When Do Students and Taxpayers See a Return? Optimal Accountability Thresholds in Higher Education

Douglas A. Webber *

8-24-19

Abstract

I conduct policy simulations to determine the optimal level at which to set thresholds that colleges are eligible to participate in the federal student loan program. Regardless of whether a policymaker wishes to maximize student outcomes or safeguard taxpayer dollars, I find that thresholds based on three-year repayment rates should be set in the 15%-20% range. Results examining students' financial welfare are relatively insensitive to modeling choices. However, thresholds aimed at minimizing taxpayer losses are sensitive to certain choices, and may be much higher (e.g. excluding more schools) for modest changes in assumptions.

JEL Codes: I21, I22, I23, I26, I28

Keywords: Higher Education Accountability; Student Loans; Returns to College

*I have greatly benefited from the advice of Robert Kelchen, Adam Looney, J. Catherine Maclean, Kenneth Megan, Barmak Nassirian, and Dubravka Ritter. Affiliation: Temple University, Email: douglas.webber@temple.edu, Address: 1301 Cecil B. Moore Ave Ritter Annex Office 877.

1 Introduction

Higher education has come under increased scrutiny in recent years due to a ballooning student debt portfolio at the national level (currently at $1.6 trillion) and countless examples of students struggling with their individual debt burdens. A number of high-profile policy solutions have been proposed to address the oft-mentioned "student debt crisis", including debt forgiveness,[1] free college,[2] and doubling federal need-based funding such as the Pell grant.[3]

Running parallel to these student-centric proposals are reforms aimed at the institutional side of the market. These range from informational improvements, which aim to provide better data on student outcomes at each school/program,[4] to formally sanctioning institutions whose students appear to receive little value-added from their postsecondary investment. This paper seeks to provide guidance to policymakers regarding this latter set of policy proposals.

There are two main rationales for the government to provide oversight and maintain a system of accountability in the market for higher education: protecting taxpayer funds, and protecting the time and monetary investment of students. Given that student loan debt cannot be forgiven through bankruptcy (except in extreme circumstances) this second rationale is especially critical, as poor outcomes can financially cripple an individual for years or decades.

Policymakers and researchers generally agree that the current accountability system which governs the higher education landscape is inadequate. For example, 709 schools have loan portfolios where less than 25% of students have made any progress paying down the principal balance on their loans three years after leaving school and entering repayment.

[1] https://medium.com/@teamwarren/im-calling-for-something-truly-transformational-universal-free-public-college-and-cancellation-of-a246cd0f910f

[2] https://www.sanders.senate.gov/download/collegeforallsummary/?inline=file

[3] https://ticas.org/sites/default/files/pub_files/pell_recs_one_pager.pdf

[4] https://www.insidehighered.com/news/2019/05/22/trump-administration-releases-new-program-level-loan-data

However, there is a disconnect between many in the policy community over what the actual goals of accountability policies are. One side tends to focus their rhetoric on the protection of taxpayers' investment in higher education, while others view the chief goal of accountability as a means of consumer (e.g. student) protection.

While both goals are certainly laudable, the disconnect can be problematic. Any policy which maximizes student outcomes likely also includes some degree of government subsidization. Similarly, a policy which maximizes taxpayers' return almost certainly harms some share students through inadequate access to financial capital. Making matters worse, there is no work to my knowledge which attempts to quantify these gains/losses, making it all the more difficult to come to a consensus. This paper attempts to fill that void.

I examine a key decision the federal government faces, how to decide which institutions are granted access to the student loan system. I first discuss the conceptual tradeoffs that exist and what an "optimal" policy looks like, depending on the objective function a policymaker is seeking to maximize. I then use data on educational outcomes and lifecycle earnings to provide concrete estimates of optimal policy thresholds under a variety of different assumptions regarding the costs and benefits (and the weights we place on them) of governmental involvement in higher education.

I specifically look at the implementation of a threshold based on an institution's student loan "repayment rate", which measures the proportion of students who have made some progress on their principal loan balance (e.g. paid down at least $1) within three years of leaving school. Such a system has been proposed by policy analysts and in legislation in recent years.

I find that the optimal repayment rate threshold lies in the 15%-20% range under the most comprehensive models, regardless of whether the objective is to maximize government revenues or individual incomes. However, the threshold which maximizes government revenues is much more sensitive to individual modeling decisions, and is considerably higher under even small changes in the assumed discount rate or fraction of tax revenue generated

from college education. These results can be interpreted as 1) schools below these benchmarks cost the government more money than they generate in loan repayments and taxes, and 2) these schools leave the average student worse off for having attended the institution.

The impact of using a repayment rate threshold in this range to dictate access to the student loan program varies considerably even within this small bandwidth. On the low end, roughly 75 schools which collectively enroll 30,000 students would face sanctions if the threshold were set a 15%. At 20%, however, the population impacted rises to 300 institutions and more than 200 thousand students. This highlights perhaps the central tension facing legislators and regulators in the higher education accountability sphere: 1) there is likely a large number of students attending schools which leave them (and taxpayers) worse off, and 2) small design changes in accountability programs can have sizable consequences in the market for higher education.

The paper is organized as follows: Section 2 discusses the institutional details of higher education accountability and the related academic literature. Section 3 presents a simple conceptual framework for evaluating optimal policy choices in the context of higher education. Section 4 describes the data and empirical methodology. Section 5 presents the results, and Section 6 concludes.

2 Previous Literature

Institutional Details of Accountability in Higher Education

The federal government began collecting and publishing institution-level measures of student loan performance in 1989 in response to concerns about persistently high student loan default rates. After a brief phasing-in period, Congress tied an institution's ability to receive student loans to a set of thresholds based on two-year Cohort Default Rates (CDR). This type

of CDR reflects the proportion of student who default on their student loans within two years of officially leaving school (either graduating or dropping out). The Higher Education Opportunity Act of 2008 slightly altered the regulations to instead depend on three-year CDR thresholds amid concerns that following students for two years was not long enough to adequately assess the financial health of a loan portfolio. The current definition of "default" within a student loan context requires no payments to be made for 270 consecutive days.

In order to lose Title IV eligibility (the ability to participate in the student loan system), an institution must have a CDR of greater than 30% for three consecutive years, or a CDR of greater than 40% for a single year. Several criticisms have emerged over the use of CDR thresholds in this context. First, default is a worst-case scenario which does not accurately reflect the financial health of a students' ability to repay their loans. A former student could pay $1 toward their loan every six months and stay out of default, but would clearly be in severe financial trouble based upon nearly any other metric.

Second, very few schools have CDR's in the vicinity of the thresholds despite having poor outcomes along many other earnings and loan repayment metrics. Only a handful of (typically small) schools per year are even at risk of CDR-based penalties.[5] Thus, despite many institutions having poor loan-based outcomes,[6] virtually no schools actually feel any pressure from CDR thresholds.

This is due in part to a dilution of the definition of default following the advent/growth in popularity of Income Driven Repayment (IDR) plans. Under IDR plans,[7] students who earn below certain income levels (typically 150% of the poverty line for a given family size) are not responsible for making any student loan payments, and thus will not be considered as defaulting even if they make no payments for years at a time. Moreover, due to the specific definition of default, it is possible for schools to manipulate their CDR by encouraging

[5] https://www.thirdway.org/memo/hea-replacement-for-cohort-default-rate-leaves-more-questions-than-answers

[6] 709 schools enrolling more than 700 thousand students have fewer than 25% of students making any progress on their principal loan balance after three years out of school.

[7] There have been many different incarnations of such plans over time, the different variations can be found here https://studentaid.ed.gov/sa/repay-loans/understand/plans/income-driven.

students to enter programs such as forbearance until they are outside of the three-year monitoring window. Indeed, at least 1,300 institutions employ consultants to help "manage" their default rates.[8]

A final concern with CDR-based thresholds is their blunt nature, denying federal financial aid to all students at an institution which barely fails the threshold test while doing nothing to schools which barely pass. While this is a limitation of any threshold-based accountability system, it may be particularly problematic in the case of higher education because of heterogeneity in outcomes across programs within an institution. For example, should nursing students be denied student loans because they attend a school over-populated with degrees which underperform in the labor market? Conversely, programs with extremely poor outcomes can be allowed to continue if they exist at an institution with stronger outcomes in other departments.

One solution proposed in both congressional legislation[9] and via think tanks[10] is to replace the CDR with a different measure of a student loan portfolio's financial health, the repayment rate. Among student loans, a repayment rate is defined as the proportion of students who have made at least $1 of progress paying down their principal loan balance. Such a change solves many of the issues with CDR's listed above. Repayment rates do a much better job of illustrating borrowers who are actually struggling with repaying their loans, are not manipulable (unless institutions are willing to pay off students' accrued interest), and are not diluted by the rise of IDR-style programs. Moreover, if thresholds are applied at the program level, rather than at the institution level, the bluntness of the penalty will be far less crucial.[11]

Figure 1a plots the overall distribution of three-year repayment rates, and Figure 1b plots separate distributions by institutional control. Both the overall spread, and the differences

[8] https://www.washingtonpost.com/news/grade-point/wp/2018/04/26/colleges-are-using-consultants-to-manipulate-student-loan-default-rates-gao-says/

[9] https://www.shaheen.senate.gov/imo/media/doc/Student%20Protection%20and%20Success%20Act%20of%202019.pdf

[10] https://www.chronicle.com/article/Student-Loan-Default-Rates-Are/228771

[11] An excellent discussion of this and other policies discussed in this paper can be found in Baum and Schwartz (2018).

across type are striking in their magnitude. While these figures should not be interpreted causally due to differences in the relative student populations (more on that issue below), they certainly raise concerns that a non-trivial fraction of students may not be receiving a positive return on the educational investment made by themselves and taxpayers. It should be noted that these figures and the total number of students included in later tables include all Title IV eligible schools (e.g. they include community colleges and four-year schools). I chose to include all institutions in these figures because accountability programs such as the CDR thresholds have historically applied to all sectors regardless of degree type. While I focus on the returns to a four-year degree in the policy simulations below, my results are similar if I instead focus on two-year degrees.

Academic Literature on Accountability

This section presents a brief summary of the literatures which are touched on by this paper. For a broader overview of the higher education fiscal landscape, see Ehrenberg (2012). This paper contributes to three related literatures: the returns to a college education, student loans, and accountability programs in higher education.

The literature which estimates the returns to college is far too broad and diverse to comprehensively describe here. Focusing on recent work, researchers have generally found large causal returns to college attendance/completion across a wide variety of empirical approaches. These include control function style approaches using detailed survey/administrative data (Webber, 2014; Andrews et al., 2016), the exploitation of exogenous discontinuities in admissions/academic probation (Zimmerman, 2014; Ost et al., 2018), and structural approaches (Bhuller et al., 2017).

In order to attain the large returns documented above, many students finance college fully or in part through the student loan system, described in great detail in the review article

Avery and Turner (2012). While the availability of student loans generally improves outcomes (Marx and Turner, 2019), the current system can hardly be characterized as functioning well. Goldrick-Rab (2016) documents myriad inequities in the current system which further disadvantages first-generation students and individuals from low-income backgrounds. Barr et al. (2018) puts the U.S. student loan repayment system in context with those of other countries, illustrating several areas where improvements could benefit students.

There are two chief justifications for having a strong accountability system within higher education: protection of the taxpayers' investment and protection of students from spending money and time on low-quality programs. If all programs yielded the same outcomes, or at least the same outcomes conditional on student ability, then there would be little to no need for such safeguards. But the literature on returns to college quality generally finds large disparities across schools (Hoekstra, 2009; Cohodes and Goodman, 2014; Andrews et al., 2016; Scott-Clayton, 2017; Canaan and Mouganie, 2018). A notable exception to this finding comes from Dale and Krueger (2014), which finds no differences in returns on average across a set of relatively selective schools. An excellent reconciliation of this literature can be found in Hoxby (2018), which examines value-added and productivity (return on spending) across the majority of the higher education landscape. She finds practically no differences among highly selective schools, but significant dispersion in quality among non/less selective schools.

Political discussions surrounding accountability often focus on the for-profit sector due to the sector having worse student outcomes (e.g. loan default rates, earnings) on average(Cellini, 2012; Cellini and Chaudhary, 2014; Cellini and Turner, 2019). This led to the "Gainful Employment" regulations implemented by the Obama administration which targeted vocational programs, which comprise much of the for-profit industry. This targeting likely doomed the ultimate implementation of the regulations, as they were rendered effectively toothless early in the Trump administration. In addition to the reforms described in this paper, other federal accountability mechanisms such as "risk-sharing" have been pro-

posed and analyzed (Webber, 2017), but as of this date none have been implemented.

At the state level, however, there has been a surge in the incidence of accountability programs and research evaluating the impacts. The modal program, known as Performance Based Funding (PBF) links state funding to outcomes such as graduation rates and enrollment of students from disadvantaged backgrounds. For an overview of PBF and evaluations of various programs, see recent work by Shin (2010), Hillman et al. (2014), and Kelchen and Stedrak (2016). For a complete accounting of all accountability systems faced by U.S. higher education, see the recent book Kelchen (2018a).

Given that this paper seeks to provide context for policymakers on where to set repayment rate thresholds, an especially important literature to be aware of is that which examines the consequences for students and institutions of failing the CDR thresholds. Darolia (2013) documents significant declines in enrollment at sanctioned institutions using a regression discontinuity design. Cellini et al. (Forthcoming) further show spillover effects for sanctioned for-profit institutions (positive for local community colleges, negative for local non-sanctioned for-profits) and an overall two percent decline in total local market enrollment in college. The latter effect is most pronounced among localities with few alternative postsecondary options. This highlights a key challenge of any accountability system: the consequences of student debt from low-performing programs can be very damaging to future financial health,[12] but the negative tradeoffs of reduced access, particularly among diverse populations, must also be considered.

[12]These negative impacts are most pronounced among minority student populations who often do not have the safety net of parental wealth to protect them from the worst outcomes such as default (Addo et al., 2016).

3 Theoretical Model

When discussing student loan policy, it is often not clear exactly which costs and benefits are being considered by policymakers. Different constituencies are likely trying to maximize distinctly different objective functions. This paper aims to provide some useful context for how to think about optimality in the context of accountability thresholds like those described in the prior section. While the main contribution of this work is in the range of explicit policy recommendations arising from the empirical analysis, it is also useful to lay out a simple theoretical framework that can guide decisionmakers.

Appeals for greater accountability in higher education generally take one of two forms: 1) safeguarding taxpayer dollars, or 2) protecting individual students from predatory/low-value programs. An appeal of the first type might argue that the government should not provide financial backing to programs where, in expectation, the investment will yield a negative return to the taxpayer. In other words, the optimal threshold should be set where the expected return on investment exactly equals the costs to the government of providing the loans.

Given the current structure of the financial aid system, and it's many generous features (such as debt forgiveness after 20 years), this is a fairly high bar. Importantly, it ignores the many private and social benefits that individuals receive from a college degree. Appeals of the second type, concerned only with individual outcomes, would prefer a threshold such that a typical student at the marginal school is just as well off whether they attend that school or not.

For those whose chief concern is the status of the government's loan portfolio, an optimal repayment rate threshold can be derived from Equation 1.

$$\int_{z}^{\infty} \delta(X)v(X)dH(X) - A[1 - H(z)] - c\int_{z}^{\infty} v(X)dH(X) \tag{1}$$

The optimal threshold, z^*, is chosen to maximize the government's investment in the

student loan program. $v(X)$ represents the total portfolio balance for schools at a particular repayment rate X, and $\delta(X)$ denotes the marginal return on investment to such schools. $\delta'(X)$ is assumed to be greater than or equal to 0. This is not a particularly strong assumption, since the repayment rate is mechanically related to return on investment in a world where students' loan repayment obligations are based on their incomes (as is becoming increasingly popular and as I assume in the policy simulations below).

$H()$ is a cumulative distribution function indexing all repayment rates. A denotes the administrative costs borne by the government of providing student loans to a set of schools, and is assumed to be fixed for a given loan portfolio size. Finally, c represents the marginal opportunity cost of the government's investment in providing student loans. This could be viewed as 0, the current interest rate on the national debt, or the return on investment from some other potential government activity (the sensitivity to such choices will be discussed in the empirical exercise below).

Applying the first fundamental theorem of calculus to Equation 1, we obtain the following condition for an optimal repayment rate threshold:

$$v(z^*)[\delta(z^*) - c] = A \tag{2}$$

Intuitively, Equation (2) states that the marginal social return (net of opportunity costs) to lowering the threshold must equal the additional administrative costs of providing loans to more students.

If instead we are interested in maximizing the net return to individuals, we would be interested in maximizing the objective function in Equation (3):

$$\int_{z}^{\infty} [u(Y_{Col}) - u(Y_{HS})] dH(Y_{col}) \tag{3}$$

where u() represents individuals' utility function, $H()$ is a cumulative distribution function indexing all repayment rates, and Y_{Col} is the net present value of earnings for college

graduates from a school with a given repayment rate.

$$Y_{Col} = \sum_{t=1}^{N} \beta^t NetEarnings_{Col,t} \tag{4}$$

Y_{HS} denotes the net present value of earnings for high school graduates.

$$Y_{HS} = \sum_{t=1}^{N} \beta^t NetEarnings_{HS,t} \tag{5}$$

The optimal repayment rate threshold in this case will satisfy the following condition

$$u(z^*) = u(Y_{HS}) \tag{6}$$

In other words, students graduating from a school which exactly meets the critical threshold will be just as well off as the average high school student.

Neither the results in Equations (2) or (6) are particularly surprising. They do, however, provide a baseline for how to think about the gatekeeping role that federal policymakers have in the market for higher education. In the remainder of the paper, I use data on institutional and student outcomes to provide estimates of the optimal repayment rate thresholds under a variety of different assumptions of the costs/benefits of college to society and individuals.

4 Data and Methodology

The goal of this paper is to provide concrete estimates of thresholds that the federal government could use to determine eligibility of an institution to participate in the student loan system, otherwise known as Title IV. Regardless of whether you are concerned with safeguarding taxpayer dollars or producing the best possible outcomes for students, the first step in such a policy simulation is to construct age-earnings profiles for a population of students.

I use a simplification of the lifecycle earnings simulation approach from Webber (2014)

and Webber (2016) to create a synthetic population which has the same distribution of earnings (both cross-sectionally and over time) as is observed in the real world. From here, I can evaluate either the government's financial position, or the relative welfare of college versus high school graduates, under any loan repayment regime which is based on the borrower's income.

Specifically, I use the five most recent waves of the American Community Survey (ACS; 2013-2017) to estimate Equation (7) for each of 9 age groups (18-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-55, 56-60, 61-65):

$$y_{ij} = \check{\beta}_0 + \check{\beta}_1^{(j)} age_{ij} + \check{\beta}_2 Black_i + \check{\beta}_3 Hisp_i + \check{\beta}_4 Gender_i + \check{\delta} educ_i + \varepsilon_{ij} \tag{7}$$

The dependent variable is the natural log of prior year earnings, and the vector educ includes indicator variables for educational attainment (for the main analyses this is limited to "some college" and "college graduate" (individuals with less than a high school diploma or postgraduate work are excluded from the sample to make it as homogeneous as possible). The coefficients on each education category within each age grouping, as well as the variance of residual log earnings, $\sigma^2_{educ,j}$, for each education category and age grouping are saved for use in the subsequent lifecycle earnings simulation.[13]

Log earnings are then simulated for each year of labor force participation based on the following equation:

[13] In an attempt to deal with cohort-specific effects on age-earnings profiles, I also repeat the analysis with a longer set of pooled cross-sections of the March Current Population Survey (CPS). Using only the recent ACS waves would impose the potentially strong assumption that both the shape and level of the age-earnings profiles are constant across all cohorts. To be of the greatest policy relevance, I wish to produce projections based only younger cohorts of students. If I use only data on recent college graduates, then forecasts of their earnings later in life would rely exclusively on ad hoc assumptions. My solution to this problem is to use older cohorts, for whom I observe the majority of their lifecycle earnings across the several decades covered in my data, to estimate the shape of the age-earnings profile. I then assume that the most recent cohort of individuals (up to age 35) will follow a parallel path. In other words, the level of my projections are determined by the most recent cohort, but the shape (e.g. peak, slope) is dictated by older cohorts. The results using this correction are qualitatively the same as using only the recent ACS waves, likely because the college wage premium has remained constant in recent decades. I present the ACS-based results as my preferred specificationdue to the smaller sample sizes (and thus less precise simulations) offered by the CPS.

$$\hat{y}_{it} = \bar{y}_j^{HS} + \check{\delta}_j^{educ} + \check{\beta}_{age}^{(j)}(t - a\bar{g}e_j) + f^{-1}(\sigma^2_{educ,j}) \; \forall \, t \in j \tag{8}$$

Equation (8) describes the simulated log earnings for individual i at each age t. The first term on the right hand side is the average log earnings of high school graduates with no college enrollment in age group j. The second term, $\check{\delta}_j^{educ}$, represents the estimated premium from the ACS for each educational category other than high school graduates in age group j. The next term adjusts the simulated earnings for age differences within each age group j. This accounts for the fact that there are returns (positive or negative) to age/experience within small age groupings without the loss of precision associated with estimating age effects for every age. The final term, the inverse normal CDF for each educational category and age grouping, generates dispersion in the simulated log earnings distribution based on the observed residual variation from the ACS samples.[14] Table 1 presents basic descriptive statistics from the pooled ACS samples, and Figure 2 presents the distributions of lifetime earnings for both college and high school graduates produced from the above model.

Individuals are assumed to spend five years in school, graduating at age 23. Despite roughly a third of current undergraduates being over age 25, data on outcomes for these so-called "non-traditional" students is relatively sparse. I thus focus my policy simulations on traditional college students for two reasons. First, this avoids having to make myriad ad hoc assumptions about how age/experience interact with the college wage premium. Second, this assumption produces conservative (downward-biased) estimates of optimal repayment rate thresholds due to individuals receiving the college premium over the longest possible time horizon. A similarly downward-biased estimate is achieved by assuming that all students graduate.

[14] In order to control the level of persistence in earnings across the life cycle, two uniform random variables are generated for use in the CDF for each individual. One is "permanent", following a given individual throughout their lifecycle, the other is transitory, and is redrawn at each new age group. A weighted average (70% permanent and 30% transitory) of the two terms is used to determine where in the earnings CDF the individual resides. These parameters were chosen to roughly match patterns observed in the National Longitudinal Survey of Youth 1979.

Students are assumed to pay $10,000 per year in out-of-pocket costs to attend college, and to finance the balance of their education with student loan debt. This equates roughly to the situation faced by the typical student attending an in-state public school.[15] Each simulated student in the analysis below accrued $30,000 in loans at the time they leave school, the current national average, and will repay their loans according to the most recent Income Driven Repayment (IDR) policies: 10 percent of discretionary income (the difference between annual income and the federal poverty line for an individual person). Any remaining student loan balance after twenty years is written off by the government. All results below are presented based on this twenty year time horizon of repayment. I model student loan repayment under IDR (rather than the standard/fixed repayment plan) for several reasons. First, income-based plans are growing in popularity and disproportionately utilized at schools with poor loan outcomes. Second, it bases the resulting estimates solely on student labor market outcomes, irrespective of factors such as family wealth. Finally, paying based on income mechanically makes it more likely that you will fail to repay any principal balance. Basing estimated optimal thresholds off of this assumption thus yield a lower bound, as a school with any given repayment rate actually produces worse "true" outcomes if students are paying down their debt with preexisting wealth rather than returns from their education.

Interest accrues at the 2018-19 academic year federal student loan interest rate of 5.05%, and all incomes increase over time at the current long-run projected inflation rate of 2.2%.[16] Taxes and after-tax incomes are calculated based on the tax laws in effect during 2019, and are based on single-filer tax rates.

We can now track the success of the student loan portfolio, or the relative welfare of individuals with/without a college degree, and benchmark these figures against any other costs (e.g. administrative) or benefits. Finally, we can scale the college premium up or down to achieve different repayment rates, producing an earnings distribution which mimics

[15] https://trends.collegeboard.org/college-pricing/figures-tables/average-net-price-over-time-full-time-students-public-four-year-institution

[16] https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/historical-data/inflation-forecasts

a school with any given quality of outcomes. For example, a simulated population with the same college premium as in the real world is predicted to have a three-year repayment rate of 41.6%, which is quite close to the average across all U.S. postsecondary institutions of 45.4%. It should be noted that while I view it as a positive that these two figures are relatively close, they are by no means directly comparable. The repayment rate generated from the policy simulation assumes that all students repay via income driven repayment, while only a third of actual student loan borrowers use this repayment mechanism. Furthermore, the college premium is estimated from all individuals with a college degree, but the real repayment rate is definitionally only calculated from individuals who have taken out student loans (on average about 70% of students).

5 Results

It is critical to note that the estimates produced in the following tables represent lower bounds (the true break-even points are likely higher) due to the modeling assumptions described above. The focus on obtaining a lower bound is intentional, given the severe nature of the proposed penalty for institutions if they fall below the threshold. In other words, my goal is to estimate thresholds, below which students and taxpayers are measurably worse off, with as much certainty as is possible.

Given the prominent role that projected forgiveness under IDR plans has played in revised estimates of the cost of the student loan program from the Congressional Budget Office,[17] and the politics surrounding loan forgiveness, a first-order concern of policymakers is likely the magnitude of forgiveness that can be expected for schools with different repayment rates. Figures 3 plots the expected percentage of borrowers who receive forgiveness (based on the policy simulation) against the institutional repayment rate. For an institution with a repayment rate of 15%, at least 50% of borrowers are expected to receive some loan

[17] https://about.bgov.com/news/student-loan-outlook-is-reversed-showing-31-billion-cost/

forgiveness after twenty years being enrolled in an IDR program. This figure drops to 18% of borrowers at an institution with a repayment rate of 30%. It should be noted that these projections assume a world in which there is no adverse selection into IDR programs because the entire population is enrolled.

Tables 2 presents the estimated optimal thresholds (without accounting for time-discounting/opportunit costs) from the perspective of the government's student loan portfolio. An institution with a repayment rate of 14.3% can be expected to generate a net-zero budgetary impact in nominal dollars, and all institutions with repayment rates below this level are predicted to have net-negative impacts.[18] The second and third columns show the number of schools (69) and students enrolled at those schools (26,936) which fall below this repayment rate. The final three columns further break down the students by sector. Data on the repayment rates and enrollments at each school, used to illustrate the impact of tying Title IV eligibility to a repayment rate threshold, are drawn from the most recent wave of the College Scorecard.

The second, third, and fourth rows of Table 2 individually add several adjustments to the most basic simulation: administrative costs, taxes, and "selection" (the notion that increased earnings among college graduates may not be entirely due to attending college in a causal sense). Administrative costs have been estimated to be roughly 1.7% of loan volume.[19] Taking these costs into account has only a small impact on the breakeven level from the government's perspective.

Tax revenue has the potential to drastically alter the calculus of the government's student loan investment, assuming that the college degree increased earnings in a causal sense. Under the set of assumptions in Table 2, virtually every school produces a net-benefit to government

[18] I choose not to present standard errors because I believe their inclusion would convey a false sense of precision in the results. As in any policy simulation of this type, I could artificially obtain standard erros arbitrarily small simply by simulating a larger population. For example, with roughly 65,000 simulated lifecycles, the standard errors on the typical threshold range from 0.1% to 0.5% when the simulation procedure is bootstrapped. The principal source of uncertainty that readers should be worried about, particularly in Tables 2 and 3, is the assumptions used to generate the thresholds, rather than any sort of sampling variation/prediction error. I would be happy to provide results under any set of alternative assumptions that is of interest to readers.

[19] https://www.newamerica.org/education-policy/edcentral/federal-student-loans-administrative-costs/

revenues when taxes are taken into account.

On it's own, non-random selection does not logically have an impact on the government's balance sheet. If a loan is being repaid, it is of no difference whether an individual's income is higher because they attended college or due to their innate ability (this is why the thresholds in row 4 of Tables 2 and 3 are unchanged). When taking tax revenue into account, however, selection is an important factor to consider because now a counterfactual has been introduced. In the results labeled with "selection" I assume that one-third of the college wage premium is due to selection, and adjust the counterfactual scenario accordingly. This is roughly the magnitude of selection estimated in prior work (Webber, 2014), and I think is a fair representation of estimates from the returns to college literature. The final row incorporates all of the above adjustments at the same time, yielding an optimal threshold of 11.6%

While the results based on nominal dollars are a useful starting point to establish a baseline, any policy must take into account the discounted nature of future payments. This is particularly important given the prevalence of Income Driven Repayment plans and the fact that the wage gains from education accrue throughout the lifecycle (rather than all up front). At the time the government receives the bulk of student loan repayments under IDR, the real value of those payments has been diluted by years of inflation. Moreover, every dollar that the government lends today is one that could be used on some other form of spending, even if it is simply paying down the federal debt.

There is some debate over how significantly to discount future repayments for the purposes of scoring the federal student loan portfolio. This paper discounts future repayments by the interest rate on 30-year Treasury Bonds. This is intuitive, as this rate supposedly reflects the longterm costs of borrowing that the government faces. In practice, this approach simply looks at repayment in real dollars, as the 30-year Treasury Bond interest rate and long-run inflation expectations are naturally quite close.

A compelling argument can also be made for the "Fair Value Accounting" methodology, which takes into account market risk/return when calculating a discount factor. For example,

there is uncertainty about how much students will repay, and there are other investments the government could make which may have a greater return. As Fair Value Accounting necessarily discounts future repayments to a much more significant degree, my method of discounting will again lead to a lower bound on the estimated thresholds. See Delisle and Richwine (2014) for a detailed discussion of Fair Value Accounting in a student loan context.

The large impact that discounting has on whether a given school contributes a net loss to the federal student loan portfolio can easily be seen in Table 3. The baseline optimal threshold, which compares dollars repaid by students to dollars lent by the government, is 43.9%. This implies that nearly half of institutions (though comprising only about a third of students), and effectively the entire for-profit sector, act as a net drain on the federal budget.

Adding in the same corrections as in Table 2 produces similar qualitative changes to the estimated optimal threshold. In the specification with all corrections, the breakeven point is estimated to be 16.9%. Table 3 in particular underscores the importance that taxes play in the government's financial interest in the student loan portfolio. Without taking taxes into account, a third or more students attend an institution which produces a net cost to taxpayers. When accounting for taxes (and assuming a third of the college premium is due to selection/signaling), this figure drops to less than half of a percentage point. Note that these figures only represent the cost to taxpayers of the student loan program, and do not include other student-centered investments such as Pell Grants or institutional subsidies, which produces a further downward bias on the optimal thresholds that I estimate.

The protection of taxpayers' investment is certainly an important consideration for policymakers, but it ignores much of the rationale often used to justify a significant government presence in the market for higher education. First and foremost, it ignores the importance of a college degree to upward economic mobility, particularly among credit-constrained individuals. In other words, it could certainly be the case that, from a social welfare perspective, even if an institution's students do not fully repay their student loans it is still in the public's best interest to incur this loss. The next two tables thus examine the college investment

decision purely from the point of view of the individual borrower, and whether their explicit and implicit (opportunity cost of time) investment is likely to pay off.

Table 4 presents optimal thresholds under a variety of specifications, including accounting for taxes, selection, and an increased emphasis on lower income individuals. The thresholds in this table should be interpreted as the point at which college graduates are equally well-off when compared to the average high school graduate. At schools with a higher repayment rate, the average college grad will be worse off, and at schools with a lower repayment rate college graduates are predicted to be worse off. The same assumptions from the above analyses, namely those which produce "lower bound" estimates, are maintained here. Note that in addition to my attempt to make assumptions which estimate a lower bound on the optimal threshold, this comparison to the average high school graduate is itself a fairly low bar to judge the success of college investment.

In this exercise, correcting for selection has an independent impact on the estimated thresholds because counterfactual earnings matter to the individual, whereas from the government's financial point of view counterfactual earnings only mattered once taxes enter the equation. Results are presented both with and without the inclusion of taxes, which in this case increase the breakeven threshold due to our progressive system. The second set of specifications use the natural log transformation when aggregating outcomes to put a greater weight on improvements among low-income workers.

Without accounting for discounting, estimated thresholds tend to lie in the 12%-16% range, impacting a relatively small number of schools. The student composition below this level tends to be mostly private (both non-profit and for-profit). After accounting for present value discounting, shown in Table 5, the estimated threshold increase somewhat to 17.2% (all corrections, log model) and 20.1% (all corrections, linear model). Given the low bar (as well off as the average high school graduate) that these thresholds are designed to approximate, it should be clear that there are a significant number of students attending schools which produce poor outcomes. It follows that any change in accountability policy which seeks

to impose a standard such as "the average college graduate is financially better off than the average high school graduate" would represent a significant disturbance to the higher education marketplace. Figures 4a and 4b plot the repayment rate (x-axis) against the cumulative number of schools and students below that repayment rate. The clear change in slope in the vicinity of the optimal thresholds described above (15%-20%) underscores the tough decisions that policymakers face.

6 Conclusion

This paper seeks to provide information to policymakers on a fundamental question in higher education policy: which institutions/programs should have access to the federal student loan program. There are two two primary considerations to this question. First, which schools leave students worse off for having made the monetary and time investment of attending college? Second, which schools have a loan portfolio which leave taxpayers worse off. Most policymakers put more weight on one or the other of these factors, and for this reason I try to provide empirical evidence on both in the analyses above.

I use population earnings data to construct estimates of lifecycle earnings and loan repayment for the distribution of both college and high school graduates. In then conduct a number of policy simulation in which I estimate the point at which a given school's students are 1) able to repay enough of their loans so that taxpayers break even, and 2) exactly as well off as the average high school student. I present results under a variety of different assumptions including the degree to which future repayment/earnings are discounted, taxes, and the causal impact of college on earnings. The quality of an institution's student loan portfolio is measured by the three-year repayment rate, a widely-reported metric used to assess the loan performance of student borrowers which has been proposed as a key part of most proposed reforms to the higher education accountability system. The three-year repayment rate measures the proportion of borrowers who have made at least $1 of progress

paying down the principal balance of their student loans three years after leaving school.

I find that regardless of whether a policymaker's intention is to minimize taxpayer losses or to maximize students' future income, the optimal repayment rate threshold lies in the range of 15%-20%. There are two significant caveats to these findings. First, as in any policy simulation, a number of assumptions must be made to reduce the complexities of the world down to a computer program. In all cases I tried to make assumptions which would lead to the most conservative threshold (e.g. a lower bound on the true breakeven points for taxpayers and students). Given the paramount importance of a college education in today's labor market, I would rather err on the side of providing too much access rather than too little.

Second, while the results for student financial well-being are relatively insensitive to modeling choices, this is not true for taxpayer welfare. Taxpayers are predicted to realize a net loss from dramatically more schools under even modestly less charitable assumptions about the causal impact of college on earnings or the magnitude of present-value discounting. This will likely not surprise those in the higher education policy community, who are well-aware that the undergraduate student loan program operates at a net loss to the taxpayer.

In light of the above evidence, I feel it is safe to conclude that both students and taxpayers are made measurably worse off by a small, but nontrivial, number of institutions. Policymakers face a tough set of choices, as there are only 75 schools (enrolling 30,000 students) with a repayment rate below 15%. This figure jumps to 300 institutions and 200,000 students below a repayment rate of 20%. Fitting my findings into the historical format that higher education accountability has taken, one possible rule could be that schools lose access to Title IV funding if they have a single three-year repayment rate below 15%, or three consecutive three-year repayment rates below a slightly higher threshold such as 25%.

One promising way to reform the current system of accountability is to use program-level metrics to determine access to Title IV funding rather than institution-level measures. Such a change would lead to better targeted sanctions, but moving enforcement to the program

level is no panacea (Kelchen, 2018b). As of the writing of this paper, the public has no information on the distribution of repayment rates at the program level, and hence it is difficult to predict how many programs and students would be impacted. Moreover, there are many administrative difficulties that exist for a program level accountability system such as how to categorize students who drop out or potential gaming by shifting students to different programs. As more data become available on student outcomes at the program level (short-term outcomes are expected soon), future work should focus on whether the potential benefits of targeted sanctions outweigh the tradeoffs inherent in structuring an accountability regime around the program rather than institutional level.

References

F. R. Addo, J. N. Houle, and D. Simon, "Young, black, and (still) in the red: Parental wealth, race, and student loan debt," *Race and Social Problems*, vol. 8, no. 1, pp. 64–76, 2016.

R. J. Andrews, J. Li, and M. F. Lovenheim, "Quantile treatment effects of college quality on earnings," *Journal of Human Resources*, vol. 51, no. 1, pp. 200–238, 2016.

C. Avery and S. Turner, "Student loans: Do college students borrow too much–or not enough?" *Journal of Economic Perspectives*, vol. 26, no. 1, pp. 165–92, 2012.

N. Barr, B. Chapman, L. Dearden, and S. Dynarski, "The us college loans system: Lessons from australia and england," *Economics of Education Review*, 2018.

S. Baum and S. Schwartz, "Unaffordable loans," 2018.

M. Bhuller, M. Mogstad, and K. G. Salvanes, "Life-cycle earnings, education premiums, and internal rates of return," *Journal of Labor Economics*, vol. 35, no. 4, pp. 993–1030, 2017.

S. Canaan and P. Mouganie, "Returns to education quality for low-skilled students: Evidence from a discontinuity," *Journal of Labor Economics*, vol. 36, no. 2, pp. 395–436, 2018.

S. R. Cellini, R. Darolia, and L. J. Turner, "Where do students go when for-profit colleges lose federal aid?" *American Economic Journal: Economic Policy*, Forthcoming.

S. R. Cellini, "For-profit higher education: An assessment of costs and benefits," *National Tax Journal*, vol. 65, no. 1, p. 153, 2012.

S. R. Cellini and L. Chaudhary, "The labor market returns to a for-profit college education," *Economics of Education Review*, vol. 43, pp. 125–140, 2014.

S. R. Cellini and N. Turner, "Gainfully employed? assessing the employment and earnings of for-profit college students using administrative data," *Journal of Human Resources*, vol. 54, no. 2, pp. 342–370, 2019.

S. R. Cohodes and J. S. Goodman, "Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy," *American Economic Journal: Applied Economics*, vol. 6, no. 4, pp. 251–85, 2014.

S. B. Dale and A. B. Krueger, "Estimating the effects of college characteristics over the career using administrative earnings data," *Journal of human resources*, vol. 49, no. 2, pp. 323–358, 2014.

R. Darolia, "Integrity versus access? the effect of federal financial aid availability on post-secondary enrollment," *Journal of Public Economics*, vol. 106, pp. 101–114, 2013.

J. Delisle and J. Richwine, "The case for fair-value accounting," *National Affairs*, vol. 21, pp. 95–111, 2014.

R. G. Ehrenberg, "American higher education in transition," *Journal of Economic Perspectives*, vol. 26, no. 1, pp. 193–216, 2012.

S. Goldrick-Rab, *Paying the price: College costs, financial aid, and the betrayal of the American dream.* University of Chicago Press, 2016.

N. W. Hillman, D. A. Tandberg, and J. P. Gross, "Performance funding in higher education: Do financial incentives impact college completions?" *The Journal of Higher Education*, vol. 85, no. 6, pp. 826–857, 2014.

M. Hoekstra, "The effect of attending the flagship state university on earnings: A discontinuity-based approach," *The Review of Economics and Statistics*, vol. 91, no. 4, pp. 717–724, 2009.

C. M. Hoxby, "The productivity of us postsecondary institutions," in *Productivity in Higher Education.* University of Chicago Press, 2018.

R. Kelchen, *Higher education accountability.* JHU Press, 2018.

——, "Exploring key questions: Program-level repayment for higher education accountability," Higher Learning Advocates, Tech. Rep., 2018.

R. Kelchen and L. J. Stedrak, "Does performance-based funding affect colleges' financial priorities?" *Journal of Education Finance*, vol. 41, no. 3, pp. 302–321, 2016.

B. M. Marx and L. J. Turner, "Student loan nudges: Experimental evidence on borrowing and educational attainment," *American Economic Journal: Economic Policy*, vol. 11, no. 2, pp. 108–41, 2019.

B. Ost, W. Pan, and D. Webber, "The returns to college persistence for marginal students: Regression discontinuity evidence from university dismissal policies," *Journal of Labor Economics*, vol. 36, no. 3, pp. 779–805, 2018.

J. Scott-Clayton, "What accounts for gaps in student loan default, and what happens after. evidence speaks reports, vol 2,# 57." *Center on Children and Families at Brookings*, 2017.

J. C. Shin, "Impacts of performance-based accountability on institutional performance in the us," *Higher Education*, vol. 60, no. 1, pp. 47–68, 2010.

D. A. Webber, "The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors," *Labour economics*, vol. 28, pp. 14–23, 2014.

——, "Are college costs worth it? how ability, major, and debt affect the returns to schooling," *Economics of Education Review*, vol. 53, pp. 296–310, 2016.

——, "Risk-sharing and student loan policy: Consequences for students and institutions," *Economics of Education Review*, vol. 57, pp. 1–9, 2017.

S. D. Zimmerman, "The returns to college admission for academically marginal students," *Journal of Labor Economics*, vol. 32, no. 4, pp. 711–754, 2014.

Figure 1a: Repayment Rate Distribution

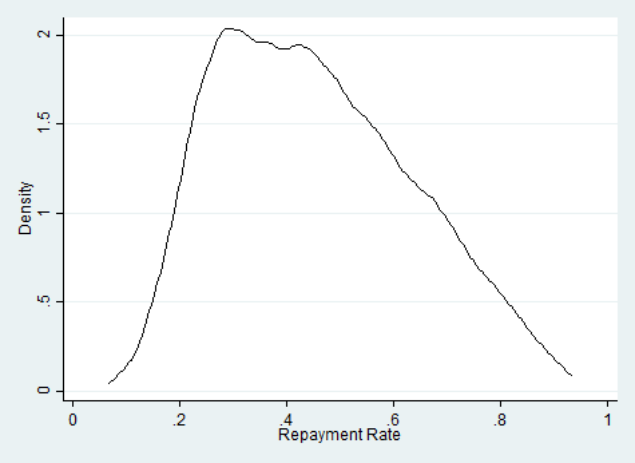


Figure 1b: Repayment Rate Distribution: Broken Down by Sector

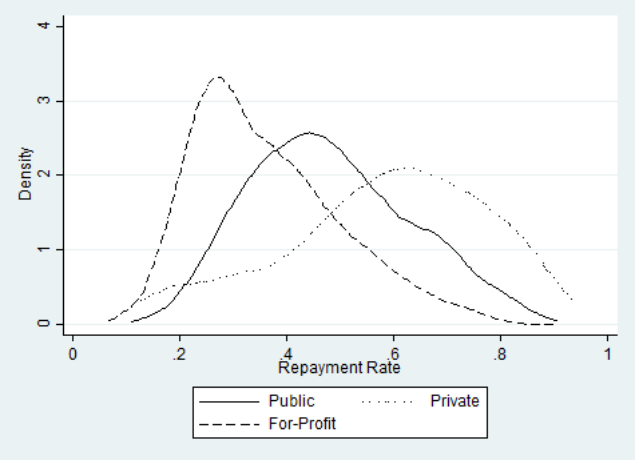


Figure 2b: Earnings Distributions

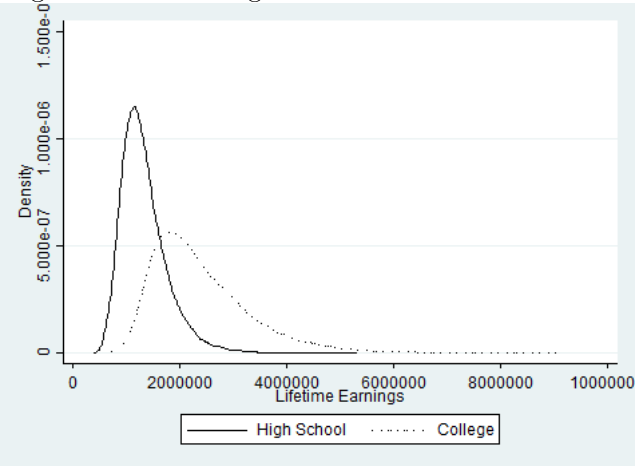


Figure 3: Projected Proportion of Borrowers Who Receive Loan Forgiveness

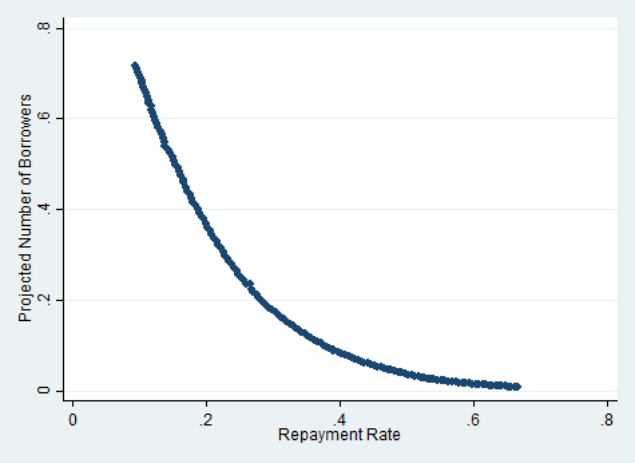


Figure 4a: Repayment Rates and a Running Total of Schools

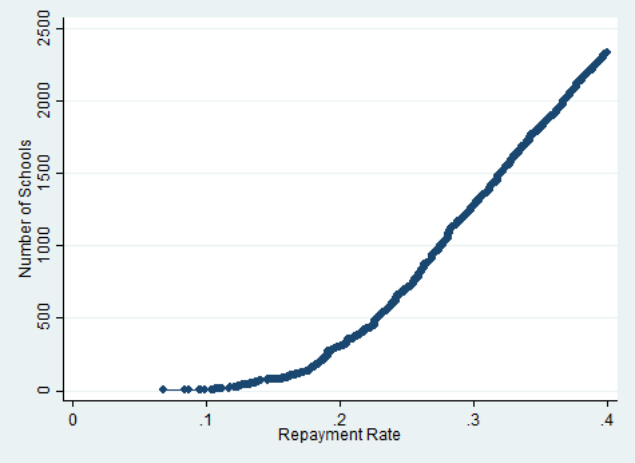


Figure 4b: Repayment Rates and a Running Total of Students

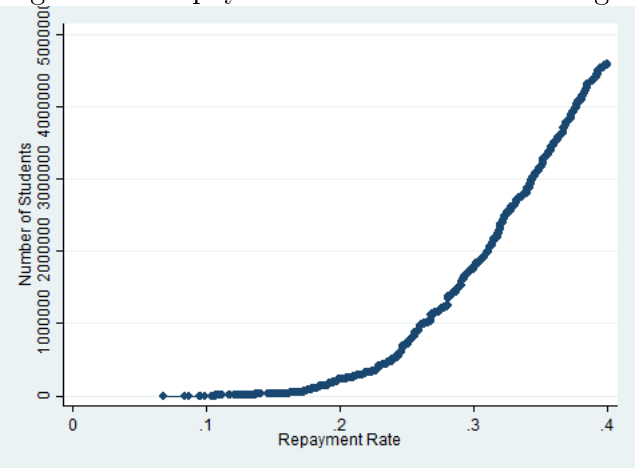


Table 1: Summary Statistics

	High School	College
Age	42.85	42.6
	(13.38)	(11.96)
Female	.43	.52
	(.5)	(.5)
Black	.12	.06
	(.32)	(.24)
Hispanic	.15	.07
	(.36)	(.26)
Earnings	36006.16	72933.59
	(35108)	(74124.26)
Social Science	0	.1
	(0)	(.31)
Arts-Humanities	0	.13
	(0)	(.34)
Business	0	.26
	(0)	(.44)
STEM	0	.26
	(0)	(.44)
Education	0	.09
	(0)	(.29)
Observations	1750411	1418885

American Community Survey waves 2013-2017

Table 2: Optimal Repayment Rate Threshold - Government Loan Portfolio, No Discounting

				Student breakdown		
	Total=	5416	15137256	11261030	2718018	1158208
	Threshold	Schools	Students	Public	Private	For-Profit
Baseline	.143	69	26936	4866	12888	9182
w/admin costs	.151	78	30408	4866	15187	10355
w/taxes	.109	13	2788	.	1245	1543
w/selection	.143	69	26936	4866	12888	9182
w/all	.114	16	3456	496	1245	1715

Each row presents estimates of the optimal repayment rate threshold (from the perspective of the government's loan portfolio), and the number of schools/students below this threshold, based on different assumptions regarding the costs and benefits of higher education. The baseline row considers only the nominal dollars repaid to the government. The row w/admin costs includes the estimated 1.7% cost of administering the student loan program. The w/taxes row assumes that all of the college premium is causal, and includes the increased tax revenue as a benefit to the government. The w/selection row assume that one-third of the college wage premium is due to non-random selection. The final row combines all of these adjustments.

Table 3: Optimal Repayment Rate Threshold - Government Loan Portfolio, with Discounting

				Student breakdown		
	Total=	5416	15137256	11261030	2718018	1158208
	Threshold	Schools	Students	Public	Private	For-Profit
Baseline	.436	2715	5633622	4280646	390567	962409
w/admin costs	.487	3253	7009345	5412307	573579	1023459
w/taxes	.148	76	28396	4866	13367	10163
w/selection	.436	2715	5633622	4280646	390567	962409
w/all	.17	120	51998	11881	21435	18682

Each row presents estimates of the optimal repayment rate threshold (from the perspective of the government's loan portfolio), and the number of schools/students below this threshold, based on different assumptions regarding the costs and benefits of higher education. The baseline row considers only the dollars (discounted according to current 30-yr Treasury Bond rate) repaid to the government. The row w/admin costs includes the estimated 1.7% cost of administering the student loan program. The w/taxes row assumes that all of the college premium is causal, and includes the increased tax revenue as a benefit to the government. The w/selection row assume that one-third of the college wage premium is due to non-random selection. The final row combines all of these adjustments.

Table 4: Optimal Repayment Rate Threshold - Individual, No Discounting

				Student breakdown		
	Total=	5416	15137256	11261030	2718018	1158208
	Threshold	Schools	Students	Public	Private	For-Profit
Linear Baseline	.128	41	13683	3373	7826	2484
w/taxes	.133	48	17180	3373	10007	3800
w/selection	.151	78	30408	4866	15187	10355
w/all	.161	99	38676	8712	17448	12516
Log Baseline	.139	65	24713	3373	12888	8452
w/taxes	.15	78	30408	4866	15187	10355
w/selection	.143	69	26936	4866	12888	9182
w/all	.145	69	26936	4866	12888	9182

Each row presents estimates of the optimal repayment rate threshold (from the maximizing overall individual welfare), and the number of schools/students below this threshold, based on different assumptions regarding the costs and benefits of higher education. The baseline row considers only the income differences (after paying college/loan costs) between college graduates and high school graduates. The w/taxes row presents results based on after-tax income. The w/selection row assume that one-third of the college wage premium is due to non-random selection. The final row combines all of these adjustments. The first four estimates use a linear utility function, thus treating every additional dollar in income equally. The second four estimates use a natural log utility function, giving increased priority to increases among low-income workers.

Table 5: Optimal Repayment Rate Threshold - Individual, with Discounting

				Student breakdown		
	Total=	5416	15137256	11261030	2718018	1158208
	Threshold	Schools	Students	Public	Private	For-Profit
Linear Baseline	.151	78	30408	4866	15187	10355
w/taxes	.162	100	38790	8712	17448	12630
w/selection	.186	205	132476	59671	30042	42763
w/all	.199	303	224089	117960	32868	73261
Log Baseline	.166	109	42165	8712	20490	12963
w/taxes	.168	116	49064	9223	21435	18406
w/selection	.163	103	39442	8712	17962	12768
w/all	.172	126	54554	11881	23146	19527

Each row presents estimates of the optimal repayment rate threshold (from the maximizing overall individual welfare), and the number of schools/students below this threshold, based on different assumptions regarding the costs and benefits of higher education. The baseline row considers only the income differences (after paying college/loan costs, discounted according to current 30-yr Treasury Bond rate) between college graduates and high school graduates. The w/taxes row presents results based on after-tax income. The w/selection row assume that one-third of the college wage premium is due to non-random selection. The final row combines all of these adjustments. The first four estimates use a linear utility function, thus treating every additional dollar in income equally. The second four estimates use a natural log utility function, giving increased priority to increases among low-income workers.