Education and Default Incentives with Government Student Loan Programs*

Lance Lochner
University of Rochester
Hoover Institution
NBER

Alexander Monge-Naranjo
Northwestern University

June, 2002
Preliminary and Incomplete

Abstract

This paper examines data on student loan default from the *Baccalaureate and Beyond Survey*. The main findings include: (1) conditional on debt, the probability of default is declining in both predicted and actual post-school earnings; (2) conditional on earnings, the probability of default is increasing in debt; (3) default rates vary across undergraduate majors, but those differences disappear when controlling for debt and earnings; and most interestingly, (4) there is a U-shaped relationship between ability and the probability of default even after controlling for debt and earnings.

We next develop a model of endogenous human capital investment and default that attempts to replicate these facts. The model incorporates a lending scheme that ties borrowing to investment in schooling but does not make repayments explicitly contingent on the subsequent earnings of borrowers. Punishments on borrowers who default follow those in practice under the current government student loan program, and limits on those punishments cause some borrowers to choose default over repayment. Within the context of the model, we ask the following questions: (1) what types of heterogeneity and market shocks explain our empirical findings? (2) Given the answer to the first question, how different are consumption and investment under the current program with respect to the optimal (uncontingent) lending program? More generally, the model is useful for studying the interaction between borrowing constraints, default, and investment in human capital. In contrast to conventional wisdom, the model suggests that credit constraints do not necessarily imply under-investment in human capital given the current lending system.

1 Introduction

In the 2000-01 academic year, over 9 million American college students borrowed a total of more than $38 billion from federal student loan programs to help finance their education (The College Board [6]). Given recent trends, one in twenty of those borrowers will default on their federal student loans within the first two years of re-payment. Overall, the total amount outstanding from defaulted student loans

*We thank participants at the Institute for Poverty Research Summer Workshop for their comments. Monge gratefully acknowledges the financial support of Northwestern University and the National Science Foundation, grant 0112943. Send comments to a-monge1@northwestern.edu and lochner@hoover.stanford.edu.

stands at \$25 billion (US Dept. of Education [29]). While default rates have declined considerably since reaching their peak of 22% in 1990, they are still quite high and entail substantial sums of money.

In this paper, we empirically examine patterns in student loan default among college graduates and develop a model consistent with those patterns. Using panel data from the Baccalaureate and Beyond Surveys (BB), which follows a random sample of 1992-93 college graduates through 1997, we analyze the relationship between individual background and ability, choice of college major, student borrowing, post-graduate employment and earnings, and loan default. Interesting empirical patterns include:

1. Conditional on education debt, default rates are generally declining in both predicted and actual post-school earnings.

2. Conditional on actual post-school earnings, default rates are generally increasing in education debt.

3. Default rates are U-shaped in SAT/ACT scores, even after controlling for actual post-school earnings and education debt.

4. Default rates vary across students with different undergraduate majors, but those differences largely disappear after controlling for actual post-school earnings and education debt.

5. Blacks and hispanics default at significantly higher rates than whites and Asians, even after controlling for actual post-school earnings and education debt.

To understand these relationships, we model the accumulation of human capital in an environment with incomplete credit markets and limited enforcement of loan contracts. Instead of studying *constrained efficient allocations* as in our previous work (Lochner and Monge [26]), we consider an economy in which young agents face a lending environment similar to the federal student loan system (e.g. the Perkins, William D. Ford Federal Direct Student Loan (FDSL), and Federal Family Education Loan (FFEL) programs). Specifically, loan amounts are limited to the cost of a student's schooling and cannot exceed an upper limit set by the government. Repayment of loans is not explicitly contingent on post-school earnings, though it effectively is through the decision to default or re-pay.

In the model, individuals not only choose their level of human capital investment, but they also choose their type of study (or college major). This choice is determined by both the expected earnings associated with each choice and individual tastes or interests. Post-graduation earnings are determined by human capital investments, choice of major, and unpredictable shocks.

Once any shocks to earnings have been realized, individuals decide whether or not to re-pay their loans. This decision will depend on the amount owed (including the interest rates on any loans) as well as the costs of default. These costs are primarily determined by the amount of earnings the government can seize from borrowers who default.[1]

Because creditors (both the government and private institutions) are restricted in the punishments they can impose on individuals that default, some borrowers may find it privately optimal to do so. When creditors can seize a fraction of earnings above some minimum amount, default should be a function of both post-school earnings and the amount of education debt. When conditioning on both, the model predicts that default should be declining in earnings and increasing in debt, consistent with empirical findings 1 and 2 above.

The option to default will affect schooling decisions. Individuals with a high probability of future default are likely to over-invest in their human capital to the extent that the punishment of default increases less with investment than the benefit does. On the other hand, limits on borrowing may restrict some students from borrowing as much as they would like causing them to under-invest in their human capital. It is also likely that many borrowers invest optimally in their human capital even when they are constrained from borrowing as much as they want. Because federal student lending programs offer loans that cover investment but not additional consumption, these individuals are able to borrow enough to finance efficient levels of investment while they are forced to consume less than they would otherwise choose if they were completely unconstrained. Taken as a whole, knowing an individual is borrowing constrained says nothing about the efficiency of his schooling choice under the current lending system.

The relationship between default and characteristics such as race and ability can be complicated, since these factors not only affect default through earnings conditional on schooling attainment, but they also affect optimal schooling choices. Furthermore, the future option to default impacts schooling and borrowing decisions differently across ability levels. For reasonable parameterizations of the model, default rates are monotonically declining in ability. This makes the empirical finding that default rates are U-shaped in SAT/ACT scores puzzling and has induced us to begin exploring extensions of the basic model.

The paper proceeds as follows. Section 2 briefly reviews the literature related to borrowing constraints and education, default, and bankruptcy. Section 3 summarizes major features of the federal student lending system. Data from the Baccalaureate and Beyond Surveys are used to analyze student borrowing and default in Section 4. We then develop a model of human capital investment, student lending, and default in Section 5 to better understand the empirical patterns we uncover. Section 6 discusses how a more efficient lending environment might look, and Section 7 concludes.

[1] The ability of creditors to seize assets can also be incorporated but it is largely unimportant for student borrowers. Reporting to credit bureaus may have some additional effects that are not directly studied here (see Lochner and Monge [26]). To the extent that the effects of such a punishment are a function of post-graduation earnings, they can be roughly approximated within our framework.

2 Related Literature

This paper builds on and synthesizes three distinct literatures, which we briefly discuss in this section. Since we analyze human capital investment decisions when credit markets are imperfect, we first review the empirical literature on education and borrowing constraints. Our model below is based on earlier models of limited enforcement, which we also review. Finally, our study of student loan default is preceded by a few empirical studies and a more general literature on bankruptcy.

2.1 Education and Borrowing Constraints

It has long been argued that the presence of borrowing constraints induces individuals with low wealth to under-invest in human capital when human capital cannot serve as collateral for financial liabilities (see Becker [5]). In the standard economic framework, credit limits are fixed and independent of the observable characteristics and decisions of individuals. Alternatively, credit 'constraints' are sometimes represented by interest rates that increase with the amount borrowed or that exogenously vary in the population (Becker [5], Cameron and Taber [9] and Card [10]). Based on these ideas, an empirical literature has developed that focuses on two tests of credit constraints in the market for human capital. One branch of the literature tests whether individuals from different family income levels have different college enrollment rates conditional on ability and other variables that may influence tastes for schooling or the ability to attend. The second branch compares the returns to schooling for individuals who are expected to face different interest rates or constraints on their borrowing. See Carniero and Heckman [11] for a recent synthesis of this research.

Disagreement about the importance of credit constraints in determining schooling levels in the U.S. abounds. Kane [20] and Elwood and Kane [14] argue that differences in family income are responsible for a sizeable difference in college enrollment rates. However, Cameron and Heckman [7, 8] find that after controlling for cognitive ability and dynamic unobserved self-selection, family income has little effect on enrollment rates. This leads them to conclude that short-term borrowing constraints at the college age are not responsible for much of the difference in college enrollment rates by family income. However, they cannot rule out the possibility that borrowing constraints impede the development of younger children from low income families, and indeed, Carniero and Heckman [11] argue that this is likely to be true. Estimating a dynamic model of education choice, Keane and Wolpin [21] find evidence of stringent credit constraints during the college years, but they estimate that relaxing those constraints would have little effect on enrollment decisions. While Card [10] argues that individuals most likely to face constraints receive higher returns to schooling (suggesting constraints do exist and prevent constrained youth from pursuing highly productive opportunities), Cameron and Taber [9] find little evidence of differential returns consistent with borrowing constraints. In summarizing this literature, Carneiro and Heckman [11] conclude that short-term credit constraints at the college age do not substantially curtail enrollment in college. However, they find stronger evidence that credit constraints may affect whether or not students delay entry into college.

In the only study that attempts to empirically estimate actual borrowing limits, Keane and Wolpin [21] find that individuals with more human capital can borrow more than those with fewer skills. While the empirical literature has paid little attention to the forces underlying constraints on borrowing, Lochner and Monge [26] develop a model of efficient lending and default incentives that generates outcomes consistent with this finding. The model embodies the idea that, on average, individuals possess few physical assets at early ages when it is most efficient to invest in human capital. They must, therefore, borrow against future earnings if they want to invest in their skills. Even after schooling, in the early stages of labor market participation, they may want to borrow additional funds to smooth consumption in anticipation of higher future earnings. The degree to which incentives to re-pay early loans line up with the needs of would-be students determines the extent of borrowing constraints at young ages. Because more able and skilled individuals can credibly commit to re-pay higher amounts of debt, the model produces the type of heterogeneity in constraints estimated by Keane and Wolpin [21].

2.2 Endogenous Borrowing Constraints

While most human capital theories that explicitly consider borrowing constraints assume that those constraints are exogenously determined (e.g. Aiyagari, Greenwood, and Seshadri [1], Becker [5], Caucutt and Kumar [12], Hanushek, Yilmaz, and Leung, [19] and Loury [27]), theoretical studies of endogenous credit constraints have mainly ignored issues related to human capital accumulation. Early studies were primarily limited to implications for risk sharing and asset prices, taking household earnings to be an exogenous, and often stationary, process (e.g. Alvarez and Jermann [2], Kehoe and Levine [22], and Kocherlakota, [23]). However, a number of more recent papers have used endogenous constraint models to study the importance of durable goods (Krueger and Fernandez [16]) and pensions (Andolfatto and Gervais [3] and Lambertini [25]) in determining life cycle consumption decisions. Also within endogenous constraint frameworks, Krueger and Perri [24] study the effect of progressive taxation on insurance markets, and Attanassio and Ríos-Rull [4] explore the impact of outside transfers on insurance for villages. The only study that analyzes the role of endogenous credit constraints in determining human capital investment decisions is Lochner and Monge [26], which we have discussed above.

2.3 Default and Bankruptcy

A few studies have examined the characteristics of students who default on their loan obligations (Dynarski [13], Greene [17], and Wilms, et. al [31]), concluding that student background characteristics, like race, family income, and high school graduation, are correlated with default. They also find a negative relationship between completion of a college program and subsequent default. Dynarski [13] further examines the relationship between earnings two years after completing college and default, finding that default declines with earnings.

Some recent studies have looked carefully at the effects of government bankruptcy laws on borrowing and lending decisions as well as the decision to file for bankruptcy. Fay, Hurst, and White [15] empirically show that households are more likely to declare bankruptcy when the financial benefits of doing so are greater. Gropp, Scholz, and White [18] estimate that households living in states with larger bankruptcy asset exemptions are more likely to have their loan applications rejected and to be discouraged from borrowing. Monge, et al. [28] and the other studies in Pagano [30] examine how incentives to default translate into borrowing constraints in different countries.

3 Federal Student Lending Programs

From the 1990-91 academic year to that of 2000-01, the percent of total federal student aid provided through government loans rose from 48 to 58 percent. By any measure, student loans are an important source of financing for higher education. In this section, we briefly discuss important features of the federal student loan system.

Historically, private lenders have provided the capital to student borrowers and the government has guaranteed those loans with a promise to cover any unpaid amounts (and, in some cases, interest payments while students are in school), now referred to as the Federal Family Education Loan Program (FFEL). Since the 1994-95 academic year, the federal government has begun to directly provide loans to students through the William D. Ford Federal Direct Student Loan (FDSL) program.[2] Regardless of the source of funds, the rules governing these two programs are the same. Both offer Subsidized and Unsubsidized Stafford Loans as well as Parent Loans to Undergraduate Students (PLUS). Prior to the introduction of Unsubsidized Stafford Loans in the early 1990s, Supplemental Loans to Students (SLS) were an alternative source of unsubsidized federal loans for independent students.

The distinction between subsidized and unsubsidized loans hinges on the treatment of interest on loans while students are enrolled. Students are not charged interest on subsidized loans as long as they are enrolled in school (the government pays the interest), while interest accrues on unsubsidized loans. In order to qualify for subsidized loans, students must demonstrate financial need, which depends on family income, dependency status, and the cost of the institution attended. Unsubsidized loans may be obtained without any show of need. In general, students under age 24 are assumed to be dependent, in which case their parents' income is an important determinant of their financial need. See Table 1

[2]Perkins Loans, which charge low interest rates and are only available for students most in need, are quite limited in scope.

for a basic description of the main federal student loan programs, since the 1987-88 academic year.

Table 1: Major Federal Student Loan Programs from 1987-88 to the Present

Academic Years	Loan Type	Eligible Students	Federal Program	Source of Funds
1987-88 to 1992-93	Stafford Loans (subsidized)	Dependent and Independent Students	Guaranteed Student Loan Program (GSL)	Private Lenders
	SLS Loans (unsubsidized)	Primarily Independent Students	Supplemental Loans to Students	Private Lenders
1993-94 to present	Subsidized Stafford Loans	Dependent and Independent Students Showing Need	Federal Family Education Loan Program (FFEL)	Private Lenders
			William D. Ford Direct Loan Program (FDSL)	U.S. Department of Education
	Unsubsidized Stafford Loans	Dependent and Independent Students	Federal Family Education Loan Program (FFEL)	Private Lenders
			William D. Ford Direct Loan Program (FDSL)	U.S. Department of Education

Dependency status and class level determine the total amount of loans a student is eligible for as seen in Table 2. Through 1992-93, dependent undergraduate students could borrow up to a total of \$17,250 over a five year period, while that limit was raised to \$23,000 in subsequent years. Interestingly, dependent students finishing college within four years are eligible for \$4,000 or \$5,500 (depending on the time period) less than those enrolled for five years. Independent students can borrow roughly twice those amounts, though most traditional undergraduates do not fall into this category.

Finally, we discuss issues related to re-payment and default. Re-payment of Stafford Loans begins six months after finishing school. Borrowers that are having difficulties making payments and that can establish financial hardship may qualify for either a forbearance or deferment, which will temporarily delay payments.[3] Loans covered by the federal system cannot generally be expunged through bankruptcy except in very special circumstances. Thus, the only way a borrower can 'avoid' re-payment when he does not qualify for a deferment or forbearance is to simply stop making payments, or default. A borrower is considered to be in default once he becomes 270 days late in making

[3]Interest continues to accrue during a forbearance, but it does not during a deferment.

Table 2: Loan Limits for Undergraduate Borrowers

Class Level	1987-88 to 1992-93 Dependent Students	1987-88 to 1992-93 Independent Students	1992-93 to present Dependent Students	1992-93 to present Independent Students
First Year	$2,625	$6,625	$2,625	$6,625
Second Year	2,625	6,625	3,500	7,500
Third Year	4,000	8,000	5,500	10,500
Fourth Year	4,000	8,000	5,500	10,500
Fifth Year	4,000	8,000	5,500	10,500
Cumulative Total	17,250	37,250	23,000	46,000

Notes:
Amounts include Subsidized and Unsubsidized Stafford Loans and SLS Loans.
Yearly limits sum to less than cumulative totals in the later period.

a payment. If the loan is not fully re-paid immediately, or if a suitable re-payment plan is not agreed upon with the lender, the default status will be reported to credit bureaus, and collection costs (up to 25% of the balance due) may be added to the amount outstanding. Up to 10% of the borrower's wages can be garnished (15% can be garnished by the Department of Education if it becomes involved in the collection process), and federal tax refunds can be seized and applied toward the balance.[4] In practice, these sanctions are sometimes limited by the inability of institutions to locate those who have defaulted. Wage garnishments are ineffective against defaulting borrowers that are self-employed. Furthermore, individuals can object to the wage garnishment if it would leave them with a weekly-take home pay of less than 30 times the federal minimum wage, or if the garnishment would otherwise result in an extreme financial hardship.

4 Empirical Patterns in Default on Student Loans

We use the Baccalaureate and Beyond Surveys (BB) to analyze patterns in default on student loans by college graduates a few years out of college. The survey has followed a random sample of about 11,000 individuals who received their baccalaureate degree during the 1992-93 academic year through 1997 (with surveys in 1993, 1994, and 1997).[5] We also disregard individuals receiving their BA at

[4]Other sanctions against borrowers who default include a possible hold on college transcripts, ineligibility for further federal student loans, and ineligibility for a deferment or forbearance. Since the early 1990s, the government has also begun to punish educational institutions with high student default rates by making their students ineligible to borrow from federal lending programs.

[5]The BB sample is a subsample consisting of all graduating respondents from the 1993 National Postsecondary Student Aid Study (NPSAS), a nationally representative sample of all postsecondary students in the U.S. All averages

age 30 or later to focus on the traditional college student (less than 15% received their BA at such late ages). Because those continuing on to graduate school are eligible for deferments in their loan re-payment, we do not observe the default choices of individuals enrolling in graduate school for more than a short time. We, therefore, focus on U.S. citizens receiving no more than 9 months of graduate education and who are no longer enrolled in any form of school as of the 1997 survey (maintaining about two-thirds of the sample).

Approximately 50% of these graduates report having borrowed money for their schooling, and we focus on them. Our main sample, therefore, consists of 2,796 undergraduate borrowers who graduated from college in the 1992-93 academic year and did not pursue more than 9 months of post-graduate education.

To measure default, we determine whether or not an individual defaulted on (and did not subsequently re-pay) or expunged through bankruptcy any federal student loan through early 1998.[6] Loan amounts are based on survey responses in 1997 and include any borrowing from federal or private sources (including the family). The 1994 and 1997 surveys ask respondents about their earnings at their current job. We calculate their annual earnings and wage rates for the job at which they were employed at the time of these surveys. While the earnings measure may not reflect their actual earnings for those years, it is the only available measure for personal earnings. Respondents also reported household income for the year prior to the 1994 and 1997 surveys.

Table A-1 of the Appendix reports background characteristics for our sample of borrowers. It is nearly equally split along gender lines, with more than 80% white. About one-fourth of the sample majored in business and management. Table 3 reports the extent of borrowing and post-school earnings/employment. Among the borrowers in this sample, average undergraduate loans totaled about $10,500, while graduate loans were negligible (reflecting our sample requirement that they have no more than 9 months of graduate study). Approximately 27% had taken out loans of less than $5,000, and 24% had loans of $15,000 or more. When spousal loans are added on for those who were married, average family educational borrowing was nearly $13,000. Four years after receiving their BA, these borrowers still owed more than $4,500 on their own student loans and their families owed a total of almost $6,000. Nearly 6% of these borrowers had defaulted on (and not subsequently re-paid) at least one of their college loans.

Average earnings for the sample of borrowers one year out of college was $20,790 (1994), increasing to $31,599 four years after graduation (1997).[7] Wage rates in the 1997 job averaged $14.77, and most had experienced little if any time unemployed. Household income averaged about $45,000 in 1996.

in the following tables are weighted to reflect the stratified sampling scheme of the original NPSAS survey as well as any attrition in later surveys.

[6]These data, collected for the BB, are based on individual loan records from the National Student Loan Data System (NSLDS) as of 1998. They report the status of each federal student loan taken out by all borrowers in the sample.

[7]These measures include those with zero earnings–only a very small fraction of the sample.

Table 3: Debt and Earnings for Undergraduate Borrowers

Variable	Mean	Standard Deviation
(a) Debt/Loans (as of 1997):		
Any graduate loans	0.012	0.108
Any outstanding loan balance	0.649	0.477
Any family outstanding loan balance	0.682	0.466
Any other debt	0.970	0.171
Total undergraduate loan amount	10,527	9,238
Total loan amount	10,687	9,493
Educational debt $1-4,999	0.265	0.441
Educational debt $5,000-9,999	0.275	0.446
Educational debt $10,000-14,999	0.216	0.411
Educational debt $15,000-19,999	0.125	0.331
Educational debt $20,000 +	0.119	0.324
Total family loan amount	12,598	12,121
Amount still owed on loans	4,562	5,904
Amount family still owes on loans	5,910	8,572
Amount of other debt	956	578
Tuition paid for 1992-93 year	4,484	4,721
Default (without re-payment)	0.056	0.231
(b) Earnings and Employment:		
1994 earnings	20,790	39,131
longest unemployment spell from BA to 1994 survey (months)	1.502	3.048
1997 earnings	31,599	21,514
longest unemployment spell from BA to 1997 survey (months)	2.651	5.211
1997 wage rate	14.77	13.79
1996 family income	44,984	27,157
1997 earnings less than $25,000	0.360	0.480
1996 family income less than $25,000	0.196	0.397
Family income less than $25,000 in 1993 or 1996	0.680	0.466
Family income higher than $25,000 in 1993 or 1996	0.819	0.385
Ratio of total education loans to 1997 earnings	3.405	196.800
Ratio of family education loans to 1996 income	1.771	97.236

Notes:
Sample includes all U.S. citizens who did not report a disability, had no more than 9 months of graduate education, were not currently enrolled in school, received their BA prior to age 30, and borrowed money for their education.

Since low incomes may indicate an inability to make loan payments, it is instructive to note that 36% of the college graduates earned less than $25,000 in 1997 while slightly less than 20% had total family income below that level during the previous year.

Table 4 reports information about student borrowing and default by gender, ethnicity, and SAT/ACT quartile for all borrowers in the sample. The final column of the table reports whether individuals had a family income above $25,000 in either 1993 or 1996, since this is likely to reflect an ability (if not desire) to make timely loan payments. Default rates for men and women are nearly identical, as are the amounts borrowed for school and the amount still owed. Among the Asian borrowers, only 2.2% had defaulted on their loans within four years of graduating, despite the fact that their average loan size was nearly $14,000. About 81% had a family income above $25,000 in at least one of the post-graduation years. In stark contrast, black borrowers defaulted at an astonishing rate of 21%, despite taking out loans that averaged more than $3,000 less. These high rates may, at least partially, reflect an inability to maintain payments, since 29% of all black borrowers in our sample had incomes below $25,000 in both 1993 and 1996. Earnings were noticeably lower for blacks than for all other ethnic groups. Hispanics borrowed the least, approximately $8,000 on average, and defaulted about 12% of the time. White borrowers defaulted 3.7% of the time, having taken out loans averaging $10,800. They owed the least ($4,406) on their student loans four years after graduating. Across all ethnic groups, borrowers owed about one-half of their original loan amounts four years after graduating.

Turning to comparisons across SAT/ACT test score quartiles, we find an interesting pattern.[8] Default rates are highest for the *most* able (quartile 4) at 8.6%, followed closely by the lowest ability quartile, which had a default rate of 7.5%. Default among those in the third ability quartile was by far the lowest with only 2.4% choosing not to re-pay their loans. Total loan amounts were quite similar across ability quartiles, with the the least able borrowing the most at $11,425 (about $1,000 more than all other ability categories). Despite the non-monotonic relationship between ability and default, family income is positively correlated with ability as shown in the final column. Fewer than 14% of all high ability borrowers had family income levels below $25,000 in both 1993 and 1996, while more than 19% of the lowest ability borrowers did.

Table 5 reports borrowing and default rates by undergraduate major. While small sample sizes within each category make it difficult to statistically identify differences across majors, a few patterns are worth mentioning. First, average total loan amounts differ by less than $2,000 across most major categories (health profession majors are the clear exception). Second, the likelihood of default varies considerably, from a low of 2.1% for those in health profession majors to a high of almost 12% for those in humanities majors. Default rates are below 4% in majors such as education, engineering, and health professions, while they are above 9.0% in public affairs and social service majors, math

[8]SAT/ACT quartile represents the individual's quartile in the test distribution of all sample SAT scores if available. If the individual did not take the SAT, then the quartile was similarly determined from his ACT score.

Table 4: Borrowing and Default by Gender, Ethnicity, and SAT/ACT Scores

Characteristic	Default	Total Education Loan Amount	Amount Owed	1993 or 1996 family income > $25,000
(a) Full Sample (N=1,252)	0.056	10,687	4,562	0.819
	(0.005)	(180)	(115)	(0.008)
(b) Gender:				
Male (N=1,252)	0.057	10,890	4,334	0.844
	(0.007)	(284)	(164)	(0.011)
Female (N=1,544)	0.056	10,504	4,769	0.796
	(0.006)	(228)	(162)	(0.011)
(c) Ethnicity:				
Asian (N=80)	0.022	13,848	5,864	0.813
	(0.019)	(1,566)	(1,225)	(0.049)
Black (N=204)	0.210	10,517	5,979	0.712
	(0.030)	(578)	(456)	(0.037)
Hispanic (N=172)	0.118	8,184	3,953	0.812
	(0.027)	(511)	(386)	(0.033)
White (N=2,308)	0.037	10,771	4,406	0.829
	(0.004)	(199)	(122)	(0.009)
(d) SAT/ACT Quartile:				
Quartile 1 (N=687)	0.075	11,444	4,716	0.806
	(0.011)	(420)	(220)	(0.017)
Quartile 2 (N=687)	0.050	10,384	4,093	0.804
	(0.009)	(359)	(204)	(0.017)
Quartile 3 (N=595)	0.024	10,190	4,532	0.812
	(0.007)	(312)	(234)	(0.017)
Quartile 4 (N=495)	0.086	10,816	4,727	0.862
	(0.014)	(468)	(340)	(0.017)

Notes:
Sample includes all U.S. citizens who did not report a disability, had no more than 9 months of graduate education, were not currently enrolled in school, received their BA prior to age 30, and borrowed money for their education.

and sciences, and humanities. Third, there is no obvious pattern between post-graduate earnings and default rates when looking across majors. While borrowers majoring in math and sciences are very unlikely to earn less than $25,000 after graduating, they have a very high default rate. On the other hand, those majoring in engineering and health professions are also likely to earn more than $25,000, but they have very low default rates.

Table 5: Borrowing and Default by Undergraduate Major

Undergraduate Major	Default	Total Education Loan Amount	Amount Owed	1993 or 1996 family income > $25,000
Business/management (N=379)	0.048	10,756	4,433	0.818
	(0.012)	(539)	(319)	(0.022)
Education (N=425)	0.038	10,658	4,663	0.768
	(0.010)	(429)	(280)	(0.023)
Engineering (N=222)	0.031	10,961	4,380	0.977
	(0.013)	(571)	(354)	(0.011)
Health professions (N=248)	0.021	13,109	5,183	0.946
	(0.010)	(685)	(430)	(0.016)
Public affairs/social serv. (N=120)	0.090	9,416	4,323	0.825
	(0.028)	(673)	(424)	(0.038)
Biological sciences (N=117)	0.043	9,937	4,218	0.734
	(0.020)	(695)	(553)	(0.043)
Math & science (N=186)	0.095	9,830	4,091	0.917
	(0.024)	(631)	(506)	(0.022)
Social science (N=261)	0.051	9,564	4,556	0.823
	(0.015)	(450)	(321)	(0.026)
History (N=50)	0.061	9,664	3,844	0.759
	(0.037)	(1,385)	(899)	(0.066)
Humanities (N=274)	0.117	10,558	4,775	0.655
	(0.021)	(475)	(313)	(0.032)
Psychology (N=87)	0.041	10,351	5,349	0.850
	(0.023)	(782)	(670)	(0.043)
Other (N=426)	0.067	10,846	4,633	0.780
	(0.013)	(532)	(331)	(0.022)

Notes:
Sample includes all U.S. citizens who did not report a disability, had no more than 9 months of graduate education, were not currently enrolled in school, received their BA prior to age 30, and borrowed money for their education.

Default should be increasing in the cost of re-payment (alternatively, the benefit of defaulting). Since the cost of re-payment is simply the amount owed, default rates should be increasing in the amount borrowed. Table 6 supports this prediction. Default rates are around 4-5% for those borrowing

Table 6: Default by Total Education Loan Amount

Loan Amount	N	Default Rate	Standard Error
\$1-4,999	696	0.051	0.009
\$5,000-9,999	756	0.042	0.008
\$10,000-14,999	662	0.059	0.010
\$15,000-19,999	351	0.062	0.014
\$20,000 +	330	0.094	0.018

Notes:
Sample includes all U.S. citizens who did not report a disability, had no more than 9 months of currently graduate education, were not enrolled in school, received their BA prior to age 30, and borrowed money for their education.

less than \$10,000, increasing to more than 9% for those borrowing \$20,000 or more. This relationship continues to hold after controlling for post-graduation earnings as we show below.

Since individuals with lower earnings are less able to make their loan payments, we might expect to observe a strong negative correlation between earnings and default. This does not appear to be the case as seen in panel (a) of Table 7, which reports on the relationship between 1997 earnings and default. Among those earning less than \$10,000 (less than 10% of the sample), fewer than four percent defaulted. However, among those earning \$10,000-20,000, default rates are as high as 12%. The likelihood of default then declines with earnings to 2.6% for those earning between \$40,000 and \$50,000, but it then rises again to nearly 5% for those earning \$50,000 or more.

The low default rates among those at the bottom of the earnings distribution may reflect the ability of these individuals to receive a deferment or forbearance on their loans, which temporarily delays payments for borrowers experiencing severe financial difficulties. In some sense, this is simply a labelling issue, since it reflects non-payment. More importantly, family income levels tend to be *higher* among those earning less than \$10,000 (78% had family income above \$25,000 in 1993 or 1996) than they are for those earning \$10,000-20,000 (only 51% had high family incomes in 1993 or 1996), so income from spousal employment may help reduce pressures to default for these borrowers. Turning to panel (c) we observe that default rates among those with very low family incomes are higher than among those with incomes ranging from ten to twenty thousand dollars. Default rates are monotonically declining in family income until the highest income category is reached.

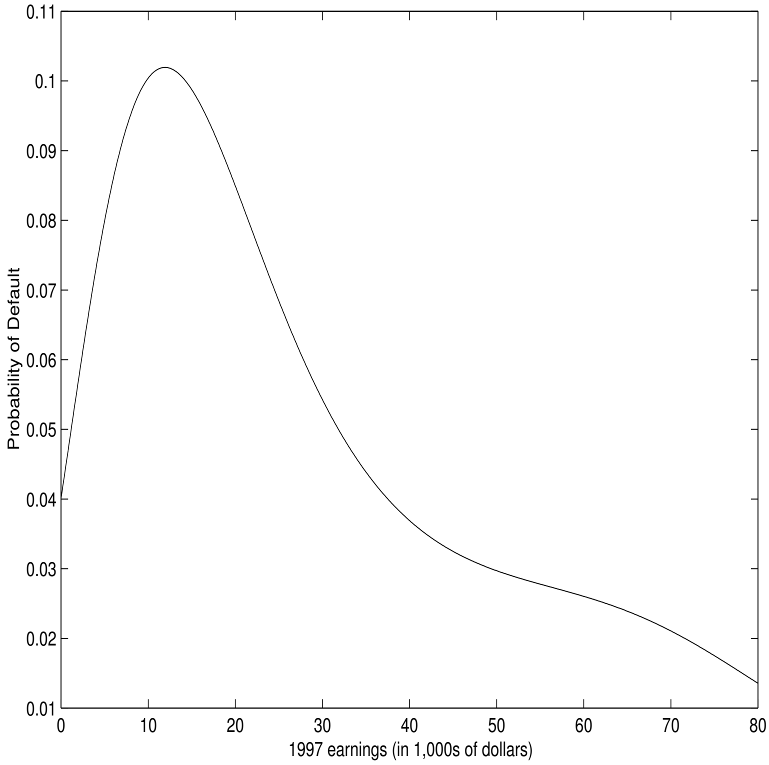
Those with high earnings (above \$50,000) also tend to have borrowed more for their education than all other income groups – on average, those earning more than fifty thousand dollars borrowed nearly \$4,000 more than those earning twenty to fifty thousand dollars. Their higher earnings should make re-payment easier, but their substantial debts may more than offset this effect. To explore this

Table 7: Default by Actual and Predicted 1997 Earnings and 1996 Family Income
(Males with Some Undergraduate Borrowing)

	Earnings/Income Category:					
	\$0-9,999	\$10,000-19,999	\$20,000-29,999	\$30,000-39,999	\$40,000-49,999	\$50,000 +
(a) Actual 1997 Earnings						
Default Rate	0.037	0.121	0.058	0.041	0.026	0.049
(Std. Error)	(0.012)	(0.018)	(0.009)	(0.009)	(0.011)	(0.016)
N	265	399	813	568	247	234
(b) Predicted 1997 Earnings						
Default Rate		0.041	0.070	0.059	0.029	
(Std. Error)		(0.012)	(0.010)	(0.008)	(0.010)	
N		331	792	983	344	
(c) 1996 Family Income						
Default Rate	0.170	0.119	0.065	0.052	0.014	0.042
(Std. Error)	(0.054)	(0.021)	(0.011)	(0.011)	(0.007)	(0.008)
N	59	279	551	503	335	792

Notes:
Sample includes all U.S. citizens who did not report a disability, had no more than 9 months of graduate education, were not currently enrolled in school, received their BA prior to age 30, and borrowed money for their education. Predicted earnings based on a Tobit specification that controls for SAT/ACT quartile, major, ethnicity, and months of graduate education.

Figure 1: Predicted probability of default by 1997 earnings (with average education debt)

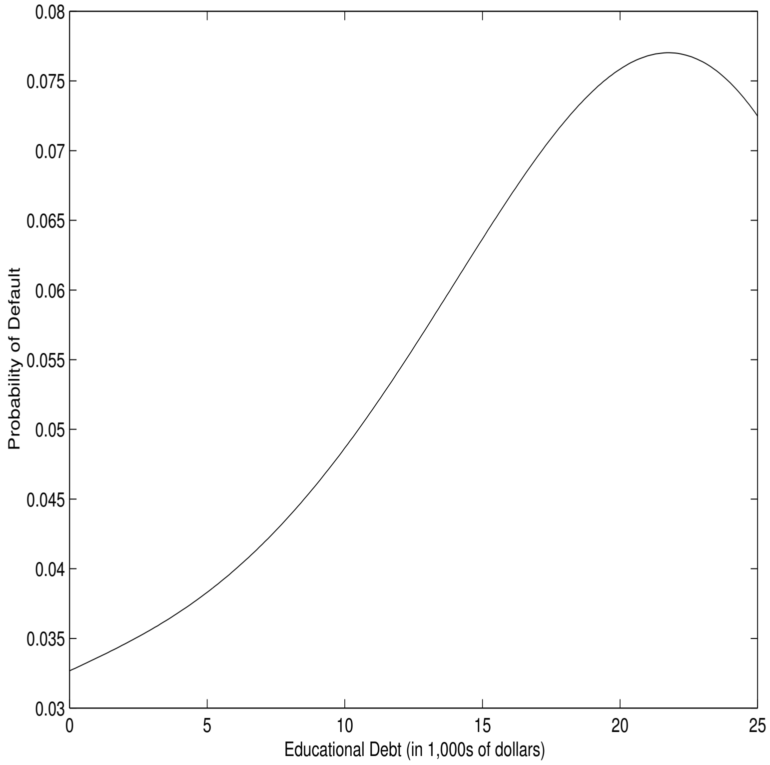


issue further, we estimate the probability of default as a function of both 1997 earnings and the total amount borrowed using a probit specification for default as a function of polynomials in both earnings and educational borrowing.[9] Figure 1 graphs the predicted probability of default by 1997 earnings for someone with the average level of educational borrowing. While the graph shows a steep increase in the probability of default at the very low end of the earnings distribution (note that 90% of the sample earned more than $12,000 in 1997), default is generally decreasing in earnings once we control for the amount borrowed. This suggests that the high default rates among those at the top end of the earnings distribution are the result of higher levels of education debt. Figure 2 shows that the probability of default is monotonically increasing in education debt after controlling for earnings.

Because earnings vary predictably with individual background and choice of college major and because earnings affect default, many individuals may know whether or not they are likely to default on their loans far before they do so. That is, black men with low SAT scores who choose to major in humanities ought to recognize that their post-graduation earnings will be low and that they are likely to be unable or unwilling to re-pay their loans. To examine this further, we estimate annual earnings in 1997 based on SAT/ACT quartile, undergraduate major, ethnicity, and the number of

[9] We estimate the model using polynomial terms through order six for both debt and earnings, since likelihood ratio tests did not reject that this specification fit the data better than lower order polynomials while tests rejected the inclusion of higher order polynomials.

Figure 2: Predicted probability of default by education debt (with average 1997 earnings)



months of graduate school (still considering those with no more than 9 months of graduate education but extending the sample to include non-borrowers as well) using a standard Tobit procedure. (See Appendix Table A-2 for the estimates.) We then form predicted 1997 earnings for each individual based on their characteristics. Panel (b) of Table 7 reports the relationship between default and this predicted earnings measure. Default rates increase from the \$10-20,000 to the \$20-30,000 income category, but then decline thereafter. It is impossible to know whether default rates would increase for borrowers with predicted earnings above \$50,000 (as is the case with actual earnings), since none of our sample members are predicted to have earnings that high.

Finally, we examine the relationship between gender, race, SAT/ACT scores, college major, and default conditional on actual earnings and total education loan amounts in Table 8. Specification (1) does not control for earnings or debt and serves as a baseline. The differences across groups shown in Tables 4 and 5 remain. All remaining specifications in the table control for sixth order polynomials in 1997 earnings and student loan amounts. Even after controlling for earnings and student debt in columns (2) and (3), default is U-shaped in SAT/ACT test scores and default rates are higher among blacks and hispanics. While there is some variation in the effects of college major on default rates, most differences are insignificant. Two obvious exceptions are math/science and humanities majors, which show the highest default rates conditional on race, gender, test scores, earnings, and debt. It is somewhat surprising how little most coefficient estimates change when we control for earnings and debt. Columns (4) and (5) condition on 1994 earnings in addition to 1997 earnings and education debt, which should help alleviate any biases due to measurement error in earnings and should help account for any effects that growth in earnings may have on default. The qualitative features with respect to test scores, race, and college major remain.

We also examine whether predicted earnings continue to affect default rates conditional on actual earnings and debt. Controlling for polynomials in debt and 1997 earnings, the estimated coefficient on predicted earnings is insignificant (-0.007 with a standard error of 0.006).

5 A Model of Schooling and Default

We develop a simple model of optimal human capital investment and default in an environment based on the U.S. federal student loan program. We begin by abstracting from the choice of college major and stigma of default.

5.1 Basic Model

Individuals are endowed with initial assets a (primarily reflecting parental transfers) that they may use for consumption, c, or investment in human capital, y. At the time of human capital investment, preferences are given by

$$u(c) + \beta \int v(A(z))dF(z). \tag{1}$$

Here, $u(\cdot)$ is the utility of consumption during the time of investment (youth), and $v(\cdot)$ represents the expected discounted utility after investment decisions have been made. Both are increasing, concave, and twice continuously differentiable functions. The parameter β represents the time discount factor, and $A(z)$ represents net resources available to the agent in the second period who has received a shock of z. $F(z)$ is the cdf of z. These shocks may take place during or after the education process, but they are not realized until *after* investment decisions have been made. A low z may, therefore, represent a worse than expected educational experience or a poor employment outcome. From an individual's perspective, shocks can also represent economy-wide conditions to the labor market.

While individuals may choose to save some of their initial assets, we focus attention on youth choosing to borrow from the federal loan program. We, therefore, assume that to finance their investments, individuals can borrow an amount d not to exceed the amount of investment or an upper limit on borrowing, $\hat{d}$. These constraints on borrowing reflect two important features of the current government lending environment: (1) loans are provided to cover schooling costs (with some provision for modest living expenses) and not consumption, and (2) students face an upper limit on the amount they can borrow. An interest rate of R is charged on these loans. For now, we ignore subsidies tied

Table 8: The Effect of Background Characteristics on Default
(Probit Coefficient Estimates)

Variable	(1)	(2)	(3)	(4)	(5)
Male	0.105	0.127	0.132	0.141	0.160
	(0.101)	(0.112)	(0.107)	(0.116)	(0.111)
SAT/ACT Quartile 2	-0.173	-0.214	-0.193	-0.201	-0.183
	(0.128)	(0.137)	(0.134)	(0.144)	(0.140)
SAT/ACT Quartile 3	-0.443	-0.523	-0.437	-0.481	-0.396
	(0.154)	(0.171)	(0.163)	(0.175)	(0.167)
SAT/ACT Quartile 4	0.020	0.006	0.102	-0.044	0.065
	(0.140)	(0.152)	(0.143)	(0.159)	(0.149)
Black	0.848	0.788	0.813	0.842	0.878
	(0.135)	(0.147)	(0.141)	(0.153)	(0.147)
Hispanic	0.408	0.475	0.531	0.418	0.489
	(0.179)	(0.188)	(0.182)	(0.200)	(0.192)
Asian	-0.024	0.216	0.301	-0.127	-0.035
	(0.343)	(0.358)	(0.338)	(0.461)	(0.439)
Business/management	0.110	0.264		0.308	
	(0.183)	(0.199)		(0.209)	
Education	0.136	0.229		0.210	
	(0.182)	(0.202)		(0.211)	
Engineering	-0.329	-0.052		-0.001	
	(0.275)	(0.301)		(0.325)	
Health professions	-0.423	-0.384		-0.571	
	(0.278)	(0.330)		(0.375)	
Public affairs/social services	0.048	0.176		0.180	
	(0.275)	(0.289)		(0.296)	
Biological sciences	0.124	0.096		0.001	
	(0.258)	(0.288)		(0.324)	
Math & science	0.347	0.574		0.651	
	(0.206)	(0.222)		(0.232)	
Social science	0.017	-0.042		-0.023	
	(0.209)	(0.234)		(0.242)	
History	0.424	0.225		0.289	
	(0.334)	(0.417)		(0.426)	
Humanities	0.508	0.594		0.568	
	(0.179)	(0.195)		(0.205)	
Psychology	0.060	-0.020		-0.011	
	(0.304)	(0.350)		(0.359)	
Log Likelihood	-392.12	-342.99	-355.05	-316.53	-329.05
Number of Observations	2,071	1,876	1,876	1,790	1,790
Earnings & Debt Controls:					
Total Education Loan Amount	No	Yes	Yes	Yes	Yes
1997 Earnings	No	Yes	Yes	Yes	Yes
1994 Earnings	No	No	No	Yes	Yes

Notes:
Standard errors in parentheses. Specifications controlling for earnings and and total loan amounts, include sixth order polynomials in those variables. All specifications control for months of graduate school.

to these loans, which introduce a wedge between the government lending interest rate and the market rate.

Given investment y, an individual of known ability e will earn a second period income of

$$w(z, e, y) = zey^{\alpha}, \tag{2}$$

where $\alpha \in (0, 1)$ and z is the education/labor market shock. The important features of the earnings process are: (1) expected earnings are increasing in human capital investments (where y may represent a greater quantity or quality of schooling), (2) expected earnings are increasing in known ability e, (3) ability and investment are complements, and (4) future earnings may be risky at the time of investment, as reflected in $z \geq 0$.

Once z is realized, individuals can decide whether or not to default on their loan obligations. If they choose to default, we assume that any of their earnings above some minimum amount $\underline{w}$ can be garnished at the rate τ. Thus, the effective garnishment schedule is given by $\max\{0, \tau(zey^{\alpha} - \underline{w})\}$. One could also assume that assets can be seized from defaulting agents, however, this will not play a role in our analysis of constrained borrowers who will have non-positive assets at the conclusion of their schooling.

For borrowers, the decision to default simply involves comparing the cost of re-payment, Rd with the cost of default. The re-payment decision is, therefore,

$$P(z, e, y, d) = \min\{Rd, \max\{0, \tau(zey^{\alpha} - \underline{w})\}\}.$$

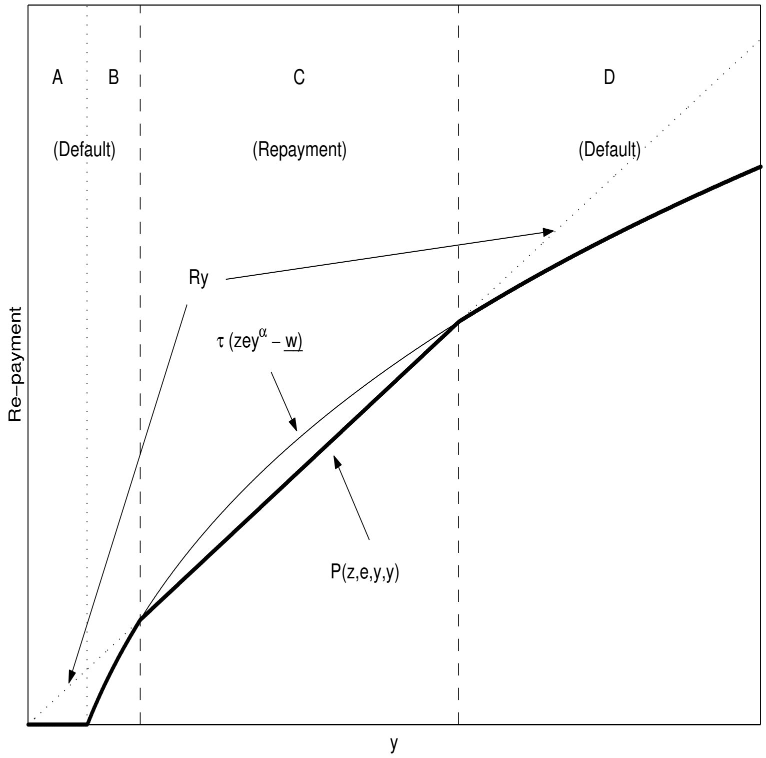
Among constrained agents (i.e. low a), the amount borrowed equals the amount invested ($d = y$) and

$$P(z, e, y, y) = \min\{Ry, \max\{0, \tau(zey^{\alpha} - \underline{w})\}\}.$$

For large $\hat{d}$, Figure 3 graphically represents a re-payment schedule as a function of y (assuming $y = d$) given (z, e). Notice that effective re-payment (dark solid line) is the lower envelope of the default (light solid line) and no-default (dotted line) payment schedules. Interestingly, given z, default occurs at the extremes – regions A, B, and D. Because of the limit on garnishments for earnings below $\underline{w}$, individuals with low investment and debt are able to keep nearly all of their earnings when they default. Thus, they are better off defaulting than re-paying their loans in regions A and B. Those with very high investment/debt (region D) are also better off defaulting, since their re-payment obligations exceed the garnishments they face. Only individuals with mid-level investment (region C) and earnings find it in their best interest to re-pay their loans.

The two default regions emerge for different reasons. The lower default region is created by the lower bound on wage garnishments. If $\underline{w}$ were equal to zero, there would be no default among those

Figure 3: Re-payment as a function of y given (z, e) when default occurs at extremes



investing very little as shown in Figure 4. The upper default region arises because of diminishing marginal returns to investment. While re-payment obligations increase linearly with investment and debt, garnishments increase linearly with earnings (above $[\underline{w}/(ze)]^{1/\alpha}$). For levels of investment slightly above $[\underline{w}/(ze)]^{1/\alpha}$, earnings increase substantially with investment until the penalty of default exceeds the cost of re-payment where region C begins. However, the concavity of earnings in investment implies that beyond some point, the cost of re-payment will again exceed the penalty of default (region D). For $\underline{w}$ high enough or τ low enough, the middle region of re-payment may not exist, in which case everyone will choose to default regardless of their debt and investment. This case is shown in Figure 5.

By choosing a sufficiently low $\hat{d}$, the government can eliminate the upper default region in Figures 4 and 5 altogether, but this must come at the expense of constraining investment. It is important to note that this upper region depends on observed ability, e, and unobserved shocks, z. While it is certainly feasible to condition upper loan limits on observed ability levels, it is not possible to condition those limits on z. Unless the range of shocks is quite limited, it would seem difficult to completely eliminate default at the upper end without seriously curtailing investment. In principle, a lower bound on investment and borrowing could also be imposed in an attempt to reduce default at that end of the investment schedule (regions A and B in Figure 3).

Figure 4: Re-payment as a function of y given (z, e) when $\underline{w} = 0$

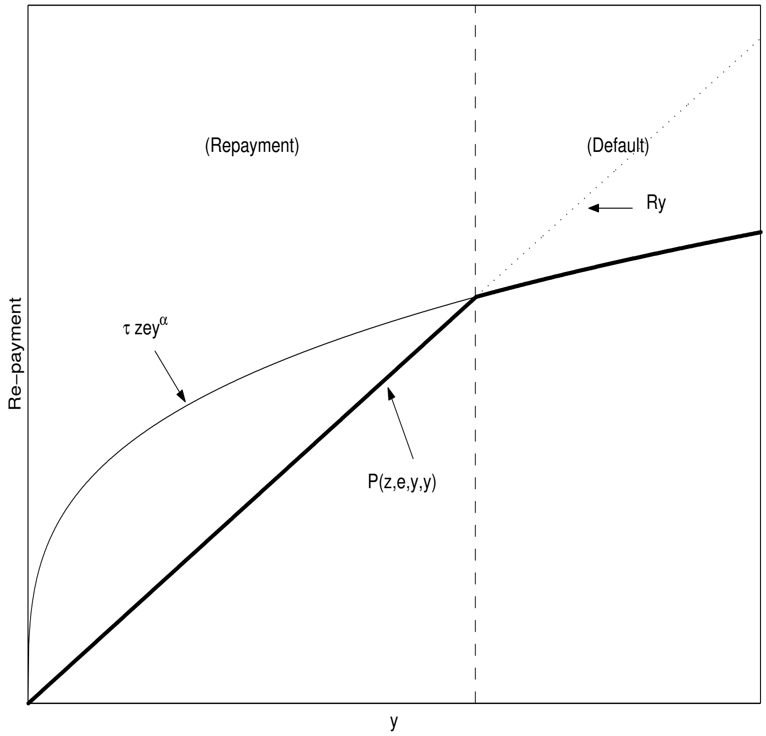
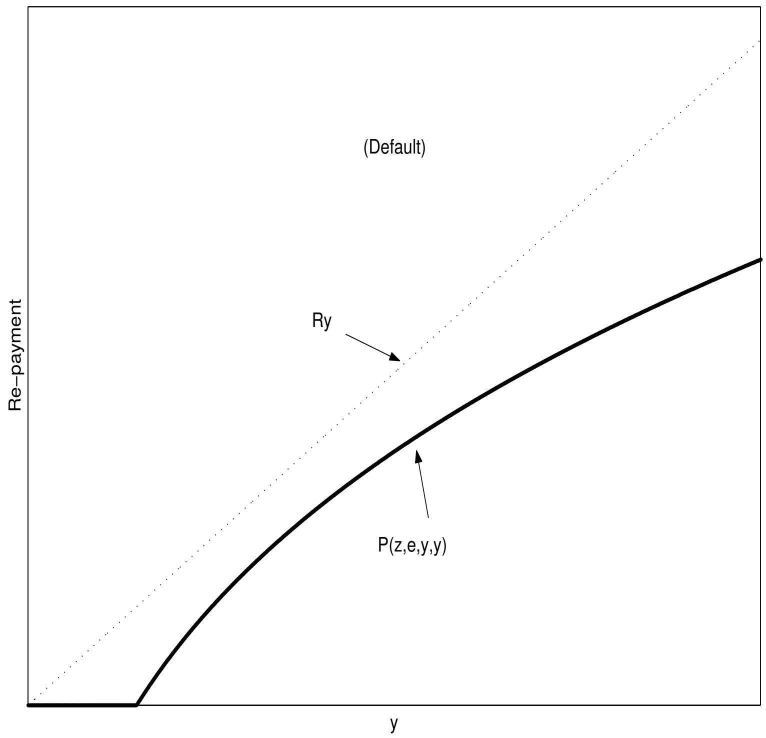


Figure 5: Re-payment as a function of y given (z, e) when everyone defaults



Given re-payment decisions, we can determine optimal investment, borrowing, and consumption decisions during the first period. Second period assets are given by

$$A(z, e, y, d) = w(z, e, y) - P(z, e, y, d). \tag{3}$$

Individuals, therefore, choose investment and consumption to maximize (1) subject to (2), (3) and $d \leq \min\{y, \hat{d}\}$.

Again focusing on constrained agents with low initial assets, the optimal investment decision maximizes expected utility for the next period. With low assets, individuals want to borrow to enhance consumption and investment. But, loans are limited to the amount of investment, so poor individuals are forced to consume from their initial assets and borrow to pay for their schooling (i.e. $c = a$ and $d = y$).

The option of default introduces a non-convexity into the problem. To find the optimum, we must consider interior optima as well as corner solutions for investment. First order conditions may not hold with equality at the optimal level of investment due to kinks in the effective re-payment function. Still *optimal investment is uniquely determined.* Assuming z has an atomless distribution, the first order condition for an interior optimum is[10]

$$\alpha e y^{\alpha-1} \int z v'(A(z, e, y, y)) dF(z) = \int v'(A(z, e, y, y)) \frac{\partial P(z, e, y, y)}{\partial y} dF(z). \tag{4}$$

At most, one value of y can satisfy this condition. To find the globally optimal investment, one has to compare the value attained from this local optimum with the one attained by investing the maximum, $y = \hat{d}$. This comparison must be made because individuals who invest an amount such that they will default with probability one are always better off investing and borrowing as much as possible, then defaulting. This will become more apparent below.

It is interesting to examine the interior first order condition to understand the implications of the option to default for optimal investment choices. Returning to Figure 3, we can distinguish four distinct re-payment regions that characterize the marginal cost of borrowing and investment conditional on (z, e). In regions A and D, observe that $\frac{\partial P(z,e,y,y)}{\partial y} < R$. On the other hand, $\frac{\partial P(z,e,y,y)}{\partial y} > R$ in region B, and $\frac{\partial P(z,e,y,y)}{\partial y} = R$ in region C. Thus, the marginal cost of borrowing is highest in region B and lowest in regions A and D.

It is quite possible that the distribution of z is such that, given optimal investment, a borrower is always in region C. In this case, the option of default has no effect on investment, and investment would be the same as in models with non-contingent loans and perfect enforcement. Still, credit constraints are binding as individuals would prefer to borrow more to increase current consumption.

[10] If one allows for mass points in the distribution of z, then instead of "=" the first order should be expressed in terms of inclusion (i.e. the left hand side term should be an element of the set defined by the right hand side, which may not be a singleton due to the kinks.)

Thus, while consumption is suboptimal, the federal lending program enables efficient schooling. This is generally consistent with the results of Keane and Wolpin [21] who find that individuals are severely constrained in their borrowing but that relaxing those constraints would have little effect on chosen schooling levels.

5.1.1 Full Certainty

Consider the case where z is known before investment decisions are made (and individuals are constrained borrowers). In this case, human capital investment entails no risk and agents maximize their private net financial return, regardless of the utility function $v(\cdot)$.

As noted above, optimal investment may be either an interior solution or at a corner, $y = \hat{d}$. Consider first the interior optimum, which must be characterized by equality between private marginal costs and benefits.

The marginal returns to investment are given by $MR(y) = \alpha z e y^{\alpha-1}$. Marginal costs are given by the derivative of $P(z, e, y, y)$ with respect to y, which is slightly more complicated as it incorporates the optimal default decision. In general, there are several regions for marginal costs. In the case underlying Figure 3,

$$MC(y) = \begin{cases} 0 & \text{if } y \leq y_0 & \text{(Region A)} \\ \tau\alpha z e y^{\alpha-1} & \text{if } y \in (y_0, y_1] & \text{(Region B)} \\ R & \text{if } y \in (y_1, y_2] & \text{(Region C)} \\ \tau\alpha z e y^{\alpha-1} & \text{if } y > y_2 & \text{(Region D)} \end{cases}$$

where $y_0 = \left(\frac{\underline{w}}{ez}\right)^{1/\alpha}$ is the cutoff where earnings just equal the minimum, $\underline{w}$. The values $\{y_1, y_2\}$ are roots of the equation $\tau(zey^\alpha - \underline{w}) - Ry = 0$. These values assume that there is a region for which $\tau\{zey^\alpha - \underline{w}\} > Ry$, but it is quite possible that ability may be low enough or $\underline{w}$ high enough that $\tau\{zey^\alpha - \underline{w}\} < Ry$ for all $y > 0$. These cases entail agents who always default.

As long as $\tau < 1$ the marginal return to investment is strictly greater than the marginal cost throughout regions A, B, and D. Thus, we will never observe investment levels within regions A or B, and investments in region D will be at the maximum allowable amount, $\hat{d}$. Investments in region C will be socially optimal, equating the marginal return to the gross interest rate.

The socially optimal investment (in a world without default) satisfies $MR(y) = R$:

$$y^s = \left(\frac{\alpha z e}{R}\right)^{1/(1-\alpha)}.$$

Privately optimal investment and default decisions will be given by

$$y^* = \begin{cases} \hat{d} & \text{if } y^s \leq y_1 & \text{(Default)} \\ \hat{d} & \text{if } y^s \in (y_1, \hat{d}) \text{ and } ze(y^s)^\alpha - Ry^s < (1-\tau)ze\hat{d}^\alpha + \tau\underline{w} & \text{(Default)} \\ y^s & \text{if } y^s \in (y_1, \hat{d}) \text{ and } ze(y^s)^\alpha - Ry^s \geq (1-\tau)ze\hat{d}^\alpha + \tau\underline{w} & \text{(Re-pay)} \\ \hat{d} & \text{if } y^s \geq \hat{d} \text{ and } ze\hat{d}^\alpha - R\hat{d} < (1-\tau)ze\hat{d}^\alpha + \tau\underline{w} & \text{(Default)} \\ \hat{d} & \text{if } y^s \geq \hat{d} \text{ and } ze\hat{d}^\alpha - R\hat{d} \geq (1-\tau)ze\hat{d}^\alpha + \tau\underline{w} & \text{(Re-pay)} \end{cases}$$

Privately optimal investment and default are more succinctly characterized in terms of ability. To simplify matters, assume $\underline{w} = 0$ so there is no minimum amount of earnings safe from creditors. Let $\underline{e}$ represent the ability level that equates the net earnings for investing $\hat{d}$ and defaulting with that associated with investing y^s and re-payment:

$$\underline{e} = \left(\frac{R}{\alpha z}\right)\left(\frac{1-\tau}{1-\alpha}\right)^{(1-\alpha)/\alpha}\hat{d}^{1-\alpha}.$$

Individuals with ability less than $\underline{e}$ will borrow $\hat{d}$ and default. Now, define $\bar{e}$ as the ability level for which $y^s = \hat{d}$ (i.e. the person who would optimally invest $\hat{d}$ when they intend to re-pay their loans):

$$\bar{e} = \left(\frac{R}{\alpha z}\right)\hat{d}^{1-\alpha}.$$

A person with ability above $\bar{e}$ is constrained from borrowing as much as he would like. Finally, let $\hat{e}$ represent the ability level for which someone is indifferent between re-paying and defaulting on a loan of size $\hat{d}$:

$$\hat{e} = \left(\frac{R}{\tau z}\right)\hat{d}^{1-\alpha}.$$

Those with ability above $\hat{e}$ will re-pay a loan of size $\hat{d}$.

If $\tau = \alpha$, then $\underline{e} = \bar{e} = \hat{e}$ and all individuals will invest and borrow $y^*(e) = \hat{d}$; those with ability below the cutoff over-invest and default while those above the cutoff under-invest and re-pay. Only those with ability equal to the cutoff invest the socially optimal amount.

If $\tau > \alpha$, then $\underline{e} < \hat{e} < \bar{e}$ and

$$y^*(e) = \begin{cases} \hat{d} & \text{if } e \leq \underline{e} & \text{(Default)} \\ y^s & \text{if } e \in (\underline{e}, \bar{e}) & \text{(Re-pay)} \\ \hat{d} & \text{if } e \geq \bar{e} & \text{(Re-pay).} \end{cases}$$

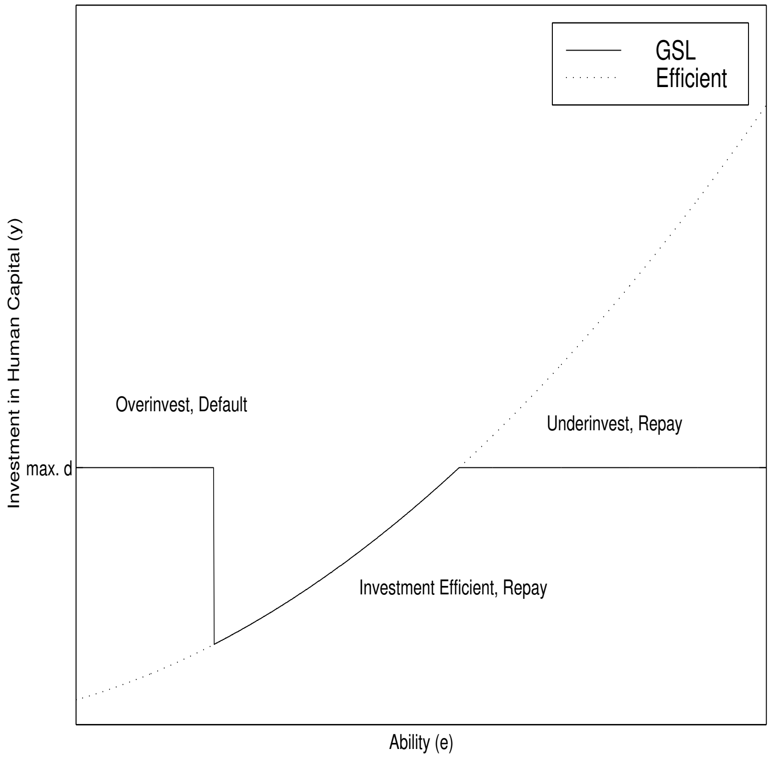
Figure 6 graphically represents this case. Investment is clearly a discontinuous non-monotonic function of ability with three distinct regions. Only the very low ability students default after borrowing the maximum. Middle ability students borrow the efficient amount and re-pay their loans. High ability students are limited to borrowing $\hat{d}$, and they re-pay their loans. Again, we observe over-investment among the least able and under-investment among the most able.

If $\tau < \alpha$, then $\bar{e} < \underline{e} < \hat{e}$ and

$$y^*(e) = \begin{cases} \hat{d} & \text{if } e < \hat{e} & \text{(Default)} \\ \hat{d} & \text{if } e \geq \hat{e} & \text{(Re-pay)} \end{cases}$$

In this case, all individuals borrow the maximum and only the most able re-pay their debts. Those with ability less than $\bar{e}$ over-invest, while those with ability above that threshold under-invest relative

Figure 6: Optimal Investment and Default Decisions by Ability



to the socially efficient amount. Interestingly, since $\bar{e} < \hat{e}$, middle ability individuals who choose to default – those with $e \in (\bar{e}, \hat{e})$ – actually under-invest relative to the socially efficient amount.

The introduction of $\underline{w} > 0$ will increase both $\underline{e}$ and $\hat{e}$, increasing default rates. When $\tau > \alpha$, the fraction of people over-investing increases. However, when $\tau < \alpha$, investment levels are unaffected and so there is no change in the efficiency of investment decisions. Once $\underline{w}$ becomes sufficiently large (as exhibited in Figure 5), all individuals borrow the maximum and default.

5.1.2 Risk and Uncertain Returns

Now, consider uncertainty in the returns to human capital investment, which may be due to shocks during or after the completion of school. For simplicity, assume that all agents face the same distribution for these shocks, z, which has the support $[0, \infty)$.

The lending program does not explicitly make repayments contingent on the realization of z. Yet, optimal repayment decisions will be made contingent not only on z but also on investment, y, ability, e, and the amount of debt, d.

As above, the optimal investment and debt decisions for someone of ability e are

$$\{d^*(e), y^*(e)\} \in argmax \left\{ u(a_0 + d - y) + \beta \int_0^\infty v\left(zey^\alpha - P(z, e, y, d)\right) dF(z) \right\}$$

subject to

$$d \leq \min\{y, \hat{d}\}.$$

Investment and default decisions interact. At the time individuals make investment decisions, agents consider their ability e and their ex-post optimal decisions in terms of repayment. Repayment decisions are based on previous investments through their effects on debt and earnings.

We now discuss the ex-post optimal default decisions, then returning to the investment decision.

Ex-Post Default

Consider an agent with ability e who has made an investment y and owes debt d. He will not repay anything (i.e. default without punishment) if the realization of z is so low that earnings are below the minimum threshold $\underline{w}$:

$$zey^{\alpha} \leq \underline{w}.$$

Thus, agents will default without punishment if

$$z \leq z_0(e, y) \equiv \left(\frac{\underline{w}}{ey^{\alpha}}\right).$$

This region depends on e and y, but it is independent of the level of debt, d, and the interest rate, R. Agents will find it optimal to fully repay their loans only if it is in their best interest. Repayment is optimal only if $\tau(zey^{\alpha} - \underline{w}) \geq Rd$, or

$$z \geq z_1(e, y, d) \equiv \frac{Rd + \underline{w}}{\tau ey^{\alpha}}.$$

This region depends on e, y, and d.

Since investment decisions affect the thresholds z_0 and z_1, they affect whether an agent defaults or not. The probability of default without punishment for an individual of ability e making investment y is given by

$$Pr(z \leq z_0(e, y)) = F\left(\frac{\underline{w}}{ey^{\alpha}}\right),$$

and the probability of default with or without punishment for debt d equals

$$Pr(z \leq z_1(e, y, d)) = F\left(\frac{Rd + \underline{w}}{\tau ey^{\alpha}}\right).$$

In principle, one could estimate the distribution of z without data on earnings, provided one knew the punishment policy, $\underline{w}$ and τ, and had data on default decisions, investment expenditures, loan amounts, ability, and interest rates. One might assume $e = exp(X\Gamma)$, where X contains variables affecting human capital productivity and market wages like SAT and ACT test scores, race, or college major. With this assumption and the assumption that $log(z) \sim N(0, \psi^2)$, one obtains a simple probit model with

$$Pr(default) = \Phi\left[\left(\frac{1}{\psi}\right) log[(Rd + \underline{w})/\tau] - X\left(\frac{\Gamma}{\psi}\right) - \left(\frac{\alpha}{\psi}\right) log(y)\right],$$

where $\Phi(\cdot)$ is the standard normal cdf. All unknown parameters, ψ, Γ, and α can be identified. If either τ or $\underline{w}$ is unknown, it can also be estimated (though both cannot be estimated simultaneously).

Data on earnings and ability (or determinants of ability) would allow for a more direct way to estimate $F(z)$. Conditional on true earnings and total debt, individuals either default or repay with probability one. However, if earnings are measured with error – either due to simple measurement error or due to the fact that we typically only observe earnings for a limited amount of time while default is likely to depend on a more permanent measure of earnings – then for any observed level of earnings and debt, the probability of default will be in $[0, 1]$. Suppose observed earnings, w, mis-measure true earnings, w^*, according to $w = w^* - \varepsilon$ and that $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. In this case,

$$Pr(default) = \Phi\left[\left(\frac{\underline{w}}{\sigma_\varepsilon}\right) + \left(\frac{R/\tau}{\sigma_\varepsilon}\right) d - \left(\frac{1}{\sigma_\varepsilon}\right) w\right].$$

With data on earnings, w, and total debt, d, one can consistently estimate σ_ε, R/τ, and $\underline{w}$. To the extent that interest rates are know, τ can be separately identified.

Given the above characterization, a simple probit specification for default as a linear function of debt and observed earnings should yield a positive coefficient on debt and a negative coefficient on earnings. The inverse of the coefficient on earnings provides an estimate of the variance in measurement error, while the ratio of the coefficients on debt and earnings provides an estimate of R/τ. When we empirically estimate such a model, we obtain parameter estimates for coefficients on earnings and debt that are of the predicted signs; however, the estimated intercept is negative, which is inconsistent with a non-negative $\underline{w}$.[11] The estimated effect of earnings on default is quite small, which produces an unreasonably large estimate of σ_ε. The implied estimate for R/τ is more reasonable at around 1.4. The fact that higher order polynomials in debt and earnings are preferred to the current specification that is linear in debt and earnings (see the discussion of Figures 1 and 2), implies a more complex punishment schedule than the simple linear 'tax' used here. Still, the general theoretical prediction that default should be decreasing in earnings and increasing in debt is supported empirically. The fact that SAT/ACT quartile and race affect default probabilities even after controlling for debt and earnings (see Table 8) suggests an important shortcoming of the model, which we discuss further below.

[11]The estimated coefficient on earnings (in thousands of dollars) is -.0057 and the coefficient on debt (in thousands of dollars) is .0081. The estimated intercept is -1.518.

Optimal Investment

We now consider optimal investment, again focusing on the case with low initial assets. Thus, $y = d$ and the default threshold for z becomes

$$z_1(e, y) \equiv \frac{Ry + \underline{w}}{\tau e y^{\alpha}}.$$

With this, optimal investment is

$$y^*(e) \in argmax \left\{ \int_0^{z_0(e,y)} v\left(zey^{\alpha}\right) dF(z) + \int_{z_0(e,y)}^{z_1(e,y)} v\left((1-\tau)zey^{\alpha} + \tau\underline{w}\right) dF(z) + \int_{z_1(e,y)}^{\infty} v\left(zey^{\alpha} - Ry\right) dF(z) \right\}$$

This problem does not have an analytical solution but can be solved numerically. Since this problem is not necessarily convex, the value of the interior optimum must be compared with the value of a corner solution, $y = \hat{d}$. This will generate interesting patterns for investment and the probability of default in relation to ability.

Ability, Investment, and Default Probabilities

If $\hat{d} = \infty$ (i.e. no limits on borrowing are in place), the optimal solution for all agents is to invest $y^*(e) = \infty$ and default with probability one. Thus, the problem is only interesting for finite values of $\hat{d}$. In this case, agents with very low ability, e, may still choose to make substantial investments, borrowing the maximum, knowing they will default with almost certainty. On the other hand, for high ability agents, the maximum loan amount may be less than they wish to invest. In this case, $y^*(e)$ may also equal $\hat{d}$ but for very different reasons. These agents will repay with very high probability. Thus, the model predicts that agents with extremely low and extremely high ability will both invest the maximum amount, but those with high ability will almost certainly repay while those with low ability will almost certainly default.

In general, when the optimal investment amount is $\hat{d}$, the default probabilities $F\left(\frac{R\hat{d}-\underline{w}}{\tau e\hat{d}^{\alpha}}\right)$ are locally decreasing in ability. This is not a global property, however, since an increase in e can induce discrete declines in $y^*(e)$ due to the non-convex nature of the problem.

Now consider interior optima $y^*(e) \in (0, \hat{d})$. At an interior, investment is generally an increasing function of e.[12] Will default probabilities decrease monotonically with e? Not necessarily. Monotonicity of the probability of default is determined by monotonicity of the cutoff z_1 evaluated at the optimum,

$$z_1^*(e) \equiv \frac{Ry^*(e) - \underline{w}}{\tau e (y^*(e))^{\alpha}}.$$

[12] One can establish this result analytically when individuals are risk neutral, and it is likely to hold more generally with risk aversion. However, investment solves a complicated first order condition that does not yield a clear cut relationship between y^* and e when individuals are risk averse.

This threshold may be non-monotonic in e, since $y^*(e)$ is typically increasing in e. For any given investment/debt level, more able agents are more likely to repay, since the punishment of default is increasing in earnings conditional on investment. However, more able agents generally choose to borrow and invest more, making repayment more costly. As a result, the net effect of ability on default is ambiguous.

At this point, it is clear that the current model has a serious limitation in terms of its ability to generate a U-shaped pattern for default in terms of ability as seen in the data. If optimal levels of debt/investment are sensitive enough to ability, it could, in principle, generate the desired default pattern. However, the high levels of default would be accompanied by much higher levels of debt and investment among the most able, which is not the case empirically (see Table 4). More simply, the model suggests that ability should not affect default once one conditions on earnings and debt; however, the estimates of Table 8 contradict this result.

To generate the empirical patterns in the data, forces that influence default without substantially affecting financial investment and borrowing must be introduced. We are currently considering a number of possible generalizations. (1) A stigma of default that is correlated with ability can generate a variety of default-ability relationships. (2) Other punishments like reporting to credit bureaus, which limits subsequent borrowing, may generate differential patterns in default by ability. (3) The ability of creditors to punish defaulting borrowers may vary by choice of occupation (e.g. it is impossible to garnish the wages of self-employed workers). Introducing an occupational decision may provide another source of variation in default by ability. (4) We are also considering models in which students can invest other inputs such as time and effort to enhance their earnings. In this case, the same level of total investment can be made with different levels of financial investments and indebtedness. Such a generalization would also enable an analysis of lending and subsidy policy on work among students. (5) We also plan to introduce the choice of college major.

6 An Efficient Lending System

To be done.

7 Conclusions

This paper is clearly a work in progress. Thus far, we have uncovered some interesting patterns in student loan default among college graduates. Our simple model of default and human capital investment is able to explain some of those patterns (the decline in default associated with earnings and increase associated with debt), but fails to explain others (namely, the effects of race and ability conditional on earnings and debt). We expect that a slightly more general, and more complex, model will remedy these current shortcomings. We envision a framework that can simultaneously explain the relationships between human capital investment, earnings, debt, choice of college major, and default, which can then be estimated using a more structural approach. Such a model should offer many insights into optimal government lending and subsidy policy and the role of credit constraints in our current environment.

References

[1] S. R. Aiyagari, J. Greenwood, and A. Seshadri. Efficient Investment in Children. *Journal of Economic Theory*, forthcoming.

[2] F. Alvarez and U. Jermann. Asset Pricing when Risk Sharing is Limited by Default. Working Paper, University of Chicago and The Wharton School, 1997.

[3] D. Andolfatto and M. Gervais. Social Security and the Payroll Tax with Endogenous Debt Constraints. Working Paper, 1999.

[4] O. Attanasio and V. Rios-Rull. Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare. University of Pennsylvania Working Paper, 2001.

[5] G. Becker. *Human Capital*. Columbia University Press, New York, 1964.

[6] College Board. *Trends in Student Aid*. College Board Publications, New York, 2001.

[7] S. Cameron and J. Heckman. Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males. *Journal of Political Economy*, 106(2):262–333, April 1998.

[8] S. Cameron and J. Heckman. The Dynamics of Educational Attainment for Blacks, Whites and Hispanics. Working Paper, 1999.

[9] S. Cameron and C. Taber. Borrowing Constraints and the Returns to Schooling. Working Paper, 1999.

[10] D. Card. Earnings, Schooling, and Ability Revisited. *Research in Labor Economics*, 14:23–48, 1995.

[11] P. Carniero and J. Heckman. The Evidence on Credit Constraints in Schooling. Working Paper, 2002.

[12] E. Caucutt and K. Kumar. Higher Education Subsidies and Heterogeneity. Working Paper, 2000.

[13] M. Dynarski. Who Defaults on Student Loans? Findings from the National Postsecondary Student Aid Study. *Economics of Education Review*, 13(1):55–68, 1994.

[14] D. Elwood and T. Kane. Who is Getting a College Education? Family Background and the Growing Gaps in Enrollment. In S. Danziger and J. Waldfogel, editors, *Securing the Future: Investing in Children from Birth to College*, chapter 10. Russell Sage Foundation, New York, 2000.

[15] S. Fay, E. Hurst, and M. White. The Household Bankruptcy Decision. Working Paper, 2000.

[16] J. Fernandez-Villaverde and D. Krueger. Consumption and Savings over the Life Cycle: How Important are Consumer Durables? Working Paper, 2001.

[17] L. Greene. An Economic Analysis of Student Loan Default. *Educational Evaluation and Policy Analysis*, 11(1):61–68, 1989.

[18] R. Gropp, J. K. Scholz, and M. White. Personal Bankruptcy and Credit Supply and Demand. *Quarterly Journal of Economics*, 112(1):217–51, Feb. 1997.

[19] E. Hanushek, K. Yilmaz, and C. Leung. Redistribution Through Education and Other Transfer Mechanisms. Working Paper, 2001.

[20] T. Kane. College Entry by Blacks Since 1970: The Role of College Costs, Family Background, and the Returns to Education. *Journal of Political Economy*, 102(5):878–911, Oct. 1994.

[21] M. Keane and K. Wolpin. The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment. Working Paper, 1999.

[22] T. Kehoe and D. Levine. Debt-Constrained Asset Markets. *Review of Economic Studies*, 60(4):865–88, Oct. 1993.

[23] N. Kocherlakota. Implications of Efficient Risk Sharing without Commitment. *Review of Economic Studies*, 63(4):595–609, 1996.

[24] D. Krueger and F. Perri. Risk Sharing: Private Insurance Markets or Redistributive Taxes? Working Paper, 2001.

[25] L. Lambertini. Social Security with Endogenous Debt Limits. Working Paper, 1999.

[26] L. Lochner and A. Monge-Naranjo. Human Capital Formation with Endogenous Credit Constraints. Working Paper, 2002.

[27] G. Loury. Intergenerational Transfers and the Distribution of Earnings. *Econometrica*, 49(4):843–67, 1981.

[28] A. Monge-Naranjo, J. Cascante, and L. J. Hall. Enforcement, Contract Design and Default: Exploring the Financial Markets of Costa Rica. In M. Pagano, editor, *Defusing Default: Incentives and Institutions*. John Hopkins University Press, New York, 2001.

[29] US Department of Education. Accountability for Results Works: College Loan Default Rates Continue to Decline. Press Release, September 19, 2001.

[30] M. Pagano. *Defusing Default: Incentives and Institutions.* John Hopkins University Press, New York, 2001.

[31] W. Wilms, R. Moore, and R. Bolus. Whose Fault is Default? A Study of the Impact of Student Characteristics and Institutional Practices on Guaranteed Student Loan Default Rates in California. *Educational Evaluation and Policy Analysis*, 9:41–54, 1987.

Table A-1: Sample Background Characteristics for Undergraduate Borrowers

Variable	N	Mean	Standard Error
Male	2,796	0.474	0.009
White	2,783	0.825	0.007
Black	2,783	0.079	0.005
Hispanic	2,783	0.061	0.005
Asian	2,783	0.029	0.003
SAT/ACT Quartile 1	2,464	0.277	0.009
SAT/ACT Quartile 2	2,464	0.290	0.009
SAT/ACT Quartile 3	2,464	0.241	0.009
SAT/ACT Quartile 4	2,464	0.193	0.008
Major: business/management	2,795	0.248	0.008
Major: education	2,795	0.125	0.006
Major: engineering	2,795	0.074	0.005
Major: health professions	2,795	0.076	0.005
Major: public affairs/social services	2,795	0.031	0.003
Major: biological sciences	2,795	0.037	0.004
Major: math & science	2,795	0.062	0.005
Major: social science	2,795	0.078	0.005
Major: history	2,795	0.015	0.002
Major: humanities	2,795	0.076	0.005
Major: psychology	2,795	0.028	0.003
Major: other	2,795	0.151	0.007

Notes:
Sample includes all U.S. citizens who did not report a disability, had no more than 9 months of graduate education, were not currently enrolled in school, received their BA prior to age 30, and borrowed money for their education.

Table A-2: Tobit Estimates for 1997 Earnings

Variable	Estimate	Standard Error
Intercept	24,214	1,032
Male	8,376	695
SAT/ACT Quartile 2	938	903
SAT/ACT Quartile 3	3,348	945
SAT/ACT Quartile 4	3,544	1,005
Black	-1,553	1,475
Hispanic	-3,993	1,606
Asian	5,065	1,893
Business/management	4,637	1,177
Education	-15,442	1,200
Engineering	10,036	1,483
Health professions	7,909	1,412
Public affairs/social services	1,149	1,862
Biological sciences	-5,118	1,800
Math & science	5,765	1,556
Social science	2,899	1,272
History	-6,025	2,460
Humanities	-3,311	1,300
Psychology	-1,259	2,006
Months of Graduate School	-2	232

Notes:
Sample includes all U.S. citizens who did not report a disability, had no more than 9 months of graduate education, were not currently enrolled in school, and received their BA prior to age 30. There are 4,019 observations and the log likelihood value is -42,364.