Institute for Research on Poverty) Vol 14, No. 3.

Mare, R.D. 1980. "Social Background and School Continuation Decisions," *Journal of The American Statistical Association*, 75 (June), 295-305.

Orfield, G, 1992. "Money, Equity, and College Access," *Harvard Educational Review* 72(3), (Fall, 1992l): 337-372.

Rosen, S., 1977, "Human Capital A Survey of Empirical Research," *Research in Labor Economics Vol. 1,* ed. R. Ehrenberg, 3-39.

Shea, J., 2000. "Does Parent's Money Matter?" forthcoming, *Journal of Public Economics.*

Taber, C., , "The Rising College Premium in the Eighties: Return to College or Return to Ability?" unpublished manuscript, Northwestern University, 1999.

Willis, R., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in O. Ashenfelter and F. Layard eds., *Handbook of Labor Economics*, Amsterdam, North Holland, 1986.

Willis, R. and Rosen, S., "Education and Self- Selection," *Journal of Political Economy*, 87, 1979.

Appendix

Downward Bias of Discount Rate with Foregone Earnings

To demonstrate the intuition as to how discount rate bias can actually bias the estimate *downward*, we follow the methodology of Imbens and Angrist (1994). Suppose schooling is binary with S equal to either zero or one (loosely college or high school) and that local labor market conditions are binary with R equal to zero or one. We consider $R = 1$ to represent a period of temporary bad labor market conditions (i.e. a recession). Using notation similar to Imbens and Angrist, we let $d(R)$ be an individual specific dummy variable equal to 1 if this individual would choose schooling level 1 in labor market condition R. That is for each student we define $d(1)$ to indicate whether they would attend college during a recession, and $d(0)$ indicate whether they would attend during a boom.

Let W_S denote the wage the student would receive if he chooses schooling level S, and W denote the unconditional wage. Assume that R is a legitimate instrument so that it is uncorrelated with $(W_1, W_0, d(1), d(0))$. The Wald estimate takes the form,

$$\begin{aligned}
&\frac{E(W \mid R=1) - E(W \mid R=0)}{\Pr(S=1 \mid R=1) - \Pr(S=1 \mid R=0)} \\
=&\frac{E(W_1 \mid R=1, d(1)=1)\Pr(d(1)=1) + E(W_0 \mid R=1, d(1)=0)\Pr(d(1)=0)}{\Pr(d(1)=1) - \Pr(d(0)=1)} \\
&- \frac{E(W_1 \mid R=0, d(0)=1)\Pr(d(0)=1) + E(W_0 \mid R=0, d(0)=0)\Pr(d(0)=0)}{\Pr(d(1)=1) - \Pr(d(0)=1)} \\
=&\frac{E(W_1 \mid d(1)=d(0)=1)\Pr(d(1)=d(0)=1) + E(W_1 \mid d(1)>d(0))\Pr(d(1)>d(0))}{\Pr(d(1)=1) - \Pr(d(0)=1)} \\
&+ \frac{E(W_0 \mid d(1)<d(0))\Pr(d(1)<d(0)) + E(W_0 \mid d(1)=d(0)=0)\Pr(d(1)=d(0)=0)}{\Pr(d(1)=1) - \Pr(d(0)=1)} \\
&- \frac{E(W_1 \mid d(1)=d(0)=1)\Pr(d(1)=d(0)=1) + E(W_1 \mid d(1)<d(0))\Pr(d(1)<d(0))}{\Pr(d(1)=1) - \Pr(d(0)=1)} \\
&- \frac{E(W_0 \mid d(1)>d(0))\Pr(d(1)>d(0)) + E(W_0 \mid d(1)=d(0)=0)\Pr(d(1)=d(0)=0)}{\Pr(d(1)=1) - \Pr(d(0)=1)} \\
=&\frac{E(W_1 - W_0 \mid d(1)>d(0))\Pr(d(1)>d(0)) - E(W_1 - W_0 \mid d(1)<d(0))\Pr(d(1)<d(0))}{\Pr(d(1)>d(0)) - \Pr(d(1)<d(0))}.
\end{aligned}$$

Given that we find that enrollment increases in bad times, Imbens and Angrist's (1994) monotonicity assumption would imply that people are only induced to increase schooling during a recession. That is there are no individuals who would attend school during a boom, but not during a recession ($\Pr(d(1) < d(0)) = 0$). Under these conditions, the Wald estimate would be equal to $E(W_1 - W_0 \mid d(1) = 1, d(0) = 0)$. The argument in the text suggests that this parameter is not likely to be strongly influenced by "discount rate bias."

Now consider relaxing the monotonicity assumption. Suppose that there are some students from borrowing constrained families that cannot afford to send their families to schooling during a recession but would send their children to college during a boom. In this case monotonicity would be violated because for these students $(d(1) = 0, d(0) = 1)$,so $Pr(d(1) < d(0))$ will be positive. To be consistent with the observation in the data that schooling rises during bad times, it must be the case that $\Pr(d(1) = 1, d(0) = 0) > \Pr(d(1) = 0, d(0) = 1)$. Under the discount rate bias hypothesis, students from these borrowing constrained families have high marginal returns to college. This means that $E(W_1 - W_0 \mid d(1) < d(0))$,would be large. Notice however that in this example this term enters the expression for the Wald estimate negatively. Thus in this case discount rate bias will bias this estimate *downward.*

Identification of Structural Model

Nonlinear models that impose linear index assumptions can often avoid the use of exclusion restrictions, so to focus on the variation arising from the exclusion restriction we consider nonparametric identification of the model. Specifically we consider identification of,

$$
\begin{aligned}
S &= 1(\log(e^{g_1(X)+\varepsilon_1} + g_2(Z_1)) + g_3(X) + g_4(Z_2) + \varepsilon_2 > 0) \\
\log(w_{1t}) &= \gamma_S + g_1(X) + g_t(Z_{\ell t}) + E_t\beta_E + E_t^2\beta_{E^2} + u_1 \\
\log(w_{0t}) &= \gamma_S + g_1(X) + g_t(Z_{\ell t}) + E_t\beta_E + E_t^2\beta_{E^2} + u_0,
\end{aligned}
$$

where Z_2 represents the exclusion restriction from local labor market conditions, Z_1 represents the exclusion restrictions that represent direct costs of schooling, $Z_{\ell t}$ represents local labor market variables at time t, and X represents other regressors that influence wages and the tastes for schooling.[24] The econometrician can observe $(X, Z_2, Z_2, Z_{\ell t}, E_t)$ and schooling

[24] The one type of variation we have not included in this specification is the local labor market variables that influence college choice. For this exercise we can fix them to a constant and only consider variation

choice S. When $S=1$, w_{1t} is observed and when $S=0, w_{0t}$ is observed.

Before one can show identification, there are some normalizations that must be imposed. First notice that we can always normalize $g_3(0)=g_4(0)=0$ by modifying the location of the error term ε_2. We need a scale normalization for g_2. As long as $g_2(0)\neq 0$,we can normalize $g_2(0)=1$. To see why we need a scale normalization, notice that for any τ we can multiply g_2 by e^{τ},add τ to ε_{1i},and subtract τ rom ε_{2i} without changing the expression. Under these normalizations and standard assumptions about support conditions, the model is identified.

Theorem 1 *Normalize $g_1(0)=g_3(0)=g_4(0)=0$ and $g_2(0)=1$. Assume that (a) the support of X does not depend on (Z_1,Z_2), (b) $(g_2(Z_1),g_4(Z_2))$has support $\Re^2$, and (c) the error terms $(\varepsilon_1,\varepsilon_2,u_1,u_2)$ are independent of the observables $(X,Z_2,Z_2,Z_{\ell t})$. Then g_1,g_2,g_3,and g_4 and the joint distribution of $(\varepsilon_{1i},\varepsilon_{2i},u_{1i})$ are identified.*

Proof: First notice that we can use a standard identification at infinite argument (see e.g. Heckman, 1990) to identify g_1. That is send $g_4(Z_2)\to\infty$,and we can identify the form of the wage equation.

Now suppose the model is not identified so that there are two distinct models that explain the data, the true model

$$1(\log(e^{g_1(X)+\varepsilon_1}+g_2(Z_1))+g_3(X)+g_4(Z_2)+\varepsilon_2>0)$$

and an alternative model,

$$1(\log(e^{g_1(X)+\varepsilon_1^*}+g_2^*(Z_2))+g_3^*(X)+g_4^*(Z_3)+\varepsilon_2^*>0).$$

We will show that these two models are equivalent.

Taking g_4 to the other side of the inequliy sign and exponentiating both sides one can see that the distribution of,

$$\omega_1=e^{g_1(X)+g_3(X)+\varepsilon_1+\varepsilon_2}+g_2(Z_1)e^{g_3(X)+\varepsilon_2}$$

from the variables that influence local labor markets during college. We are assuming that Z_2 help predict local labor market levels $Z_{\ell t}$ during college.

must be the same as the distribution of,

$$\omega_1^* = f\left(e^{g_1(X)+g_3^*(X)+\varepsilon_1^*+\varepsilon_2^*} + g_2^*(Z_1)e^{g_3^*(X)+\varepsilon_2^*}\right),$$

where f is defined so that $e^{-g_4(X)} = f\left(e^{-g_4^*(X)}\right)$. Notice that

$$\log\left(\frac{\partial E\left(\omega_1 \mid X, Z_1, Z_2\right)}{\partial Z_1}\right) = \log(g_2^{'}(Z_1)) + g_3(X) + \log(E(e^{\varepsilon_2}))$$

which is separable in Z_1 and X. The only way that $\log\left(E\left(\omega_1 \mid X, Z_1, Z_2\right)\right)$ can take this separable form is if f is linear. The fact that $g_4(X)$ has full support and that $g_4(0)$ is normalized to zero implies that $g_4 = g_4^*$.

Since

$$\log(g_2^{'}(Z_2)) + g_3(X) + \log(E(e^{\varepsilon_2})) = \log(g_2^{*'}(Z_1)) + g_3^*(X) + \log(E(e^{\varepsilon_2^*}))$$

g_3 and g_3^* can only differ by a location parameter, so since $g_3(0)$ is normalized to zero, g_3 is identified.

Since $g_1$and g_3 are identified it must be that

$$g_2^*(Z_1) = \frac{E(e^{\varepsilon_2^*})}{E(e^{\varepsilon_1})} g_2\left(Z_\cdot\right)$$

But since $g_2(0) = g_2\left(0\right) = 1$, $g_2 = g_2^*$. Thus g_1, g_2, g_3,and g_4 are all identified.

Now consider identification of the joint distribution of the error terms. Fixing $X = 0$, $Z_1$to a particular value z_1, and varying $g_4(Z_2)$ we can identify the joint distribution of $(-\log(e^{\varepsilon_1} + g_2(z_1)) - \varepsilon_2, u_1)$from

$$\Pr(s_i = 1, u_1 < y) = \Pr\left(-\log(e^{\varepsilon_1} + g_2(z_1)) - \varepsilon_2 < g_4(Z_2), u_1 < y\right).$$

Thus for any $(t_1, t_2) \in \Re^2$ we can identify,

$$\begin{aligned} E\left(\exp\left\{i\left(t_1 e^{\log(e^{\varepsilon_1}+g_2(z_1))+\varepsilon_2} + t_2 u_1\right)\right\}\right) &= E\left(\exp\left\{i\left(t_1 e^{\varepsilon_1+\varepsilon_2} + t_1 g_2(z_1) e^{\varepsilon_2} + t_2 u_1\right)\right\}\right) \\ &= \varphi(t_1, t_1 g_2(z_1), t_2), \end{aligned}$$

where φ characteristic function of $\left(e^{\varepsilon_{1i}+\varepsilon_{2i}}, e^{\varepsilon_{2i}}, u_{1i}\right)$. By varying t_1, t_2,and $g_2(z_1)$we can identify the characteristic function and thus the joint distribution of $\left(e^{\varepsilon_1+\varepsilon_2}, e^{\varepsilon_2}, u_1\right)$. From this we can identify the distribution of $\left(\varepsilon_1, \varepsilon_2, u_1\right)$. ■

7.1 Identification of $\log(R)$ from $E(\theta_1 \mid \theta_1 - \log(R))$

We suspect that this result is not new although we have not found a source that proves it.

Proposition 2 *Suppose that (a) R is independent of θ_1, (b) the distribution of $(\theta_1 - \log(R))$ is identified and (c) $E(\theta_1 \mid \theta_1 - \log(R))$ is identified. Then the distribution of R and θ_1 are identified.*

Proof: Let φ_θ and $\varphi_R$be the characteristic functions of θ_1 and $\log(R)$ respectively. Then we can identify,

$$E\left(e^{it(\theta_1 - \log(R))}\right) = \varphi_\theta(t)\varphi_R\left(-t\right).$$

We can also identify

$$\begin{aligned} E\left(E(\theta_1 \mid \theta_1 - \log(R))e^{it(\theta_1 - \log(R))}\right) &= E\left(\theta_{1i}e^{it(\theta_1 - \log(R))}\right) \\ &= E\left(\theta_{1i}e^{it(\theta_1)}\right)\varphi_{R_i}\left(-t\right). \end{aligned}$$

Thus we can identify the ratio,

$$\begin{aligned} \frac{iE\left(\theta_1 e^{it(\theta_1)}\right)\varphi_R\left(-t\right)}{\varphi_\theta\left(t\right)\varphi_R\left(-t\right)} &= \frac{E\left(i\theta_1 e^{it\theta_1}\right)}{\varphi_\theta\left(t\right)} \\ &= \frac{\partial \log(\varphi_\theta\left(t\right))}{\partial t}. \end{aligned}$$

But this means that φ_θ and thus the distribution of θ_1 is identified. Given this and $\varphi_\theta(t)\varphi_R\left(-t\right)$, then φ_R and thus the distribution of $\log(R)$ is identified.

■

Table 1

Summary Statistics

Primary Variables

Variable	Mean	Standard Deviation	Sample Size
Years of Schooling	12.802	2.487	2404
College in County	0.867	0.340	2404
Local Income at 17	12.845	2.385	2404
Black	0.314	0.464	2404
Hispanic	0.186	0.389	2404
AFQT Score	0.579	22.140	2404
Math Score	0.181	8.439	2404
Word Knowledge	0.172	8.439	2404
General Science	0.098	5.388	2404
Automotive Knowledge	0.039	5.443	2404
Highest Grade Father	10.503	4.103	2404
Highest Grade Mother	10.780	3.211	2404
Number of Siblings	3.734	2.626	2404
Family Income	3.356	2.065	2404
Experience	7.536	3.169	13762
Local Income	13.919	3.479	13762

Table 2
First Stage Regression of
Schooling on Explanatory Variables
(Standard Errors in Parentheses)§

Variable	(1)	(2)	(3)	(4)
College in County	0.417			0.435
	(0.152)			(0.148)
Local Income at 17		-0.023	-0.182	-0.183
		(0.021)	(0.074)	(0.074)
Mean Local Income over			0.130	0.123
Working Life			(0.058)	(0.058)
Black	0.688	0.723	0.691	0.677
	(0.124)	(0.124)	(0.123)	(0.124)
Hispanic	0.334	0.359	0.348	0.345
	(0.138)	(0.143)	(0.141)	(0.141)
Math Score	0.167	0.168	0.167	0.166
	(0.010)	(0.010)	(0.010	(0.010)
Word Knowledge	0.050	0.050	0.050	0.050
	(0.010)	(0.009)	(0.010)	(0.010)
General Science	0.065	0.066	0.066	0.065
	(0.015)	(0.015)	(0.015)	(0.015)
Automotive Knowledge	-0.065	-0.069	-0.065	-0.066
	(0.011)	(0.010)	(0.011)	(0.011)
Highest Grade Father	0.051	0.053	0.052	0.050
	(0.016)	(0.016)	(0.016)	(0.016)
Highest Grade Mother	0.038	0.039	0.039	0.039
	(0.022)	(0.022)	(0.022)	(0.022)
Number of Siblings	-0.044	-0.047	-0.047	-0.046
	(0.016)	(0.016)	(0.016)	(0.016)
Family Income	0.090	0.094	0.095	0.095
	(0.022)	(0.022)	(0.022)	(0.022)
Constant	11.320	11.710	11.993	11.886
	(0.273)	(0.335)	(0.371)	(0.321)
Geographic Controls	Yes	Yes	Yes	Yes
Cohort Controls	Yes	Yes	Yes	Yes
Sample Size	2404	2404	2404	2404

§ The standard errors are constructed to allow for arbitrary correlation between individuals from the same county at age 17

Table 3
Results for Log Wage Regressions
OLS and IV Estimates
Using Foregone Earnings as Instrument
(Standard Errors in Parentheses)[§]

	OLS	IV1	IV2	OLS	IV1	IV2
Schooling	0.058	0.083	0.110	0.074	0.107	0.134
	(0.004)	(0.042)	(0.086)	(0.004)	(0.034)	(0.061)
Local Income	0.027	0.035	0.036	0.025	0.031	0.034
	(0.003)	(0.006)	(0.005)	(0.003)	(0.005)	(0.005)
Experience	0.054	0.065	0.091	0.055	0.074	0.083
	(0.006)	(0.018)	(0.023)	(0.006)	(0.019)	(0.022)
Experience Squared	-0.002	-0.002	-0.004	-0.002	-0.002	-0.004
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
Black	-0.063	-0.085	-0.115	-0.148	-0.162	-0.178
	(0.024)	(0.031)	(0.061)	(0.022)	(0.022)	(0.028)
Hispanic	-0.020	-0.033	-0.041	-0.061	-0.074	-0.085
	(0.030)	(0.029)	(0.034)	(0.030)	(0.029)	(0.032)
Highest Grade Father	-0.003	-0.005	-0.008	-0.001	-0.005	-0.013
	(0.004)	(0.005)	(0.010)	(0.004)	(0.006)	(0.014)
Highest Grade Mother	-0.005	-0.008	-0.012	-0.001	-0.007	-0.015
	(0.005)	(0.006)	(0.011)	(0.005)	(0.007)	(0.014)
Number of Siblings	0.002	0.003	0.005	-0.001	0.001	0.005
	(0.003)	(0.004)	(0.005)	(0.005)	(0.004)	(0.007)
Math Score	0.010	0.007	0.002			
	(0.002)	(0.005)	(0.014)			
Word Knowledge	0.002	0.001	-0.000			
	(0.002)	(0.002)	(0.004)			
General Science	-0.004	-0.005	-0.006			
	(0.003)	(0.003)	(0.006)			
Automotive Knowledge	0.012	0.014	0.016			
	(0.002)	(0.003)	(0.007)			
Family Income	0.036	0.034	0.031	0.042	0.039	0.034
	(0.005)	(0.005)	(0.007)	(0.005)	(0.006)	(0.005)
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort Controls	Yes	Yes	Yes	Yes	Yes	Yes
Instruments:						
Local Income at 17		Yes	Yes		Yes	Yes
Current Local Inc, County 17		Yes	Yes		Yes	Yes
Age and AGE^2		No	Yes		No	Yes
Number of Individuals	2225	2225	2225	2225	2225	2225
Number of Wage Years	13762	13762	13762	13762	13762	13762

[§] The standard errors are constructed to allow for arbitrary correlation between individuals from the same county at age 17

Table 4a

Results for Log Wage Regressions

OLS and IV Estimates

Using Presence of Local College as Instrument

(Standard Errors in Parentheses)§

	OLS	IV1	IV2	OLS	IV1	IV2
Schooling	0.062	0.228	0.193	0.058	0.057	0.061
	(0.004)	(0.109)	(0.084)	(0.004)	(0.115)	(0.076)
Local Income				0.027	0.035	0.035
				(0.002)	(0.006)	(0.005)
Experience	0.054	0.124	0.075	0.054	0.054	0.102
	(0.006)	(0.047)	(0.020)	(0.006)	(0.049)	(0.019)
Experience Squared	-0.002	-0.003	-0.004	-0.002	-0.002	-0.005
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Black	-0.029	-0.111	-0.119	-0.063	-0.073	-0.082
	(0.026)	(0.066)	(0.067)	(0.023)	(0.063)	(0.059)
Hispanic	0.009	-0.023	-0.020	-0.020	-0.028	-0.029
	(0.035)	(0.045)	(0.044)	(0.026)	(0.037)	(0.036)
Highest Grade Father	-0.003	-0.014	-0.015	-0.003	-0.003	-0.003
	(0.004)	(0.009)	(0.010)	(0.004)	(0.008)	(0.008)
Highest Grade Mother	-0.002	-0.016	-0.017	-0.005	-0.006	-0.006
	(0.005)	(0.010)	(0.011)	(0.005)	(0.010)	(0.010)
Number of Siblings	-0.001	0.003	0.005	0.002	0.003	0.003
	(0.004)	(0.005)	(0.005)	(0.003)	(0.005)	(0.005)
Math Score	0.010	-0.008	-0.010	0.010	0.010	0.010
	(0.002)	(0.013)	(0.014)	(0.002)	(0.013)	(0.013)
Word Knowledge	0.003	-0.001	-0.002	0.002	0.002	0.002
	(0.002)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
General Science	-0.005	-0.012	-0.013	-0.004	-0.003	-0.003
	(0.003)	(0.006)	(0.006)	(0.003)	(0.006)	(0.005)
Automotive Knowledge	0.009	0.017	0.018	0.012	0.012	0.012
	(0.002)	(0.005)	(0.006)	(0.002)	(0.005)	(0.005)
Family Income	0.039	0.033	0.032	0.036	0.035	0.034
	(0.005)	(0.007)	(0.007)	(0.005)	(0.006)	(0.007)
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort Controls	Yes	Yes	Yes	Yes	Yes	Yes
Instruments:						
College in County		Yes	Yes		Yes	Yes
Current Local Inc, County 17		No	No		Yes	Yes
Age and AGE^2		No	Yes		No	Yes
Number of Individuals	2225	2225	2225	2225	2225	2225
Number of Wage Years	13762	13762	13762	13762	13762	13762

§ The standard errors are constructed to allow for arbitrary correlation between individuals from the same county at age 17

Table 4b

Results for Log Wage Regressions
OLS and IV Estimates
Using Presence of Local College as Instrument
(Standard Errors in Parentheses)§

	OLS	IV1	IV2	OLS	IV1	IV2
Schooling	0.054	0.027	0.042	0.074	0.052	0.065
	(0.004)	(0.095)	(0.060)	(0.004)	(0.078)	(0.046)
Local Income	0.026	0.035	0.034	0.025	0.031	0.031
	(0.002)	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)
Experience	0.054	0.042	0.106	0.055	0.043	0.104
	(0.006)	(0.046)	(0.017)	(0.006)	(0.044)	(0.016)
Experience Squared	-0.002	-0.002	-0.005	-0.002	-0.002	-0.005
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Black	-0.081	-0.077	-0.087	-0.147	-0.154	-0.160
	(0.021)	(0.066)	(0.062)	(0.021)	(0.026)	(0.026)
Hispanic	-0.040	-0.043	-0.043	-0.061	-0.065	-0.064
	(0.026)	(0.040)	(0.040)	(0.026)	(0.032)	(0.033)
Highest Grade Father	-0.004	-0.002	-0.001	-0.001	0.001	0.002
	(0.004)	(0.009)	(0.009)	(0.004)	(0.010)	(0.010)
Highest Grade Mother	-0.006	-0.005	-0.005	-0.001	0.001	0.000
	(0.005)	(0.008)	(0.008)	(0.005)	(0.011)	(0.010)
Number of Siblings	0.002	0.003	0.003	-0.001	-0.001	-0.001
	(0.003)	(0.004)	(0.004)	(0.003)	(0.006)	(0.006)
AFQT Score	0.005	0.006	0.006			
	(0.000)	(0.004)	(0.003)			
Family Income	0.036	0.036	0.036	0.042	0.044	0.043
	(0.005)	(0.006)	(0.006)	(0.005)	(0.009)	(0.009)
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort Controls	Yes	Yes	Yes	Yes	Yes	Yes
Instruments:						
College In County		Yes	Yes		Yes	Yes
Current Local Inc, County 17		Yes	Yes		Yes	Yes
Age and AGE^2		No	Yes		No	Yes
Number of Individuals	2225	2225	2225	2225	2225	2225
Number of Wage Years	13762	13762	13762	13762	13762	13762

§ The standard errors are constructed to allow for arbitrary correlation between individuals from the same county at age 17

Table 5

Regression of Instrumental Variables
on other Determinants of Schooling and Wages
(Standard Errors in Parentheses)

	Dependent Variable	
Covariate	College in County	Local Income at Seventeen
College in County		0.038
		(0.091)
Local Wage at 17	0.006	
	(0.014)	
Mean Local Income over	0.013	0.748
Working Life	(0.011)	(0.017)
Black	0.032	-0.094
	(0.035)	(0.070)
Hispanic	0.005	-0.007
	(0.037)	(0.078)
Math Score/10	0.030	-0.052
	(0.016)	(0.033)
Word Knowledge/10	0.014	0.022
	(0.014)	(0.054)
General Science/10	0.023	0.022
	(0.024)	(0.054)
Automotive Knowledge	-0.053	-0.006
	(0.017)	(0.043)
Highest Grade Father	0.006	-0.004
	(0.002)	(0.007)
Highest Grade Mother	-0.001	-0.001
	(0.003)	(0.007)
Number of Siblings	-0.002	0.007
	(0.003)	(0.007)
Family Income	-0.001	0.035
	(0.003)	(0.012)
Sample Size	2413	2413

Table 6

Determinants of College Attendance
and High School Dropout
Average Derivatives from Probit Model
(Standard Errors in Parentheses)

	Dependent Variable	
Covariate	College Attendance	High School Dropout
College in County	0.166	-0.035
	(0.036)	(0.030)
Local Income at 17	-0.043	0.028
	(0.022)	(0.016)
Mean Local Income over Working Life	0.035	-0.009
	(0.017)	(0.013)
Black	0.170	-0.109
	(0.037)	(0.023)
Hispanic	0.144	-0.034
	(0.041)	(0.024)
Math Score	0.032	-0.020
	(0.003)	(0.002)
Word Knowledge	0.017	-0.007
	(0.003)	(0.002)
General Science	0.010	-0.006
	(0.004)	(0.003)
Automotive Knowledge	-0.013	0.002
	(0.003)	(0.002)
Highest Grade Father	0.013	-0.003
	(0.004)	(0.003)
Highest Grade Mother	0.010	-0.004
	(0.007)	(0.004)
Number of Siblings	-0.016	0.004
	(0.005)	(0.003)
Family Income	0.011	-0.023
	(0.007)	(0.005)
Geographic Controls	Yes	Yes
Cohort Controls	Yes	Yes
Sample Size	2404	2404

Table 7

Schooling Regressions
on College in County interacted with
Alternative Covariates
(Standard Errors in Parentheses)

Covariate	(1)	(2)	(3)	(4)
College in County	0.596	3.149	0.103	0.592
	(0.166)	(0.072)	(0.236)	(0.229)
Local Income at 17	-0.102	-0.074	-0.107	-0.104
	(0.059)	(0.038)	(0.059)	(0.058)
Mean Local Income over	0.067	0.038	0.068	0.065
Working Life	(0.046)	(0.031)	(0.046)	(0.046)
Black	0.968	0.333	0.690	0.689
	(0.211)	(0.086)	(0.125)	(0.125)
Hispanic	0.914	0.018	0.347	0.344
	(0.379)	(0.105)	(0.142)	(0.142)
Math Score	0.165	0.077	0.166	0.166
	(0.010)	(0.006)	(0.010)	(0.010)
Word Knowledge	0.050	0.013	0.050	0.050
	(0.010)	(0.007)	(0.010)	(0.010)
General Science	0.067	0.042	0.065	0.066
	(0.015)	(0.012)	(0.015)	(0.015)
Automotive Knowledge	-0.067	-0.031	-0.066	-0.066
	(0.011)	(0.007)	(0.011)	(0.011)
Highest Grade Father	0.051	-0.040	0.051	0.051
	(0.016)	(0.028)	(0.016)	(0.016)
Highest Grade Mother	0.038	0.074	0.038	0.038
	(0.022)	(0.053)	(0.022)	(0.022)
Number of Siblings	-0.046	-0.013	-0.046	-0.012
	(0.016)	(0.011)	(0.016)	(0.038)
Family Income	0.092	0.068	-0.008	0.095
	(0.022)	(0.017)	(0.068)	(0.022)
Black×	-0.330			
College in County	(0.220)			
Hispanic×	-0.635			
College in County	(0.390)			
Highest Grade Father×		0.071		
College in County		(0.030)		
Highest Grade Mother×		-0.065		
College in County		(0.055)		
Family Income×			0.111	
College in County			(0.067)	
Number of Siblings×				-0.039
College in County				(0.041)
Sample Size	2404	2404	2404	2404

Table 8

Results From Structural Schooling Model
Estimated Without Borrowing Constraints
(Standard Errors in Parentheses)[§]

$$V_s = \alpha_1 \log\left(\left(\frac{1}{R_i}\right)^S e^{\tilde{\gamma}_S + X'_{iW}\beta_W + X'_{i\ell 0}\tilde{\beta}_{\ell S}} - \sum_{t=0}^{S}\left(\frac{1}{R_i}\right)^t X'_{ic}\beta_c\right)$$
$$+\alpha_2(S) + \alpha_3(S)\log(R_i) + X'_{iT}\beta_{TS} + u_{iTS}.$$
$$\log(w_{ist}) = \gamma_S + X'_{iW}\beta_W + X'_{i\ell t}\beta_\ell + E_{it}\beta_E + E_{it}^2\beta_{E^2} + \alpha_S\theta + \nu_{it},$$

	Graduate High School	Attend College	Graduate College
Black	0.855	1.215	1.570
	(0.169)	(0.202)	(0.252)
Hispanic	0.136	0.488	0.539
	(0.187)	(0.212)	(0.279)
AFQT Score	0.045	0.054	0.101
	(0.005)	(0.008)	(0.019)
Father Dropout	-0.055	-0.123	-0.028
	(0.144)	(0.158)	(0.198)
Father Some Coll	0.076	0.233	0.603
	(0.258)	(0.273)	(0.293)
Father Coll Grad	0.445	0.618	1.322
	(0.349)	(0.359)	(0.357)
Mother Dropout	-0.142	-0.227	-0.392
	(0.140)	(0.157)	(0.209)
Mother Some Coll	0.077	-0.162	0.376
	(0.295)	(0.318)	(0.319)
Mother Coll Grad	0.200	0.332	0.907
	(0.396)	(0.406)	(0.417)
Number of Siblings	-0.029	-0.092	-0.124
	(0.042)	(0.047)	(0.055)
Family Income	0.162	0.121	0.136
	(0.043)	(0.045)	(0.054)
Region	Yes	Yes	Yes

(Continued on Following Page)

Table 8 (Continued)

	Attend High School	Graduate High School	Attend College	Graduate College
Local Labor Market Parameters[‡] (β_l)				
Intercept			5.689	5.236
			(0.417)	(0.338)
Local Income 17-18	0.044	0.009		
	(0.047)	(0.022)		
Mean Loc Inc over	0.302	0.293	0.300	0.300
Working Life	(0.017)	(0.008)	(-)	(-)
Additional Parameters				
Scale		1.281		
		(0.140)		
High School Nesting (ρ_H)		0.452		
		(0.146)		
College Nesting (ρ_C)		1.000		
		(-)		

[†] In this set of estimates we fixed the interest rate so so that $1/(1+r) = \delta = 0.97$.

[§] Since this model overparamaterized the standard errors are likely to be misleading

[‡] The high school wage is the local Income at ages 17 and 18, the Long Term Income is the average income during working age in age 17 county

Table 9

Results for Tests of Borrowing Constraints
From Structural Model
(P-value of Psuedo Likelihood Ratio Test in Parentheses)[†]

Specification	Rate of Return[‡]	Negative Log-Likelihood
Unrestricted:		2547.69
Everyone	1.031	
Racial Groups:		2547.35
Whites	1.031	(0.712)
Blacks	1.000	
Hispanics	1.000	
Parents Education:		2546.41
Both College Educated	1.031	(0.278)
Father 12, Mother 12	1.000	
Father 12, Mother Coll	1.000	
Father Coll, Mother 12	1.034	
Family Income:		2546.90
Top Third	1.031	(0.454)
Middle Third	1.032	
Bottom Third	1.037	
Number of Siblings*:		2547.68
Zeros	1.031	(0.990)
Two	1.031	
Four	1.031	

[†] These p-values are not strictly correct since we have not corrected for the fact that the coefficients of the wage regression were estimated

[‡] The "market rate of return" is fixed to be 1.03

* In the other cases we interacted the rate with dummy variables. In this one we just interact it with the number of siblings.

Table 10

Results from Structural Schooling Model
Allowing For Unobserved Heterogeneity in Borrowing Rates

$$V_s = \alpha_1 \log\left(\left(\frac{1}{R_i}\right)^S e^{\widetilde{\gamma}_S + X'_{iW}\beta_W + X'_{i\ell 0}\widetilde{\beta}_{\ell S}} - \sum_{t=0}^{S}\left(\frac{1}{R_i}\right)^t X'_{ic}\beta_c\right)$$
$$+\alpha_2(S) + \alpha_3(S)\log(R_i) + X'_{iT}\beta_{TS} + u_{iTS}.$$
$$\log(w_{ist}) = \gamma_S + X'_{iW}\beta_W + X'_{i\ell t}\beta_\ell + E_{it}\beta_E + E^2_{it}\beta_{E^2} + \alpha_S\theta + \nu_{it},$$

Distribution of Borrowing Rates (R_i)

Rates	1.030	1.071
Probability	1.00	0.00

Wage Equation

AFQT Score	0.007
Local Income	0.027
Black	-0.095
Hispanic	-0.047
Experience	0.064
Experience^2	-0.002
Intercept-Dropout	1.046
Ingercept-HS Grad	1.139
Intercept-Some Coll	1.031
Intercept-Coll Grad	1.324
Factor Loading-Dropout(α_0)	0.176
Factor Loading-Dropout(α_1)	0.155
Factor Loading-Dropout(α_2)	0.153
Factor Loading-Dropout(α_3)	0.069
Family Background	Yes
Geographic Controls	Yes

Taste Parameters (β_T)

	Graduate High School	Attend College	Graduate College
Black	0.632	1.725	1.873
Hispanic	0.057	0.838	0.793
AFQT Score	0.043	0.088	0.128
Father Dropout	-0.144	-0.235	-0.108
Father Some Coll	0.051	0.678	0.900
Father Coll Grad	0.317	1.243	1.683
Mother Dropout	-0.166	-0.482	-0.662
Mother Some Coll	0.144	-0.271	0.336
Mother Coll Grad	0.198	0.650	1.244
Number of Siblings	-0.265	-0.113	-0.099
Factor Loading (δ_T)	0.000	2.130	2.130
Region	Yes	Yes	Yes

(Continued on Following Page)

Table 10 (Continued)

	Attend High School	Graduate High School	Attend College	Graduate College
Local Labor Market Parameters (β_l)				
Intercept			2.746	3.225
Local Income 17-18	0.408	0.258		
Mean Loc Inc Work Life	0.245	0.172	0.300	0.300
Additional Parameters				
Scale		0.231		
High School Nesting (ρ_H)		0.452		
College Nesting (ρ_C)		1.000		

† In this set of estimates we fixed the interest rate so so that $1/(1+r) = \delta = 0.97$.