The Effectiveness of the Student Loan Safety Net: An Evaluation of Income-Driven Loan Repayment

by Audrey Peek

B.A. in French Language and Literature, Western Society and Culture, May 2010, Concordia University
M.P.A. in Public Administration, June 2013, University of Washington—Seattle

A Dissertation submitted to

The Faculty of
The Graduate School of Education and Human Development
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Education

May 20, 2018

Dissertation directed by

Daniel J. Klasik
Assistant Professor of Higher Education Administration

The Graduate School of Education and Human Development of The George Washington University certifies that Audrey Peek has passed the Final Examination for the degree of Doctor of Education as of March 13, 2018. This is the final and approved form of the dissertation.

The Effectiveness of the Student Loan Safety Net:
An Evaluation of Income-Driven Loan Repayment

Audrey Peek

Dissertation Research Committee:

Daniel J. Klasik, Assistant Professor of Higher Education Administration, Dissertation Director

Lindsey E. Malcom-Piqueux, Associate Director for Research and Policy at the Center for Urban Education, University of Southern California, Committee Member

Sandy Baum, Senior Fellow, Urban Institute, Committee Member

Acknowledgements

The author wishes to thank her family, friends, mentors, writing buddies, coworkers, bosses, professors, cohort members, and the patient coffee shop owners of DC. She also wishes to thank her committee for their generous guidance and support.

Abstract

The Effectiveness of the Student Loan Safety Net:
An Evaluation of Income-Driven Loan Repayment

One in five federal student loan borrowers today is enrolled in income-driven loan repayment (IDR), a set of safety net programs in which loan payment amounts are tied to borrowers' incomes. Policymakers across the political spectrum support expanding IDR as a way of reducing loan default and encouraging borrowers to work in public service careers, though there is little evidence of the effects of IDR on borrowers' outcomes. Through the lenses of human capital theory and risk aversion theory, this study investigates two key gaps in our knowledge about IDR in comparison to other repayment plans: whether IDR borrowers' make different career choices and whether IDR borrowers are more successful at paying their loans. Using the National Center for Education Statistics' Baccalaureate and Beyond dataset (B&B:08/12), I weighted borrowers by their propensity to enroll in IDR, based on family, institutional, communities, and political and economic characteristics. This analysis found that IDR borrowers are statistically significantly more likely to be women, of Hispanic origin, and from low-income households. IDR borrowers are no more likely to pursue public service careers four years after graduation, but they are substantially more likely to be in repayment as opposed to being in default, forbearance, or any other loan status. The total estimated costs of loan repayment are more varied for IDR borrowers than for borrowers in other plans. Borrowers of color in IDR are likely to pay more than White borrowers in IDR. I conclude with a discussion of the implications for policy, research, and practice, including how policy makers and researchers should interpret the tradeoffs in these results.

Table of Contents

Acknowledgements iii

Abstract iv

List of Figures viii

List of Tables ix

Chapter One 1

Context 5

IDR options for federal student loans. 5

Borrower experiences in Income-Driven Repayment 11

Enrollment in Public Service Loan Forgiveness 12

Growing enrollment in IDR 13

International IDR programs. 16

Definitions 18

Problem Statement 21

IDR participants unknown. 21

Unclear relationship with PSLF. 22

Default assumption untested. 23

Possible differences in total payment amounts 23

Purpose Statement and Research Questions 24

Statement of Potential Significance 25

Limitations and Assumptions 26

Chapter Two 27

Literature Review 28

IDR selection. 29

IDR and loan repayment success. 31

Total repayment amounts. 34
IDR and career choice. 36
IDR and other post-graduation outcomes. 39
Discussion of empirical studies. 40
Conceptual Framework 43
IDR selection. 43
Loan repayment status. 47
Career. 47
Chapter Three 50
Research Questions 50
Research Approach 52
Causal estimation. 52
Propensity score analysis. 55
Data 57
Data preparation. 58
Statistical Models 59
Methods to Address Bias 61
Analysis Steps 65
Step 1: Choice of covariates for propensity score model. 66
Step 2: Balance of propensity scores across treatment and control groups. 71
Step 3: Balance of covariates across treatment and control groups. 74
Step 4: Choice of propensity score analysis strategy. 75
Step 5: Balance of covariates across treatment and control groups after matching. 77
Step 6: Estimation of treatment effects. 79
Chapter Four 82

RQ1. What are the characteristics of borrowers who enroll in IDR? 82

RQ2. Are IDR borrowers more likely than borrowers in standard loan repayment to work in public service organizations? 97

RQ3. Are IDR borrowers less likely than borrowers in standard loan repayment to be in repayment? 101

RQ4. How do the present values of the estimated total amounts repaid differ between IDR borrowers and borrowers in other repayment plans? 106

Summary of Results 112

Chapter Five 114

Implications 114

Contributions to the Literature 120

Contributions to Theory 121

Limitations 122

Implications for Future Research 124

Implications for Policy 125

Expanding IDR to more borrowers. 125

Eliminating Public Service Loan Forgiveness. 127

References 128

List of Figures

Figure 1. Growth in IDR borrowers and total debt from 2013 to 2017 14

Figure 2. Typology of income-related financing policies 18

Figure 3. Conceptual model 49

Figure 4. Distribution of propensity scores by treatment group 69

Figure 5. QQ plot of propensity scores by treatment status 70

Figure 6. Ratio of standardized mean differences to variance before and after adding stabilized inverse propensity weights 75

Figure 7. Standardized mean differences before and after adding stabilized inverse propensity weights 76

List of Tables

Table 1. Comparison of IDR options for federal student loans 7

Table 2. Summary statistics for propensity score model covariates.....70

Table 3. Cumulative distribution of propensity scores by treatment status 71

Table 4. Demographic characteristics of survey sample 81

Table 5. Academic characteristics of survey sample 85

Table 6. Institutional characteristics of survey sample 87

Table 7. Financial aid characteristics of survey sample 89

Table 8. Post-graduation characteristics of survey sample 91

Table 9. Mantel-Haenzel sensitivity analysis for public service outcome model 96

Table 10. Public service career outcome models 97

Table 11. Mantel-Haenzel sensitivity analysis for repayment model 100

Table 12. Student loan repayment outcome models 102

Table 13. Net present value of estimated total payments by treatment group 105

Table 14. Net present value of estimated total payments by gender, IDR borrowers only 106

Table 15. Net present value of estimated total payments by ethnicity, IDR borrowers only 107

Table 16. Net present value of estimated total payments by race and ethnicity, IDR borrowers only 108

Chapter One

A rapidly growing share of federal student loan borrowers is enrolled in Income Driven Loan Repayment (IDR) plans, which allows borrowers' loan payments to shift as their income level changes. When I began this study in early 2014, 17.10 million borrowers were enrolled in an IDR program, or 11% of all borrowers in the Direct Loan portfolio. By late 2017, that number had grown to 21.33 million, or 22% of all borrowers. These borrowers currently hold $679.4 billion in debt, or 37.9% of the Direct Loan portfolio (U.S. Department of Education, 2017a).

IDR is one solution to rising college prices and increases in student debt, sometimes called the student debt crisis. Although IDR does not address tuition prices or directly reduce debt amounts, IDR may help alleviate borrowers' anxiety about debt and help borrowers avoid loan default (Chapman, 1997; Hillman & Gross, 2014). Evidence regarding IDR's effectiveness is largely absent from policy debates about expanding IDR to a larger number of student loan borrowers, due to the paucity of empirical research on the topic. In this study, I investigated two key gaps in the evidence about IDR: IDR borrowers' loan repayment success and career choices.

Many policy makers from across the political spectrum support IDR expansion, including former Senator Tom Harkin (D-IA), former Representative Tom Petri (R-WI), former President Obama, and President Trump. Student advocates have also encouraged expanding IDR by automatically enrolling all new student loan borrowers (New America, Young Invincibles, & NASFAA, 2014). IDR is equally popular with the public, as evidenced by the frequency of articles in the non-trade press touting IDR's benefits, with headlines like "There's a way to dramatically lower student debt payments, but hardly

anyone uses it" (Douglas-Gabriel, 2014), "7 Ways to stop student loans from ruining your life" (Halpert, 2014), "7 ways to make paying down your student loans a little easier" (Kobliner, 2017), and "How this millennial paid off over $100,000 of debt in 6 years" (Hess, 2017).

There are a number of possible reasons for IDR's popularity. Some have argued that IDR encourages equitable access by making prices seem more reasonable, as well as encouraging post-college success by lowering the risk of defaulting on loan payments. An analysis of the legislative texts that created IDR programs in the United States found that, among multiple goals, the programs aimed to support debt management and on-time payments for borrowers (Hillman & Gross, 2014). If IDR increases borrowers' abilities to make on-time payments, then the rate of delinquency and default should drop (Chapman, 1997). Increasing affordability, particularly for low-income students, is a primary goal of Australia's national IDR system as well (Chapman, 1997). In sum, IDR is assumed to improve students' upfront perceptions of their abilities to pay for college as well as their post-college loan repayment behaviors.

The organizations and researchers who have spoken out against expanding IDR argue that IDR programs do not distribute public subsidies for postsecondary education equitably or efficiently. On the one hand, IDR may provide an unnecessary windfall to borrowers who attend professional graduate programs and incur large amounts of debt (Delisle & Holt, 2012). These borrowers can expect large returns on their investment and thus do not need IDR to mitigate the risk of low future earnings. On the other hand, some have cautioned that IDR can lead borrowers to pay more than they would in other plans due to the fact that interest accrues over a longer period (Asher, Cheng, & Thompson,

2014). In particular, women and people of color may take longer to repay their loans under IDR due to their relatively lower average wages (Canadian Federation of Students, 2012). These arguments assume that IDR should make student loan payments affordable in multiple ways: payments should be low in the short term if borrowers cannot afford the standard plan, and payments should be equal to or less than the standard plan in the long term if borrowers take more time to pay off their loans. In this scenario, IDR makes loans affordable in terms of cash on hand and in terms of expected cash flows over time.

Do policymakers agree that IDR should be affordable in the short and long term? Given the wide variety of IDR policies in the US and around the world, it is not clear that policymakers all have the same view. For example, IDR programs in the US have different ways of subsidizing and capitalizing accrued interest, which affects borrowers' overall payment amounts. IDR borrowers would pay less if the accrued interest is not automatically added to the loan balance, as is the case in Income Based Repayment (IBR), Pay As You Earn (PAYE), and Revised Pay As You EARN (REPAYE). Even though borrowers in the standard repayment plan take less time to pay off their loans, they may pay more interest than borrowers in IBR, PAYE, and REPAYE. In contrast, accrued interest is added to the loan balance annually under Income Contingent Repayment (ICR) and monthly under the United Kingdom's income-driven loan plan.

Differences between the five available IDR programs in the US as well as the programs in Australia and the United Kingdom make it difficult to compare borrower outcomes across programs. Further complicating these comparisons, IDR borrowers in the U.S. have access to a related benefit: Public Service Loan Forgiveness (PSLF). Borrowers who work in public service organizations may have their loans forgiven after

ten years of repayment through PSLF. This benefit is not limited to IDR borrowers, but they do have a distinct advantage. While many borrowers in the standard loan repayment plan will have paid off their loans after ten years, IDR programs do not forgive borrowers' remaining debt until 20 or 25 years, depending on the program. Furthermore, PSLF includes a wide array of qualifying organizations. While many other federal loan forgiveness programs are limited to a particular profession or region, PSLF includes borrowers in any role who work for a 501(c)3 organization or for the government (Hegji, Smole, & Heisler, 2014). IDR borrowers would have two incentives to work in the public sector: income-variable payments to adjust for lower salaries and a shorter time to forgiveness. Although not directly a part of IDR, PSLF is an important factor when examining IDR borrowers' outcomes.

In the case of supporters and detractors alike, many claims about IDR are based on theories about borrowers' perceptions and behaviors rather than observations of borrower outcomes. In this study, I aim to ground the policy debate in empirical evidence of the relationship between repaying student loans through IDR and borrowers' financial outcomes, including loan repayment success and career choice. With a more robust evidence base, the debate on expanding or reforming IDR in other ways can be built on a baseline estimate of the status quo. Are IDR borrowers more likely to avoid defaulting on their loans? Do they feel more comfortable taking low-paying but rewarding public service jobs? Existing research cannot answer these questions. This study provides the first evidence of whether IDR improves borrowers' lives, using the first available nationally representative data on loan repayment plans. This new evidence extends the literature on student loan repayment in a new direction, in order to help researchers and

policymakers understand whether increasing enrollment in IDR may increase loan repayment success. By investigating the relationship between IDR and public service careers, this study bridges the gap between research on postsecondary finance and students' workforce outcomes. Colleges may benefit from understanding whether increasing enrollment in IDR can help them encourage more students to give back to the community.

Context

IDR options for federal student loans.

There are currently five IDR options for federal student loans: Income Contingent Repayment (ICR), Income Based Repayment (IBR), New Borrower IBR, Pay As You Earn (PAYE), and Revised Pay As You Earn (REPAYE). Each differs based on a number of factors, such as the percentage of income that borrowers pay toward their debt, how income is defined, the number of years they are required to pay, whether or not loan balances are forgiven after a certain number of years of repayment, and which loans are eligible for the repayment programs.

Revised Pay As You Earn (REPAYE) is the most recent iteration of IDR. Borrowers who enroll in the REPAYE program have their loans capped at 10% of their discretionary income, calculated as the amount by which the borrower's income exceeds 150% of the poverty level. After 20 years of repayment, the remainder of an undergraduate borrowers' unpaid balance is forgiven. For borrowers who take on debt for graduate school, the loan balance is forgiven after 25 years. If the borrower has worked in a public service position for ten years, she may be eligible to have her loan balance forgiven after ten years instead (U.S. Department of Education, n.d.).

Table 1 presents the details of these plans (adapted from U.S. Department of Education, 2016a, 2016b).

Table 1. Comparison of IDR options for federal student loans

	ICR	ORIGINAL IBR	NEW BORROWER IBR	PAY AS YOU EARN	Revised Pay As You Earn (REPAYE)
Year Implemented	1994	2009	2014	2012	2015
Eligible Borrowers	Direct Loan borrowers (FFEL borrowers may qualify through consolidation)	Direct Loan and FFEL borrowers	Direct Loan new borrowers on/after 07/01/2014	Direct Loan new borrowers on/after 10/01/2007 who receive a Direct Loan on/after 10/01/2011 (FFEL new borrowers on/after 10/01/2007 may qualify through consolidation)	Direct Loan borrowers regardless of when borrowing began

	ICR	ORIGINAL IBR	NEW BORROWER IBR	PAY AS YOU EARN	Revised Pay As You Earn (REPAYE)
Eligible Loans	All Direct Loan types *except* Parent PLUS and Direct PLUS Consolidation Loans made before 07/01/2006. **Note:** Direct Consolidation Loans made on/after 07/01/2006 that repaid Parent PLUS may be repaid under ICR.	All Direct Loan and FFEL loan types *except* Parent PLUS and consolidation loans that repaid Parent PLUS.	All Direct Loan types *except* Parent PLUS and consolidation loans that repaid Parent PLUS.	Same as New Borrower IBR.	Direct Subsidized and Unsubsidized Loans Direct PLUS Loans made to graduate or professional students Direct Consolidation Loans that did not repay any PLUS loans made to parents Eligible if consolidated: Subsided and Unsubsidized FFEL, FFEL PLUS loans made to graduate or professional students, FFEL Consolidation Loans that did not repay any PLUS loans made to parents, Federal Perkins Loans
Income Requirement to Qualify	None	Must have partial financial hardship (PFH) to initially qualify and to continue making payments based on income.	Same as Original IBR.	Same as Original IBR.	None

	ICR	ORIGINAL IBR	NEW BORROWER IBR	PAY AS YOU EARN	Revised Pay As You Earn (REPAYE)
PFH Definition	N/A	Annual amount due on eligible loans under 10-year standard plan must exceed 15% of discretionary income.	Annual amount due on eligible loans under10-year standard plan must exceed 10% of discretionary income.	Same as New Borrower IBR.	None
Monthly Payment[1]	Lesser of: 20% of discretionary income, divided by 12; or Amount repaid annually over 12 years under standard plan multiplied by income percentage factor corresponding to AGI and marital status, divided by 12.	*PFH periods:* 15% of discretionary income, divided by 12. *Non-PFH periods:* 10-year standard plan payment based on amount owed when entered IBR.	*PFH periods:* 10% of discretionary income, divided by 12. *Non-PFH periods:* 10-year standard plan payment based on amount owed when entered IBR.	*PFH periods:* 10% of discretionary income, divided by 12. *Non-PFH periods:* 10-year standard plan payment based on amount owed when entered Pay As You Earn.	10% of discretionary income, divided by 12, regardless of income level.
Discretionary Income	Difference between AGI and 100% of poverty line amount for family size and state.	Difference between AGI and 150% of poverty line amount for family size and state.	Same as Original IBR.	Same as Original IBR.	Same as Original IBR.

[1] Under the ICR Plan, if the calculated monthly payment is greater than $0, but less than or equal to $5, the required monthly payment is $5.
Under the Original IBR, New Borrower IBR, and Pay As You Earn Plans:

- If the calculated monthly payment is less than $5, the required monthly payment is $0.
- If the calculated monthly payment is greater than or equal to $5, but less than $10, the required monthly payment is $10.

	ICR	ORIGINAL IBR	NEW BORROWER IBR	PAY AS YOU EARN	Revised Pay As You Earn (REPAYE)
Repayment/ Forgiveness Period	Remaining balance forgiven after 25 years of qualifying repayment (includes periods of economic hardship deferment).	Same as ICR.	Remaining balance forgiven after 20 years of qualifying repayment (includes periods of economic hardship deferment).	Same as New Borrower IBR.	Remaining balance forgiven after 20 years for undergraduates and 25 years for borrowers with graduate school loans
Interest Subsidy	None.	If payment does not cover interest on subsidized loans, government pays remaining interest for 3 consecutive years (excluding periods of economic hardship deferment) from repayment start date under plan.	Same as Original IBR.	Same as Original IBR.	If payment does not cover accrued interest, the government pays the full amount of the difference on subsidized loans for the first three years and half of the difference afterward, as well as half of the difference on unsubsidized loans.
Interest Capitalization	If monthly payment doesn't cover all interest, unpaid interest is capitalized annually. Amount capitalized under this condition limited to 10% of original principal balance at time borrower entered ICR Plan.	If monthly payment doesn't cover all interest, unpaid interest is capitalized when borrower no longer has PFH or leaves plan. No limit on amount that may be capitalized.	Same as Original IBR.	Same as Original IBR, except that amount capitalized under these conditions is limited to 10% of original principal balance at time borrower entered Pay As You Earn Plan.	Unpaid interest is capitalized if borrower is removed from the plan for failing to recertify income annually or voluntarily leaves the program. No limit on the amount of unpaid interest that may be capitalized.

Borrower experiences in Income-Driven Repayment

This study will focus on ICR and original IBR, the IDR options available to Direct Loan and FFEL borrowers[2] in the Baccalaureate and Beyond dataset (B&B:08/12) I used in this study. Although there were fewer IDR options in 2009, the IDR enrollment process was nonetheless complicated. Then as now, all borrowers were automatically enrolled in a standard, 10-year repayment plan. Direct Loan borrowers could then choose to enroll in IBR by contacting their loan servicer to determine whether or not their income level made them eligible. The loan servicer, the organization that collects payments and manages the loan repayment process, determined whether borrowers qualified for a "partial financial hardship" based on the ratio of their repayment amounts in the standard plan to their current income level. If the borrower earned less than 150% of the federal poverty level for her state and family size, then her payments would be $0 a month. If her income fell above the poverty level, her payment amount would be set to 15% of the difference between her income and the poverty level. Direct Loan and FFEL borrowers could choose to enroll in ICR regardless of their income levels. ICR borrowers' payment amounts would be either 20% of their discretionary income or based on the standard loan repayment plan amount, whichever was lower. Under both ICR and IBR, borrowers' monthly payment amounts would never be larger than the standard loan repayment amount.

[2] The Federal Family Education Loan program (FFEL) was discontinued in 2010. FFEL loans were guaranteed by private banks through a public/private partnership with the federal government. Since July 1, 2010, all new federal loans are originated and guaranteed by the federal government.

Once the borrower was deemed eligible for and enrolled in IBR or chose to enroll in ICR, she would be required to submit proof of her family size and income level annually, even if this information had not changed from the previous year. In order to do so, borrowers had to re-submit the application they used to enroll in IBR or ICR the first time, with updated information. Borrowers could also choose to re-submit the application before the annual recertification if their family size or income level changed. Loan servicers were required to remind borrowers about the annual recertification process to ensure that borrowers remained enrolled in IBR or ICR. Borrowers would be placed on a standard loan repayment plan if they did not submit the paperwork to recertify annually.

Enrollment in Public Service Loan Forgiveness

Enrollment in Public Service Loan Forgiveness (PSLF) is comparatively simple. In order to qualify for PSLF, borrowers must submit proof of ten years of loan repayment and concurrent full-time employment for a public-service organization beginning on or after October 1st, 2007. Eligible loans include Stafford, Grad PLUS, and Consolidation Loans in the Direct Loan program. Borrowers may be enrolled in ICR, IBR, or in the standard loan repayment program, but not the graduated or extended repayment plans. Eligible public-service organizations include 501(c)3 organizations, the federal government, state governments, local governments, tribal governments, the military, public schools, public colleges and universities, Americorps, and the Peace Corps. A borrower's employment may have also been deemed eligible if she did not work for "a business organized for profit, a labor union, a partisan political organization, or a non-profit organization engaged in religious instruction, worship services, or any form of proselytizing;" or if her employer provided "Emergency management, military service,

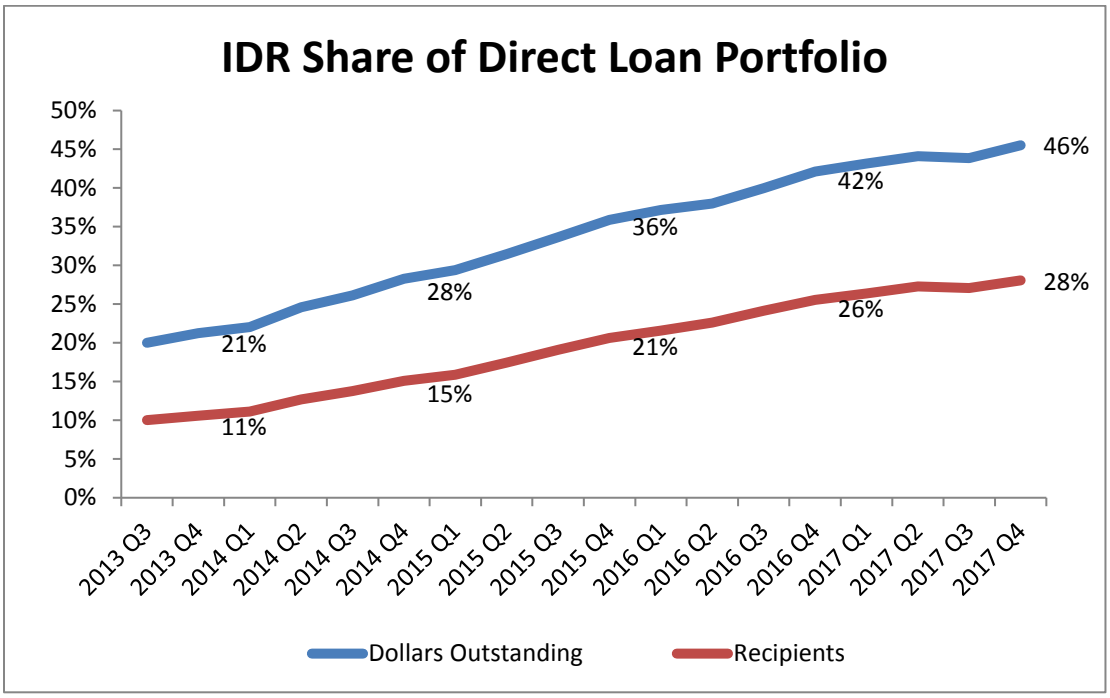
public safety, law enforcement, public interest law services, early childhood education (including licensed or regulated health care, Head Start, and State funded prekindergarten), public service for individuals with disabilities and the elderly, public health (including nurses, nurse practitioners, nurses in a clinical setting, and full-time professionals engaged in health care practitioner occupations and health care support occupations, as such terms are defined by the Bureau of Labor Statistics), public education, public library services, school library or other school-based services" ("Federal Perkins Loan Program, Federal Family Education Loan Program, and William D. Ford Federal Direct Loan Program; Final Rule. 34 C.F.R. § 674, 682, and 685.," 2008).

Borrowers must make 120 qualifying payments to be eligible, although the payments do not need to be consecutive. Although PSLF was implemented in 2007, ED did not release a form to document qualifying payments and employment until 2012. The first PSLF borrowers were eligible for forgiveness in October 2017.

Growing enrollment in IDR

As Figure 1 demonstrates, enrollment in IDR has increased almost threefold since early 2013. The total amount of debt held by these borrowers has also increased rapidly.

Figure 1. Growth in IDR borrowers and total debt from 2013 to 2017



Eligibility for IDR expanded to all Direct Loan borrowers in late 2015 with the advent of REPAYE. Many support making IDR the standard or sole repayment option for all student loans (Boatman, Evans, & Soliz, 2014; Dynarski & Kreisman, 2013; New America et al., 2014). The arguments for expanding enrollment include making loan payments affordable in proportion to income and reducing the incidence of default (Chapman, 1997; New America et al., 2014). However, IDR has been available since 1993, and neither enrollment patterns nor research evidence support these hypotheses. Why has enrollment in IDR taken so long to grow, if policymakers and researchers expect such positive outcomes for borrowers?

One possible explanation is the complexity of the programs and confusion about what they entail. The current version of IDR, REPAYE, has only been in place since late

2015. At least four changes have been made to IDR since 2009, and many more have been proposed or promised. Borrowers and others report that enrolling in and repaying through IDR is extremely complicated (Nelson, 2012). To participate in IDR, borrowers are required to submit proof of their financial hardship eligibility, income and family size every year to their loan servicer. Although borrowers may transfer information from the Internal Revenue Service (IRS) to the Department of Education website to determine their eligibility, they are not able to directly submit their IRS data to their loan servicers.

Since 1991, IDR proponents have suggested establishing a relationship with the IRS to collect loan repayments, rather than placing the responsibility on borrowers to report their income each year. In 1983, 1985, 1991, 1993, 2006, and 2007, former Representative Petri (R-WI) introduced the Income-Dependent Education Assistance (IDEA) Act, under which all federal student loan borrowers would repay their debt as a share of their income. Although his proposal did not include any loan forgiveness options and repayment levels were set at a higher percentage of income than the current program, the proposal would have made it easier for borrowers to maintain enrollment.

Some believe enrollment in IDR may be low due to a lack of awareness about the program. In the fall of 2013, the US Department of Education emailed 3.5 million borrowers to inform them they were eligible to enroll in IDR. Enrollment grew by 210,000 borrowers during that period (Stratford, 2014), although it is unclear this growth was directly related to the information campaign. The Department of Education has continued these efforts to increase enrollment in IDR by linking information about IDR for eligible taxpayers through a partnership with TurboTax (Field, 2014).

International IDR programs.

The United States is not the only country with IDR programs, nor was it the first. The US began offering IDR nearly 30 years after student loans were first available from the federal government. The U.S. federal student loan program was established in the mid-1960s in order to make it easier for middle-class students to go to college. While low-income students had access to federal grants, such as the Pell Grant, and high-income students could afford to pay tuition and living expenses out of pocket, middle-income students lacked the financial capital to cover these upfront costs. As reliance on the federal loan system grew through the 1970s and 1980s, so did concern that students might not be able to manage their debt. Current and prospective college students could not be certain that their investment would pay off. What was needed was some sort of guarantee that the risk of taking on student debt would be worth it in the long run. The policy solution for this problem was IDR.

Other countries offered IDR for their national student loan programs, although in different contexts. Australia was the first country to implement IDR on a large scale. In 1989, Australia's conservative governing coalition simultaneously introduced tuition fees and established the Higher Education Contribution Scheme (HECS), which allowed students to defer all of their tuition fees by offering loans at zero interest. Borrowers repay these loans as a portion of their income after they reach a minimum income threshold. The National Tax Office manages repayments. Some reforms have occurred over the past nearly thirty years, including the addition of more flexible repayment options, differential tuition fees and variable repayment rates. Despite the fact that the program, now called HECS-HELP, was designed to make upfront costs more manageable

for low-income students, low-income student enrollment has not changed (Chapman, 1997). However, HECS-HELP has been successful in increasing overall revenue for higher education, which allowed for significant expansion of the system (Jackson, 2003).

The United Kingdom's IDR program was introduced in 1990. Because there were no tuition fees at that time, the original IDR program was intended to supplement need-based grants for costs of living. In 1998, the New Labour government introduced tuition fees and replaced the grants with loans. Need-based grants have since been re-established, but tuition fees have more than doubled. Similar to the Australian program, loan payments are automatically collected through employer withholding and borrowers enter repayment after reaching a minimum income threshold. Unlike in Australia, students may not pay their tuition upfront. Because the IDR policies have changed many times since they were first implemented, there is some concern that complexity has inhibited their effectiveness. One survey of college students indicated that underrepresented minority students were less likely to know current tuition fee amounts and financial aid options (Adnett & Tlupova, 2008). It seems likely that the frequent reforms to IDR in the US have similarly disadvantaged low-income students, just as the complexity of financial aid applications has been a barrier for students in the rest of the higher education system (Bettinger, Long, Oreopoulos, & Sanbonmatsu, 2009).

In addition to Australia and the United Kingdom, several other countries have IDR programs. New Zealand's program is similar to Australia's, in that all borrowers are eligible regardless of income, the program has a zero real interest rate, and payments are collected through payroll withholding (Hillman & Gross, 2014). Chile, Hungary, Sweden, and South Africa also offer IDR to student loan borrowers (Hillman, 2013).

Because of their scope and longevity, much of the literature on IDR comes from Australia and the UK. Understanding the unique features of these programs is important for interpreting this literature and determining its applicability in the US context, as I discuss in Chapter Two.

Definitions

In academic and policy literature, many terms are used to describe variations in IDR policies, and policies with similar features may be referred to by very different names. To clarify the main differences between these policies, I propose four broad categories: income share agreements, graduate taxes, income-contingent loans, and IDR. The following typology explains the differences between these categories, although it is important to note that the study I propose focuses exclusively on one category, IDR.

Figure 2. Typology of Income-Related Financing Policies

	Mandatory	**Optional**
Caps on amount paid	Income-Contingent Loans	Income-Driven Repayment
No caps on amount paid	Graduate Taxes	Income Share Agreements

One way in which these categories differ is in whether or not borrowers' payments are capped. Income-contingent loans, such as those in Australia and the United Kingdom, as well as IDR in the United States all require student loan borrowers to pay only until their balance and any accrued interest has been repaid or forgiven. In contrast, graduate taxes, sometimes referred to as Pay It Forward programs, and income share

agreements, sometimes referred to as human capital contracts, require borrowers[3] to pay for a pre-determined period of time, regardless of the amount the borrower pays during this time. Borrowers may benefit under plans that do not have caps on the total amount paid if they pay less than they otherwise would have under a traditional loan. For example, a borrower who attends an expensive school but earns below-average wages afterward might pay less under a graduate tax or an income share agreement than if she had paid for school with a traditional loan (American Institutes for Research, 2016; Delaney & Dharmapala, 2015). In contrast, borrowers who earn a lot after college may end up paying substantially more than the amount they originally received. This has been called the "Mick Jagger effect" in reference to the musician's one-year stint at the London School of Economics (p. 197, Hansen, 2008). If he had signed an income share agreement, Mick Jagger could have paid a small fortune for an education that, some might argue, did not contribute to his success. The same could be said for other successful college dropouts, such as Mark Zuckerberg and Bill Gates. Under income-related financing programs without caps on amounts paid, borrowers' benefits depend both on the upfront cost of school and their relative success on the job market. These programs would be most favorable to borrowers who attend expensive schools and earn little afterward or have high debt to income ratios. Borrowers who attend inexpensive schools and earn above-average wages after school would benefit more in an income-

[33] Proponents of income share agreements and graduate taxes would argue that participants are not "borrowers" because they do not owe any particular amount of money. However, it is unlikely the participants themselves would make this distinction given that the only post-school financial obligations related to school costs in today's system are loans. I refer to all participants as borrowers for simplicity.

driven financing program with caps on the amount paid. In this sense, uncapped payment models are more equitable than capped (Delaney & Dharmapala, 2015).

Another way to characterize the difference between these policies is by the scope of their participants. In the literature and in practice, income-contingent loans and graduate taxes are mandatory programs, involving students in entire regions or countries. Students are automatically enrolled in income-contingent loans in Australia, the United Kingdom, and New Zealand (Asher, Cheng, & Thompson, 2014). Pay It Forward is also often proposed as a financing program for all public postsecondary institutions in a state (Smith, Gibson, & Burbank, 2016). While students have the option to opt out of income-contingent loans in some countries and under some Pay It Forward proposals, both sets of policies rely on wide participation in order for high-income borrowers to subsidize low-income borrowers. In order for these policies to be financially sustainable, governments must be able to minimize their risk by expecting participants to experience some degree of income inequality. In contrast, IDR and income share agreements are optional and co-exist with other loan programs. These models are financially viable because they can be subsidized through other sources of revenue. For example, the costs of the income-driven loan repayment options for US federal student loans are expected to be very high when loan forgiveness kicks in (Delisle, 2016), and they will have to be paid through tax increases or through cost-savings in other budget items. Similarly, income share agreements could prove costly for individual investors if they are not able to accurately predict students' future earnings or if their investment is compromised by other factors. These small-scale policies are inherently more risky than programs that spread the risk over a larger number of individuals, although the losses might be small overall.

These categories are neither all-inclusive nor static, as illustrated by changes to IDR in 2015. REPAYE made IDR available to all federal student loan borrowers, regardless of their income. REPAYE also removed caps on the amounts borrowers pay each month. It is possible for borrowers' monthly payments to be higher in REPAYE than they would be in a standard 10-year repayment plan. Expanding eligibility could make IDR as common as income-contingent loans, even though REPAYE is not mandatory. Removing caps on payment amounts also makes REPAYE similar to an income share agreement, although REPAYE does not require borrowers to pay more than their principle balance and accrued interest. Future changes to IDR programs may continue to blur the lines between income-related financing options. Nonetheless, a typology such as the one in Figure 3 will help to explain the directions of these changes.

Problem Statement

This study addresses three key problems: understanding IDR enrollment patterns, comparing IDR borrower outcomes to borrowers' outcomes in other repayment plans, and estimating differences in total amounts paid between IDR and other plans. These three problems lie at the core of the literature gap on IDR.

IDR participants unknown.

There have been few empirical evaluations of student outcomes in IDR due to the scarcity of detailed longitudinal data on student loan borrowers. Some have attempted to overcome the difficulty of collecting individual-level longitudinal data on loan repayment by projecting income over time by levels of educational attainment (Beer & Chapman, 2004; Debande & Vandenberghe, 2007). Another study uses survey evidence from borrowers to understand how enrollment in postsecondary education has changed since

the implementation of IDR reforms (Callender, 2002). However, only one of these studies (Herbst, 2018) controlled for participation in IDR programs at the individual level or examined how borrower outcomes are related to alternative financial aid options to determine whether IDR was related to observed borrower outcomes. The present study uses nationally representative individual-level data that includes information about what type of loan repayment plan borrowers chose after graduating from college. This is the first time this data has been collected in the United States, and thus offers a unique look at who chooses IDR. Furthermore, this dataset disaggregates IDR enrollment by socio-economic status, gender, race, and ethnicity. These descriptive statistics will provide a baseline for understanding which borrowers would be affected by policy changes.

Unclear relationship with PSLF.

Much of the empirical literature on IDR focuses on the effects of the program on borrowers before they enroll in college (Callender, 2003; Chapman & Ryan, 2005; Caetano, Patrinos, & Palacios, 2011). These studies examine whether IDR is a more effective signaling device to ensure prospective students of affordability than standard loan repayment options. There are many reasons that students choose to enroll in college, and these studies are unable to address the interaction between the availability of IDR and a host of other reasons that affect student enrollment. More problematically, many of these studies attempt to compare enrollment decisions before and after major policy changes that introduced IDR (Beer & Chapman, 2004; Chapman & Ryan, 2005). In these studies, borrowers often do not have the option to choose between IDR and another loan repayment plan. Comparing student decisions across historical cohorts introduces a large number of confounding factors that may be wholly unrelated to loan repayment plans.

In contrast, this study compares the post-graduation outcomes of borrowers who choose to pay their loans through IDR to those of borrowers who pay through other repayment plans. These post-graduation outcomes include career choice and loan repayment success. Theory suggests that borrowers who choose IDR would be more likely to pursue public service careers, because lower-paying jobs would not jeopardize their ability to make monthly loan payments and because IDR borrowers have an implicit incentive to seek PSLF. This study uses propensity score analysis to statistically control for borrowers' propensity to enroll in IDR compared to other loan repayment plans. This procedure allows me to estimate whether IDR is related to borrowers' occupational choices.

Default assumption untested.

Furthermore, IDR borrowers should be less likely to have trouble repaying their loans than borrowers in other repayment plans. IDR allows payment amounts to vary in proportion to borrowers' incomes, meaning that loan payments will always be a predictable percentage of borrowers' budgets. Furthermore, IDR payment amounts for the borrowers in this study are lower than repayment amounts for similar borrowers in other plans. The IDR borrowers in this study were enrolled in ICR or original IBR, both of which have a partial financial hardship requirement for eligibility. The flexibility of payment amounts and the partial financial hardship requirement should both contribute to lower rates of forbearance and default among IDR borrowers.

Possible differences in total payment amounts.

One purpose of IDR is to make borrowers' payments predictable and proportional to their incomes, or in other words, affordable over the short term. IDR is not intended to

decrease the amount of money that borrowers pay over the long term. Although some IDR borrowers may pay less than borrowers in the standard 10-year repayment plan over time due to loan forgiveness, I believe it is more likely that IDR borrowers will pay more than borrowers in other plans due to the way that interest is capitalized. In this study, I examine differences in the average predicted amount paid between borrowers IDR and in other plans. I also investigate whether or not there are differences in the average predicted amount paid among IDR borrower groups, including race, ethnicity, gender, and family income level before college. The purpose of these estimations is to determine how much IDR borrowers will pay in the long term for decreasing their loan payments in the short term, as well as whether certain groups of IDR borrowers pay more than their peers. These estimates assume that all IDR borrowers remain enrolled for the entire length of their repayment, and that all borrowers in both plans have the same internal discount rate. Although the added long-term cost of IDR may be worth it to help some borrowers avoid default or delinquency, it is helpful for policymakers and administrators to understand whether all IDR borrowers face the same costs.

Purpose Statement and Research Questions

This study examines the relationship between repaying loans through IDR, student loan repayment success, and early career outcomes. I answer four research questions.

- First, what are the demographic characteristics (i.e., race, ethnicity, gender) of borrowers who enroll in IDR programs compared to borrowers who are eligible for IDR but repay through the standard ten-year plan?

- Second, are IDR borrowers more likely than borrowers in standard loan repayment to work in public service organizations?
- Third, are IDR borrowers less likely than borrowers in standard loan repayment to be delinquent or in default on their student loan payments?
- Fourth, how do the estimated total amounts repaid differ between IDR borrowers and borrowers in standard repayment as well as among IDR borrowers?

Statement of Potential Significance

There are many potentially important implications of the results of this study on scholarly research, public policy, and practice. Addressing the field of scholarly research, this study adds empirical evidence regarding student outcomes to fill the gap in existing literature on IDR. Additionally, the study tests the extent to which existing federal datasets from the Department of Education can be used to evaluate the outcomes of IDR programs. Borrower information is generally stored in the National Student Loan Data System (NSLDS), which is not typically accessible to non-governmental researchers. With regard to policy, the results of this study help inform policymakers' decisions about expanding or otherwise reforming IDR. Finally, financial aid practitioners and loan servicers may benefit from this study through a better understanding of which borrowers benefit most from these programs. Servicers could dissuade borrowers with expected lower than average incomes from enrolling in IDR if these programs will be more costly for them in the long run.

The results of this analysis will allow researchers and policymakers to answer key questions about the effectiveness of IDR compared to standard loan repayment plans.

Although roughly 50% of today's federal student loan borrowers are eligible for IDR, only 19% participate in IDR programs (U.S. Government Accountability Office, 2015). The Department of Education has made substantial efforts to increase IDR participation, in part because it may help decrease rates of loan default and PSLF uptake.

Limitations and Assumptions

Given the constraints of the dataset used, this study focuses on a narrowly delimited population. The students are all graduates from four-year colleges. Loan default is more common among borrowers who do not complete credentials (Looney and Yannelis, 2015), and these borrowers are not represented in my dataset. In order to limit the sample to students who have only undergraduate debt and would be eligible for jobs at the same level of seniority and pay, the sample also does not include graduates who attended post-baccalaureate programs. Thus, these estimates are not be applicable to all college students, nor do they apply to borrowers who enroll in the latest IDR program, REPAYE, because it does not have a financial hardship requirement for eligibility. While these limitations affect the scope of the study, they actually underscore its potential impact. Because my estimates are based on borrowers who are *less* likely to default on their loans, I actually underestimate the potential effectiveness of IDR for borrowers who do not earn credentials or who are otherwise more at risk of default than the borrowers in the dataset used for this study.

Chapter Two

Even as support for IDR grows among policymakers, the evidence upon which IDR policies are based remains weak. There have been no rigorous empirical studies in the United States of the effects of IDR on student enrollment, persistence, or completion rates (Hillman & Gross, 2014) and very few investigating the ways in which IDR affects borrowers after leaving school (Herbst, 2018; Mueller & Yannelis, 2017). Evaluations of IDR in other countries have focused on the distribution of public subsidies (Beer & Chapman, 2004; Dearden, Fitzsimons, Goodman, & Kaplan, 2008; Flannery & O'Donoghue, 2011; Vandenberghe & Debande, 2005), estimates of administrative costs (Callendar, 2002), and changes in enrollment decisions (Callendar, 2002; Chapman & Ryan, 2005). One international study examined repayment success across countries' loan programs (Shen & Ziderman, 2009). This review synthesizes the available empirical research on IDR borrower outcomes as a basis for future evidence-based policymaking. In particular, I emphasize studies that have examined how IDR influences student behavior and outcomes after leaving college. The conceptual framework connects the empirical evidence to theory in a model explaining how IDR policies are intended to influence students and how effective IDR policies are at meeting these goals, according to academic and policy research.

To conduct this literature review, I searched ERIC, EBSCO, JSTOR, and Google Scholar using the terms "income contingent AND loans," "income based AND loans," and "income driven AND loans." I also used genealogical search methods to find additional sources from reference lists. In order to gather as many studies as possible, I did not limit the date ranges in this search. However, I gave priority to studies based on

empirical observations of IDR policies, which narrowed the number of relevant articles considerably. Many of the empirical studies are based on IDR policies in Australia or the United Kingdom, due to the fact that the policies have affected a large number of students and have been in effect long enough to study those students' outcomes after leaving school. I included some non-peer reviewed white papers and policy briefs to provide a more substantial U.S. perspective, though these papers were unable to draw upon the same breadth of evidence as the international studies. Overall, the U.S. literature included in this review relies heavily on simulations, student surveys, and small field experiments. Thus far, there have been only two U.S. studies of observed students' choices among multiple loan repayment options and their outcomes once they begin repaying their loan debts (Herbst, 2018; Mueller & Yannelis, 2017). Both studies use administrative data and provide the best evidence of how IDR affects borrower behavior after leaving school.

Literature Review

Although IDR policies are well established, the empirical evidence of the effectiveness of IDR is rather scarce. Many studies of IDR in the U.S., Australia, and the United Kingdom rely on small sample sizes, use estimates of student and borrower behavior rather than observed outcomes, or do not control for selection into the program. However, the literature does provide some insight into how students view loans, in particular how students manage risk. To supplement these studies, I have also included literature on loan forgiveness programs for particular professions. These studies can provide an indication of what evidence is needed in IDR, as well as how individuals in specific settings interact with similar programs.

The main findings in this review are: 1) knowledge about IDR varies, and students may not make economically rational decisions with this information; 2) IDR affects aspects of students' long-term financial decision-making; and 3) IDR could encourage borrowers to pursue public service careers. Taken together, these themes may help inform future research on IDR, as well as provide policymakers and practitioners with hypotheses for ways IDR could be improved.

IDR selection.

There is limited evidence on the factors that contribute to borrowers' decisions to select IDR. Beth Akers and Matthew Chingos (2014) estimate that 52% of U.S. student loan borrowers are eligible to participate in IDR, and about 10% would be eligible for loan forgiveness. However, actual enrollment in IDR lagged far below eligibility at only 19% (U.S. Government Accountability Office, 2015). Even today's enrollment, at 28% of the Direct Loan portfolio, is below this estimate. One possible explanation for low IDR enrollment may be that borrowers are sensitive to how loans are described. Thus, changing how IDR is described may be sufficient to increase enrollment. This was demonstrated in an experiment that found that labeling financially equivalent loan options differently influences students' willingness to borrow (Caetano, Patrinos, & Palacios, 2011). The researchers conducted an experiment with students in Chile, Columbia and Mexico who were randomly assigned to two groups that received hypothetical financial aid packages with different labels. The treatment group description contained the labels "loan" and "human capital contract" (in this case meaning a loan with IDR up to a capped amount), while the control group received the same options without these labels. After controlling for country, gender, age, level of educational

attainment, and children, the study found that eight percent of participants chose the human capital contract option because of its label. Even after explaining to participants that the options were financially equivalent, those in the treatment group were 18 percentage points more likely than participants in the control group to prefer the IDR option to the loan option. These results indicate that some students are averse to loans for non-financial reasons, which perhaps have to do with their perceptions of loan programs themselves. If this is the case for US borrowers as well, then the federal government might be able to advertise IDR as a way to increase participation in federal loan programs. This could potentially help loan averse students who might not otherwise go to college.

A variation of the Caetano, Patrinos, and Palacios study investigated loan averse young adults interest in income-driven financing options in the United States. Peek, Mason, and Soldner (2016) found that loan averse participants' families' and friends' experiences with student loans were critical in shaping their views of student debt. Negative experiences with loans, such as siblings who had experienced bankruptcy brought on by debt, were offered as explanations for participants' loan aversion. Though we discussed two alternatives to traditional student loans, IDR and income share agreements, participants had little to say about IDR. The fact that IDR was described as a loan, whereas ISAs were not, may have had something to do with this.

Both studies focus on hypothetical choices among financing options with slightly different definitions of IDR. They offer the best available evidence regarding the factors that prospective students consider important when making these choices. Participants in both studies indicate that the perceptions of policies play a meaningful role.

Understanding the contexts in which participants make these decisions, and the information they draw upon, will be an important component of any study on IDR choice.

IDR and loan repayment success.

Many believe that IDR borrowers benefit financially from the program, including the borrowers themselves. Rasmussen (2006) examined the ways that Australian students think about IDR and long-term debt. Through interviews with 16 students, Rassmussen learned that students largely preferred IDR to upfront tuition payments because they did not have to rely on their families or their savings to pay the costs. Students had varying levels of understanding about long-term costs and repayment. Most students were confident they would be able to repay their loans.

The only study, to my knowledge, that has examined differences in repayment success in the US at the borrower level found that IDR is related to significantly and substantially higher repayment rates (Herbst, 2018). This study used data from a large loan servicer, matched with data from a credit bureau, to examine IDR uptake and subsequent outcomes for delinquent federal student loan borrowers. Using a difference-in-differences design, Herbst demonstrated that student loan borrowers who chose IDR were 21 percentage points less likely to be delinquent on their payments than borrowers in other payment plans. Though switching to IDR reduced borrowers' minimum payments, IDR borrowers were actually likely to pay off more of their debt than borrowers in the standard repayment plan, a difference of $90 a month on average. It is unclear whether this extra payment amount allowed more IDR borrowers to make payments on principal, rather than simply paying down a larger portion of accrued interest. This effect disappeared after fifteen months in IDR, when IDR borrowers

returned to making the minimum required payment amounts. Herbst pointed out that switching to IDR increases borrowers' cash on hand, which they may decide to use to make purchases rather than retiring their debt more quickly.

There is limited evidence on whether all borrowers, not just those who have experienced delinquency, are actually successful in repaying their loans under IDR. The Government Accountability Office found that between 81% and 87% of IDR borrowers who entered repayment between 2010 and 2014 were in active repayment, compared to just 56% of borrowers in standard repayment plans (U.S. Government Accountability Office, 2015). In contrast, Janice & Voight (2016) note that loan repayment rates in a small sample of institutions are lower for IDR borrowers than for borrowers in other repayment plans, indicating that IDR borrowers are not making equal progress towards retiring their debt. This difference is unsurprising, as the GAO study categorized borrowers as in active repayment if they were not in deferment, forbearance, or default, while Janice and Voight counted borrowers as in repayment only if they were decreasing their loan principal. IDR borrowers might be avoiding some negative repayment outcomes, like default, while still negatively amortizing on their debts.

Unfortunately, ED does not regularly report default rates disaggregated by loan repayment plan, so it is unclear how default, forbearance, and delinquency rates differ for IDR borrowers compared to borrowers in standard repayment. At a servicing summit in 2014, ED presented delinquency rates by repayment plans showing that IDR borrowers typically have lower delinquency rates than borrowers in standard repayment plans. IDR delinquency rates ranged from 2.1% to 15.4%, depending on the IDR plan, compared to 14.9% for borrowers in standard repayment plans (U.S. Department of Education,

Federal Student Aid, 2014). This is despite the fact that lower-earning borrowers are expected to have adverse loan repayment outcomes. Looney and Yannelis (2015) found that low earning borrowers were more likely on average to default on their loans, but that does not appear to be the case for IDR borrowers.

At the country level, one international study examined differences in repayment success by loan type and repayment option. Shen and Ziderman (2009) explain that the amount of a government's loan expenses that are recuperated, or the repayment ratio, depends on a number of factors. If governments lend at below-market rates, do not collect interest during certain periods of time, do not link repayment to inflation, offer long amortization periods or offer forgiveness for various reasons, students essentially receive a "hidden grant" (p. 317) in the form of a subsidy. Measuring hidden grants and repayment ratios separately for US subsidized loans, unsubsidized loans, Perkins loans, and loans in income-contingent repayment (ICR), the authors found that the hidden grant was highest for ICR loans, and the repayment ratio was lowest for ICR loans, of the US loan programs. It is unclear when these data were collected or what time period they measure. The authors note that while many believe that income-driven repayment can lower default, their study does not bear this out. The study examined default rates on IDR programs in the US, Ethiopia, Ghana, and South Africa. In these countries, the repayment ratios were lower once default was taken into account, particularly in the African countries. Default rate data were not presented for Australia, England, or New Zealand, which were included in other parts of the study.

Total repayment amounts.

Research on total payment amounts under IDR suggests that IDR is expensive for the federal government, though estimates vary depending on the total repayment period of the program. Akers and Chingos (2014) simulated total payment amounts for IDR borrowers based on Baccalaureate and Beyond (B&B:93/03) and found that between 7% and 12% of all undergraduate degree borrowers would be eligible for some form of loan forgiveness under IDR. For borrowers in PSLF, the average amount of debt forgiven is $16,305 under IBR and $18,319 under PAYE, while for borrowers who get forgiveness at the end of the repayment period, the average amount forgiven is $44,221 for IBR borrowers and $32,584 for PAYE borrowers. These calculations assumed that all eligible borrowers would choose to participate in IDR.

Using a simulation based on NSLY79 and NSLY97 data, Chen (2017) predicted that 15% of student loan borrowers will have debt forgiven under IDR. For borrowers in PSLF, the average amount of debt forgiven is $6,180, while for borrowers who get forgiveness after 20 years of payments, the average amount forgiven is $4,919. In Chen's model, all borrowers repay through IDR. He does not split the borrower population between IDR and other repayment plans, so these estimates are likely different than what will be observed. In another simulation of NSLY97 data, Ji (2017) predicted that IDR reduces repayment burdens by roughly half compared to standard loan repayment plans. In his simulation, almost no borrowers experienced loan forgiveness after 25 years of payments.

Many countries introduced IDR at the same time they began charging tuition fees, and so international research focuses on repayment amounts rather than repayment failure

rates. Compared to free or zero tuition policies, international researchers have confirmed that students pay more for their education under IDR, as would be expected. In the United Kingdom, policy changes that introduced tuition and replaced grants with loans repaid through IDR resulted in greater loan burdens for low-income students than for wealthier students (Callender, 2002). When the grants were reinstated several years later and tuition more than doubled, most borrowers' overall costs increased again (Dearden, Fitzsimons, Goodman, & Kaplan, 2008). Similarly, in 2005, Australia increased tuition and increased the minimum income threshold at which borrowers must begin repaying loans (Beer & Chapman, 2004). Based on a hypothetical model of repayments based on income estimates, Beer and Chapman found that all borrowers would pay more towards their loans due to the reforms. In these cases, the total amount that borrowers paid for higher education increased after the establishment of IDR, although none of these studies indicates that IDR is the cause of students' cost increases. It seems likely that policymakers choose to establish IDR when tuition fees rise or grant funding falls, as a means of mitigating cost increases. Furthermore, none of the international literature examines whether implementing IDR, regardless of cost increases, resulted in large numbers of loan defaults.

However, compared to other forms of loan repayment, there is evidence that IDR distributes subsidies more equitably among borrowers. In the United Kingdom, IDR was related to smaller loan amounts among ethnic minority students (Callender, 2002) and lower costs for the lowest income women borrowers (Dearden, Fitzsimons, Goodman, & Kaplan, 2008). In Australia, policy reforms resulted in a decrease in the net present value of total loan repayment costs for low-income borrowers and women (Beer & Chapman,

2004). The loans in the Australian system also have a real interest rate of zero, meaning that the borrower does not accrue interest on the loan. Because the study assumed that low-income borrowers and women would take longer to reach the minimum income threshold for repayment, the value of the total loan repayment cost drops due to inflation and discounting. Overall, it seems that under some IDR policies, borrowers who typically earn less after graduation will pay less compared to their higher-earning peers. The disproportionate share of subsidies directed to high-need students under IDR has been strongly emphasized by some researchers as evidence of the distributional justice of the policy reforms (Flannery & O'Donoghue, 2011).

IDR and career choice.

A third major theme in the literature is career choice. One of the few IDR studies in the United States examined the enrollment decisions and outcomes of students who were admitted to NYU Law School (Field, 2009). This study differs considerably from the international literature, because participants were not responding to new tuition fees in addition to IDR. A randomized experiment was conducted in which students were encouraged to enter public service careers through two types of incentives: scholarships that converted to loans if graduates left the public sector or loans that were repaid by the school as long as the graduate worked in the public sector. The loan assistance repayment plan randomly assigned some incoming law students to tuition scholarships that became loans if the student did *not* pursue a public service career after school, while others were assigned to a loan program that forgave the debt of students who *did* work in the public service after school. Because the incentives were financially equivalent, the experiment was able to measure whether risk aversion influenced students' decisions to enroll in the

school and choose public service careers. The study found that scholarship recipients were 30% more likely than loan recipients to enroll at NYU Law School the first year the program was offered, and nearly twice as likely to enroll in the second year. Minority students who received scholarships were less likely to enroll than loan recipients, although the researcher noted that this may have been due to the small sample size of minority students. The study also found that scholarship recipients were 36% more likely to enter public service careers than loan recipients.

In contrast, research on a larger sample of student loan borrowers suggests that IDR does little to encourage public sector work. Using a simulation based on NSLY79 and NSLY97, Chen (2017) predicted just 1% of borrowers would switch to public sector work as a result of IDR. This is a smaller difference than the percentage point difference in public sector employment between the two cohorts studied and roughly equivalent to year-to-year changes in the actual data. Chen speculates that few people will change their employment sector because only those with high debt to income ratios benefit from PSLF, and these borrowers make up a small share of the total borrower population.

Research on loan forgiveness provides further insight into how borrowers' perceptions of debt are related to their career choices. There is extensive literature on loan forgiveness for health care workers, particularly for those who choose to work in rural regions. A meta-analysis of 43 studies on a variety of financial incentives, including loan forgiveness and scholarships with service requirements, for health service professionals to work in underserved areas found a pattern of positive responses to such programs (Bärnighausen & Bloom, 2009). Many of the studies reported that program participants were more likely than non-participants to continue working in an

underserved area beyond the time the program required, although it was not always the same area in which the participants were originally placed. None of the studies in the meta-analysis met the criteria for providing causal evidence, but the authors emphasize that there is an indication that the financial incentive programs were an important factor in participants' career decisions over the long term.

Additional analyses have indicated that loan forgiveness is a successful recruitment tool for organizations seeking to attract professionals to lower-paying public service careers. Based on thirty years of data from a survey of American Association of Medical Colleges graduates, (Phillips et al., 2009) found that medical students who participate in the National Health Service Corps loan forgiveness program were statistically significantly more likely to choose a primary care career, to work in family medicine, to work in a rural health center, or to practice in a rural or underserved area. These outcomes extended beyond the service requirement for loan the loan forgiveness program. Another evaluation examined the National Institutes for Health, which offers a loan repayment program for biomedical scientists who study HIV/AIDS, are from economically underserved backgrounds, and for medical doctors and fellows conducting general research within NIH's centers (Glazerman & Seftor, 2005). The evaluation found that, compared to similar non-participants, scientists who participated in the loan repayment program were more likely to accept job offers from NIH and to remain there for several years. The effects of participation on retention over the long term, however, were less evident for clinical researchers and researchers in AIDS-related fields.

Loan forgiveness programs are also common for recruiting teachers to high-need schools. A survey of participants in a loan forgiveness program for STEM teachers found

that teachers of color believed the program was more influential in their decisions to teach in underserved areas than white teachers (Liou & Lawrenz, 2011). In that survey, teachers also responded that larger financial incentives were more influential. There is unfortunately little research on the long-term effects of loan forgiveness programs for teachers (Kolbe & Strunk, 2012). It seems likely that teachers would respond in the same manner as health care workers, although their debt to income ratios are probably smaller than professionals with advanced medical training who choose to work in underserved communities. If the size of the financial incentive is an important factor in teachers' responses, then the effectiveness of loan forgiveness programs for teachers may vary based on student loan debt amounts.

The theoretical relationship between student loans and job choice also ties into borrowers' earnings. Another simulation based on NSLY97 (Ji 2017) predicted that IDR borrowers would spend longer looking for jobs and earn more than borrowers in standard repayment plans. Ji posits that because IDR removes some of the risk inherent in the job search process, IDR borrowers can be more selective in the job search process and hold out for jobs with higher wages. Thus, the greater protection from labor market shocks that IDR borrowers receive allows them to be more successful in loan repayment over the long term.

IDR and other post-graduation outcomes.

There is little literature examining the relationship between IDR and post-graduation outcomes outside of career choice. Herbst (2018) found that IDR borrowers were more successful than borrowers in other plans on multiple financial measures. IDR borrowers had higher credit scores, higher average credit card balances, and more credit

cards in the first two years after switching repayment plans, indicating that IDR increased short-term measures of consumption. He also found that IDR borrowers were slightly more likely to hold mortgages.

Marks (2009) used a panel study of Australian households, similar to the PSID in the US, to model the effects of Australia's IDR system on post-graduation outcomes. His analysis found that graduates with more debt were slightly more likely to leave their parents' homes and moderately more likely to delay having children. He did not find any difference in home purchases. It is worth remembering that all loans are income-driven in Australia, so the counterfactual in this study would be students who paid out of pocket for college. Considering the implications of this study in the American context, it is unclear whether debt amount or repayment plan type would explain these effects.

Discussion of empirical studies.

Although students enroll in IDR after leaving college and begin repaying their loans, most research has focused on how IDR influences student choices around college enrollment. For example, during the transition between high school and college, information about IDR may influence some students' perspectives on financial aid and willingness to take on student debt (Caetano, Patrinos, & Palacios, 2011; Callender, 2002; Field, 2009). Additionally, IDR may influence students' decisions on whether to enroll in college as well as their selection among colleges (Callender, 2002; Chapman & Ryan, 2005; Field, 2009). Fewer studies focus on the influence of IDR as students transition from postsecondary education to work. Researchers and policymakers assume that IDR influences career choice (Field, 2009), and total loan amount paid (Beer & Chapman, 2004; Dearden, Fitzsimons, Goodman, & Kaplan, 2008; Vandenberghe &

Debande, 2007). Some also assert that IDR will help smooth consumption and make repayment easier, thereby reducing the likelihood of loan default (Barr, 2004; Boatman, Evans, & Soliz, 2014; Hillman, 2013). However, none of these theories regarding borrowers' post-graduation outcomes has been tested on U.S. undergraduate students.

IDR may also alter borrowers' career decisions, because borrowers will not have to worry whether working in a public service organization or other low-paying career will affect their ability to make their loan payments. Furthermore, IDR borrowers are able to take advantage of Public Service Loan Forgiveness (PSLF), a federal program that eliminates borrowers' remaining loan balances after ten years of repayment if the borrower worked for a 501(c)(3) organization or for the government. Although borrowers in standard repayment plans may also qualify for PSLF, only IDR borrowers are likely to have repayment timelines that extend beyond ten years. Due to the implicit incentive in PSLF to participate in IDR, as well as IDR's function as a safety net, borrowers in IDR may be more likely to pursue lower-paying public service careers. A major difficulty with applying the evidence from these studies to IDR is the lack of a direct connection between IDR and PSLF. It is difficult to determine whether borrowers choose IDR in order to take advantage of PSLF, or if they decide to use IDR and then realize they may be eligible for PSLF. In order to make a causal argument that IDR leads borrowers to public service, borrowers must choose IDR first.

These empirical studies provide a picture of the changes, or lack thereof, that resulted from implementing large-scale IDR policies, using the policy change as a sort of natural experiment that could alter student behavior. However, none of these studies controlled for participation in IDR programs at the individual level, or examined

differences in students' decisions based on alternative financial aid packages to determine whether IDR might help risk-averse students and encourage enrollment. This finding is particularly important for low-income students, who are believed to be more sensitive to student loan debt than their peers. If students know that their loan repayments will always be affordable, the prospect of taking on debt may become less daunting. The idea that IDR is a solution to risk aversion is based on the assumption that students are deterred from borrowing to finance postsecondary education or even enrolling at all because of the risk of being unable to repay their loans afterward (Chapman & Greenaway, 2006; Chapman, 2006; Oosterbeek, 1998; Simpson, 1987). IDR policies promote access by removing this risk (Barr, 2009; Jacobs & van der Ploeg, 2006). However, if this were the case then students should choose IDR over other repayment options. Low participation rates in IDR programs may be due to the fact that students are unaware of their options or do not have adequate information about the benefits of IDR (Boatman et al., 2014).

In the aggregate there may be several possible explanations for the gaps in the literature on IDR described above. First, the necessary data may be difficult to access. A rigorous evaluation of an IDR program would require longitudinal data that tracks borrowers through postsecondary education and into the labor force, including information about the borrower's academic program, grades, career planning, employment, salary, and other forms of wealth. In the United States, such linked datasets are just beginning to be developed, but they are not yet connected to data about loan repayment (Hillman & Gross, 2014). Another possible explanation is that IDR programs are just one part of a complex set of factors that students consider when enrolling in postsecondary education and persisting to graduation. In order to isolate the effect of IDR

on enrollment and graduation, researchers could attempt to compare IDR participants to identical students who did not choose IDR. In Australia and the United Kingdom, this would be difficult because almost all students enroll in IDR automatically. In the United States, the problem is just the opposite: until recently very few borrowers chose to repay their loans through IDR. These methodological issues may explain why IDR has so far received little attention from higher education researchers.

Conceptual Framework

Students choose a loan repayment plan soon after leaving school, roughly at the same time as they are making decisions about their careers and how to manage their balance their loan payments with other financial obligations. A variety of factors influence all of these major financial decisions, including students' families, cultures, incomes, and educational contexts. In this study, I draw on multiple theories of decision-making to model the relationship between IDR, loan repayment status, and occupation.

IDR selection.

There is no established model for understanding how borrowers choose loan repayment plans, however it is likely that IDR choice is similar to other choices related to student loans. Student loan decisions are grounded in human capital theory and risk aversion theory. Human capital theory assumes that individuals are resources with economic value, similar to physical or natural resources. Just like physical and natural resources can be transformed to increase their value, individuals can also increase their value by gaining skills and knowledge (Becker, 1962; Mincer, 1974). Individuals invest in postsecondary education because they expect a personal return on that investment in the form of increased earnings over their working lives. Often this investment requires

loans, because students must pay the substantial upfront costs of postsecondary education, such as tuition, as well as its indirect costs, such as lost earnings while in school. Human capital theory offers a simplistic view of college going, in which students are primarily concerned with making the best economic choice. After weighing their preferences for future income increases and their estimation of their own ability to realize those gains against their current opportunities, students make decisions about whether to forgo immediate earnings in return for individual long-term benefits (Becker, 1993; Goldrick-Rab, Harris, & Trostel, 2009).

Risk aversion theory developed in a response to human capital theory's limited view of human decision-making. Borrowing from psychology, McFadden (1999) argues that people make decisions based on a variety of contextual factors and with imperfect information about products. Unlike taking out a mortgage to purchase a house, students have little ability to assess quality of their investment beforehand, and little ability to improve the quality of their education. Furthermore, the value of a degree is not guaranteed to appreciate over time (Barr, 2004; Palacios, 2002). According to risk aversion theory, students may also avoid loans because they are overly sensitive to the risk of failing to realize returns (Tversky & Kahneman, 1973). Anxiety and a belief that the student loan process is overwhelming are common among many student loan borrowers (Mueller, 2013), although this risk is greater for students from low-income backgrounds (Barr, 2004), who have less access to the cultural and social capital necessary to judge educational quality. These borrowers' perceptions of the risk associated with student loans are based on the experiences of the comparatively small number of people who have gone to college.

The earliest conceptualization of IDR brings risk aversion theory and human capital theory together. Friedman (1955) suggested that markets could develop a program to pay for higher education based on students' future earnings, called a human capital contract. Human capital contracts would encourage students to invest in education by minimizing the risk of low earnings after school. Unlike buying a house or car, in which case a loan may be secured against physical capital that can be repossessed, investment in human capital has no collateral. If the borrower defaults on her loan, she still gets to keep her knowledge. Friedman argued that, because college graduates typically earn more than high school graduates, loans could be insured by the likelihood of a wage increase after graduation. By tying loan repayment to income, the lender could be fairly certain the loan would be repaid (Woodhall, 1983). Thus, IDR programs are based on the idea that tying payments to borrowers' income levels will make risk-averse students more comfortable about taking on debt. Because of its basis in risk aversion, IDR is grounded in the idea that a variety of factors influence students' decisions about financial aid, including students' families, cultures, incomes, and educational contexts. Individual decisions about student loans are seen embedded in multiple layers of individual, social, and organizational contexts (Perna, 2006; Tierney & Venegas, 2009).

Borrowers' decisions about how to repay their loans are similar to other financial aid decisions they have made during their postsecondary careers, such as choosing to enroll in college and to persist. Perna's (2006) conceptual model of college choice has been widely used in research on whether or not financial aid is related to students' decisions to enroll in college (Goldrick-Rab, Harris, & Trostel, 2009). Perna's model illustrates students' decisions about going to college nested within multiple layers: the

social, economic, and policy context; higher education context; school and community context; and habitus. The core of this model, habitus, is comprised of students' demographic backgrounds, social capital, and cultural capital. This conceptual model merges human capital theory, in which students make individual decisions about college, and social and cultural capital theory, in which people, institutions, and policies influence students.

Similarly, Tierney and Venegas's (2009) cultural ecological framework explains how multiple overlapping contexts influence financial aid decisions related to college choice. However, Tierney and Venegas suggest that the framework could also be applied to "create a systemic and longitudinal framework for information about financial aid" (p. 384). Noting that these decisions are non-linear and non-recursive, the authors argue that students' decision-making processes are different in 9th grade than in 12th grade. By extension, it could be said that students' decisions about financial aid in their first and final years of college are also different. Unfortunately, no studies have used this framework to examine students' decisions about financial aid after they have enrolled in college.

To understand post-enrollment financial aid decision-making, we can turn to St. John, Paulsen, & Starkey's (1996) college choice-persistence nexus. This conceptual model connects the factors that influence college choice with those that influence students' decisions to remain enrolled each year. Initial enrollment and subsequent persistence decisions are all related to finances, particularly tuition and other costs of attendance. Because the nexus model demonstrates that there are some common elements in the ways that students make financial decisions throughout college, we can assume that

these elements might also affect students' decisions about repaying their loans as they exit college. The nexus model helps explain loan repayment decisions. It also provides a theoretical bridge between the conceptual models linking financial aid and college choice (i.e., Perna's model, Tierney and Venegas's cultural ecological framework) and this study. I chose to use these models, rather than a simple human capital investment model, in order to include a wider variety of possible confounding factors in the propensity and borrower outcome models.

Loan repayment status.

A confluence of contextual factors can also explain loan repayment status. Loan default is related to degree completion, institutional characteristics, borrowers' race and ethnicity, age, family structure, parental education level, income, and academic preparation for college (Gross, Cekic, Hossler, & Hillman, 2009). Longitudinal research on student loan default indicates that the three year window used to calculate loan default rates may be insufficient, as a substantial share of borrowers default on their debts Loan delinquency, or being behind on loan payments, is also disproportionately high among older borrowers and borrowers who attend for-profit institutions (Cunningham & Kienzl, 2011). These factors fit into the conceptual model in Figure 3, and I will include them in the statistical model of borrowers' repayment statuses.

Career.

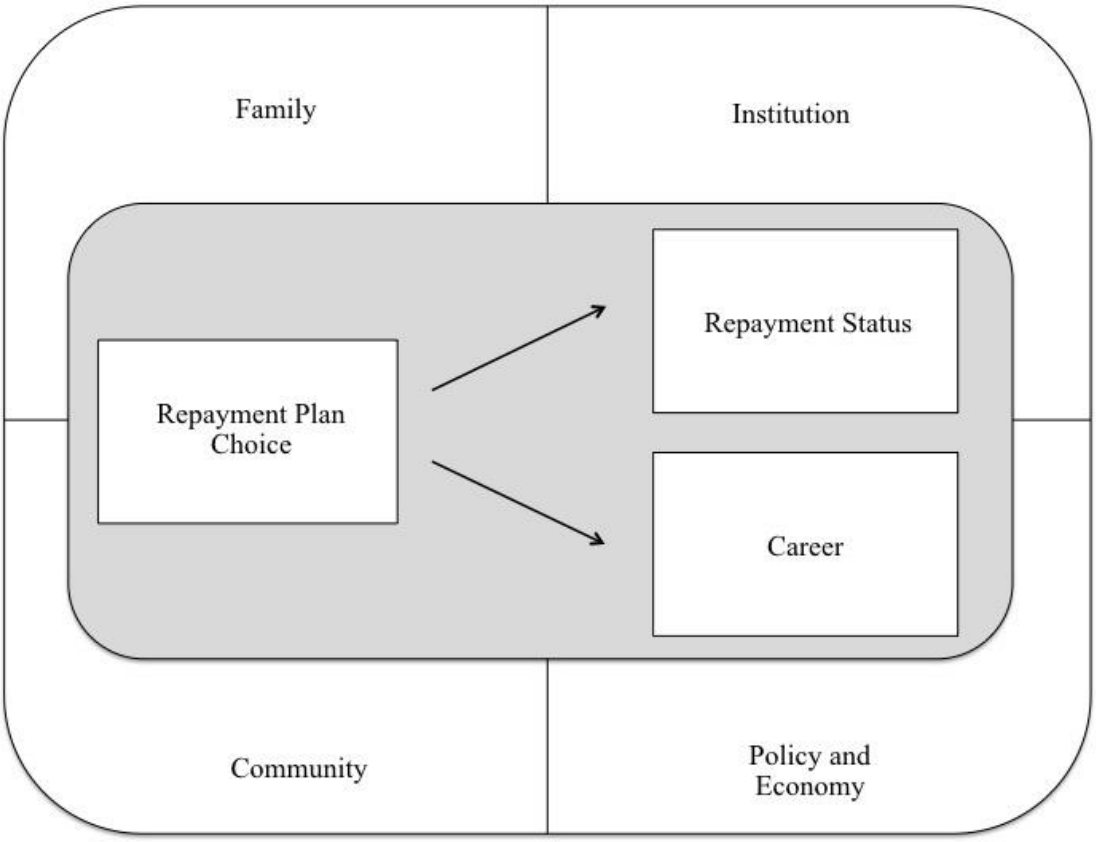
This study also examines whether IDR borrowers are more likely to choose public service careers over similar occupations in the private sector. I assume that this decision is similar to borrower's decisions to repay their loans through IDR as well as the decisions related to avoiding loan delinquency and default. First, the data used in this

study were collected fairly soon after borrowers graduated from college. Their contexts, and the influence of these contexts, are likely to be similar just four years after graduation. Second, public service motivation theory (Perry, 2000) indicates that similar contextual factors influence these choices. According to public service motivation theory, workers choose public sector jobs based on their social, political and institutional contexts, as well as their own values.

Combined model.

This confluence of contextual and individual factors fits well with the model in Figure 3. Family, community, institution, and policy and society combine to form the contexts in which borrowers make decisions about repaying their loans. These decisions are causally related to borrowers' future loan statuses and careers, which themselves are embedded in a combination of contexts. I used this conceptual model to choose variables for my analyses, as I describe in Chapter Three. For example, the family context is represented by borrowers' family income as well as a variable for whether borrowers consulted family about financial aid decisions. Both of these variables indicate the kind and amount of social and cultural capital that borrowers bring to their decisions. To represent the community context, I included a measure of what the borrower offers to the community, in terms of community service, as well as what the community is able to offer to the borrower, in terms of racial segregation. Neighborhood effects research indicates that racial segregation is related to educational outcomes for students of color. Segregation leads to the geographic concentration of wealth in White neighborhoods, thus creating a resource divide between public schools. (For a comprehensive review of these effects, see Johnson, 2010.)

Figure 3. Conceptual model



Chapter Three

In this quantitative evaluation, I investigated whether income-driven loan repayment (IDR) achieves three goals: 1) increasing the likelihood that borrowers will successfully repay their debts, 2) eliminating barriers to public service, and 3) stretching loan repayment over time to make it more affordable for borrowers. If IDR does achieve these goals, IDR borrowers would be less likely to default on their debts and more likely to enter public service than their peers. We would also expect that IDR borrowers would pay roughly the same amount at their peers but over a longer period of time. I used a statistical method called propensity score analysis to compare IDR borrowers with borrowers in standard repayment plans. Using a measure of how likely borrowers were to choose IDR, called a propensity score, I weighted the treatment and control groups based on borrowers' family backgrounds, educations, and other similar experiences. I then estimated the effect of IDR on borrowers' public service career choices, loan repayment success, and projected total payment amounts.

Research Questions

This study answers four questions:

1. What are the characteristics of borrowers who enroll in IDR?

To answer this question, I illustrated the differences between the IDR borrowers to borrowers in standard repayment plans in terms of demographics, education, financial resources, and other characteristics. I also compared these borrower groups to borrowers who are eligible for IDR but choose not to enroll. The differences between IDR participants and IDR-eligible non-participants suggest some reasons why borrowers are unaware of the IDR program or simply choosing not to participate in it.

2. Are IDR borrowers more likely than borrowers in standard loan repayment to work in public service organizations?

To answer this question, I estimated the relationship between the type of organization a borrower works for and variables related to choosing a public service career, including civic engagement and major in college. I estimated this relationship separately for IDR borrowers and borrowers in standard loan repayment after estimating the likelihood they would choose each repayment program.

3. Are IDR borrowers less likely than borrowers in standard loan repayment to be delinquent or in default on their student loan payments?

To answer this question, I estimated the relationship between several loan repayment statuses, including delinquency and default, and variables related to loan repayment, including total debt amount and whether a borrower has a mortgage. I estimated this relationship separately for IDR borrowers and borrowers in standard loan repayment after matching, and then I compared the results of these analyses.

4. How do the present values of the estimated total amounts repaid differ between IDR borrowers and borrowers in standard repayment?

To answer this question, I calculated the total amount that borrowers will pay on their loans over time, including payments on principle and interest. Because people tend to value the money they currently have more than the money they will have in the future, I adjusted these total amounts to reflect their net present value. I calculated these total amounts separately for IDR borrowers and borrowers in standard loan repayment after matching, and then I compared the results of these analyses.

Research Approach

Causal estimation.

The purpose of this evaluation is to understand whether the IDR program is successfully achieving several of its goals, namely: making it easier to keep up with payments, making public service careers more appealing, and helping borrowers avoid paying a lot more than their peers in the standard repayment program. Evaluations are useful when it is important to measure whether a program is, on average, better than business as usual or better than an alternative program (Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007). Ideally, an evaluation would tell us whether a specific person in a specific program is better off than if that same person had *not* participated in that same program.

Of course, this is impossible. We cannot observe two different outcomes for the same person at the same time. Holland (1986) called this the "fundamental problem of causal inference" (1986, p. 947). To overcome this problem, we can compare people who participated in the program with people who did not participate, and we can attempt to isolate the program's effect by identifying and eliminating as many alternative explanations for the differences in their outcomes that we possibly can. These alternative explanations could include program participants' previous knowledge and experiences, as well as external events that occur during the study. For example, in this study, we know that borrowers' families' financial experiences with financial aid shape borrowers' own views about how to manage financial aid. For borrowers whose families and friends experienced extreme financial shocks related to debt, income-driven options for managing student loans may be appealing (Peek, Mason, & Soldner, 2016). We also

know that the Obama Administration's marketing and policy efforts to expand IDR occurred at the same time as the program saw substantial growth in uptake. Some borrowers' willingness to enroll in IDR may be related to their responses to particular forms of advertising, rather than particular elements of the IDR program. Family experiences and advertising campaigns are alternative explanations for any differences we may find between borrowers in standard loan repayment plans and borrowers in IDR. If we do not take these differences into account, they may be conflated with the outcomes we want to examine.

Accounting for nonrandom, unobservable student characteristics, or the dynamic interaction between financial aid, students' perceptions of aid, and borrowing decisions, is a common challenge in statistical research on student loan behavior (Dowd, 2008; Cellini, 2008). It is not always possible to measure borrowers' motivations for their decisions, nor can we assume that they are the same across different types of borrowers. Assuming that college graduates who already have taken on loan debt would be well-informed, capable of managing their debt, and confident about their career opportunities in comparison to most college students, it is likely that including measures for these variables would inflate estimates of the effectiveness of IDR participation. Such latent factors lead to unobservable heterogeneity between IDR borrowers and non-IDR borrowers, or omitted variable bias (Cellini, 2008). Possible omitted variables in this study include students' access to information about loan repayment programs, their personal or familial experience with student loans, and their assessment of the likelihood that they will succeed in their chosen career. All of these factors could influence borrowers' decisions to enroll in IDR, as IDR borrowers may be better able to navigate

the bureaucracy associated with the student loan system. This navigational capital, or the "skills of maneuvering through social institutions" (Yosso, 2005, p. 80) may be related to parents' education levels, whether or not borrowers had siblings in college, and information about the family members and staff who helped students apply for financial aid, all of these indicating the supporting relationships inside and outside of colleges that help students learn about postsecondary systems (Teranishi & Briscoe, 2006). One example of the relationship between navigational capital and student outcomes can be seen in transfer student credential completion rates. Transfer student completion rates are higher than those of first time in college students (Ginder, Kelly-Reid, & Mann 2017), perhaps because transfer students have experience with academic bureaucracy and are able to access the institutional resources they need to achieve their goals. Similarly, IDR borrowers may be more successful repaying their loans because they are familiar with servicing processes and resources.

How can we eliminate alternative explanations in order to address the specific impact of IDR on the outcomes of borrowers who choose to enroll in it? One common method is to conduct a randomized experiment, in which participants are randomly assigned to participate in a program, or "treatment", or a status quo condition, otherwise known as the "control". When using random assignment to place participants in programs, we can assume that any characteristics that might influence participants' outcomes are randomly distributed across participants and non-participants. In the case of this study, we could assume that families' experiences with debt or borrowers' responses to advertising campaigns would be about equally present for IDR borrowers and borrowers in standard loan repayment. This means that these characteristics are less

likely to interfere with our ability to measure the effects of the program (Schneider et al., 2007). Randomized experiments are the gold standard in social science research, because they neatly eliminate many of the challenges of studying real people with complex lives.

It is often impossible to conduct a study using random assignment due to logistical or ethical barriers. In this study, it would be infeasible to assign borrowers to a repayment plan and observe their behavior over many years. Borrowers have the option to choose among several different loan repayment programs as long as they qualify for them. It would be impossible, as well as unfair, to limit borrowers to only IDR or only non-IDR plans. Instead, this study uses a design called propensity score analysis.

Propensity score analysis.

Propensity score analysis is a type of quasi-experimental design that approximates random assignment by measuring the likelihood that participants will choose a treatment and then using this measurement to create similar treatment and control groups. Logistic regression analysis can estimate treatment effects and find the difference between the predicted outcomes for both groups (Stuart & Rubin, 2007). Propensity score analysis methods are best used in combination with regression analysis so that matching can approximate treatment assignment and regression analysis can estimate a treatment effect (Stuart, 2010). Whereas research that relies on correlation, such as OLS regression analysis, estimates the average treatment effect (ATE) for an entire sample by assuming that treatment is assigned randomly, propensity score analysis also allows us to estimate the average treatment effect for the treated (ATT), a more salient measure for determining program effectiveness (Titus, 2007).

Propensity score analysis addresses two baseline requirements for quasi-experimental designs: the strongly ignorable treatment assignment assumption and the stable unit treatment value assumption. This assumption necessitates that the pretreatment covariates are not affected by treatment assignment (Ho, Imai, King, & Stuart, 2007). In other words, the covariates used to model the treatment assignment must be measured before the treatment is administered, something that can be accomplished by limiting the covariates in the propensity score specification model to only those that were measured prior to participants selecting a treatment option. Propensity score analysis cannot, however, ensure that unobservable characteristics are evenly distributed across treatment and control groups. By assuming that the observed variables that affect the treatment are related to the unobserved variables that affect the treatment, we can use matching to meet the strongly ignorable treatment assignment assumption (Stuart, 2010).

The strongly ignorable treatment assignment assumption specifies that treatment assignment is independent of participants' outcomes. When participants can self-select into a treatment group, this assumption is often violated. The error estimates in a regression model may be correlated, inflating the Type I error rate in a regression analysis. Propensity score analysis minimizes selection bias, and thus helps ensure that treatment assignment is independent of participants' outcomes, by mimicking self-selection prior to estimating the ATE or ATT. The strongly ignorable treatment assignment assumption also requires that treatment and control group members could all theoretically have participated in the treatment. Applied to propensity score analysis, this means that both the treatment and control group should have positive, overlapping propensity score density curves (Stuart & Rubin, 2007).

Another pillar of quasi-experimental research is the stable unit treatment value assumption (SUTVA). The two parts of this assumption are that there is only one treatment assigned to each individual and that there is no interference in treatment assignment between individuals (Stuart & Rubin, 2007). In other words, in order for SUTVA to be true, the treatment assignment mechanism should not influence individual outcomes, and one person's treatment assignment should not affect another's at any time (Ho et al., 2007; Schneider et al., 2007). In many cases, these assumptions do not hold true in education research. First, the treatment assignment mechanism may influence outcomes if individuals respond to the treatment assignment. For example, a student may behave differently if they are told that they received a financial aid award based on need rather than merit. Differences in observed behavior might then be due to the way the treatment was described to the student, rather than due to the funds themselves. Second, treatment assignments may be correlated if individuals in the control group are aware that they are not receiving treatment and behave differently than they otherwise would. For example, members of a control group who know that they did not receive a financial aid award based on merit may be motivated to perform better, thereby influencing their final outcomes. In the case of this study, career preferences might influence borrowers' decisions about loan repayment plans, rather than the other way around. Although it is impossible to completely eliminate these possibilities, I took steps to decrease them in the study design.

Data

This evaluation aims to understand the causal relationships between IDR participation and borrower outcomes in order to determine whether IDR is, on average,

better than the status quo, standard loan repayment. This study uses the National Center for Education Statistics' Baccalaureate and Beyond 2008-2012 dataset (B&B:08/12), which tracks student loan borrowers who completed a bachelor's degree between July 1st, 2007 and June 30th, 2008 until four years after graduation. B&B:08/12 is the second follow-up to the NPSAS 2007-2008 study, which was used to identify a group of bachelor's degree recipients. B&B:08/12 includes the first nationally representative, borrower-level data regarding loan repayment plan participation, and thus, this dataset makes an evaluation of IDR borrower outcomes nationwide possible for the first time. This dataset also contains rich information about borrowers' academic and family backgrounds, which I used to model selection into IDR, as well as information regarding career choice and loan repayment behavior, which I used to model borrowers' outcomes. I limited the sample to respondents who borrowed federal student loans and did not enroll in another postsecondary program after graduation.

Data preparation.

Before using this dataset, I took several steps to clean and prepare the data. First, I merged data from additional sources, including the Census and the Bureau of Labor Statistics. I ensured that the merged variables matched to each observation. Second, I investigated patterns of missing data and found none in my sample. Third, I created new variables necessary to complete the analyses.

I also limited the sample to borrowers who did not go on to further education after finishing college. This ensured that the borrowers in this study are focused on paying their loans and not on paying for additional years of school. I also attempted to limit the analysis to IDR borrowers who qualified as having a partial financial hardship, as defined

in Figure 1 in Chapter One. Because I cannot tell from the dataset whether borrowers were required to meet a partial financial hardship standard to enroll in IDR, I cannot know which variation of IDR borrowers chose. Unfortunately, applying a more conservative standard by limiting the analysis only to borrowers with a partial financial hardship made the propensity score model impossible to run. This suggests that a sizable number of borrowers did not participate in the IDR plan that limited borrowers based on financial need. Thus, the control group measured in this analysis includes both borrowers who would have been eligible for IDR but chose not to participate as well as borrowers who were ineligible based on their financial means.

The final sample contains 2,011 observations, representing a subpopulation sample of 249,397.8 borrowers, 14.53% of whom participated in IDR.

Statistical Models

I estimated the conceptual framework in Figure 4 in two steps. First, I modeled borrower's decisions to repay their loans through IDR, or their propensity to select IDR rather than standard loan repayment, using logistic regression. I estimated the probability that borrower *i* chooses IDR according to the binary logistic regression equation:

$$\mathrm{P}(IDR_i) = \frac{\exp\{\beta_0 + \beta_k x_i\}}{1+\exp\{\beta_0 + \beta_k x_i\}}$$

In this equation, *x* represents *k* number of covariates for *i* individuals, including institutional, demographic, academic, and financial variables, as well as interactions between race, gender, and class variables. I then used the estimated propensity scores for each borrower to create balanced groups, using the pscore package in Stata.

Second, I separately estimated the loan repayment outcomes and career choices for the treatment and control groups (i.e., IDR borrowers and non-IDR borrowers). These

multivariate regression models include additional controls for variables that are related to borrowers' outcomes, but not to their decisions about IDR. The difference in the estimated effects between the treatment and control groups reveals whether IDR effectively changes borrowers' behavior.

Borrowers' decisions to pursue careers in public service were modeled as a binary dependent variable in which 1 represents military, nonprofit, and government employers, and 0 represents all other employers, expressed as:

$$\mathrm{P}(EmployedPublic_i) = \frac{\exp\{\beta_0 + \beta_k x_i\}}{1 + \exp\{\beta_0 + \beta_k x_i\}}$$

In this equation, *x* represents *k* number of covariates for *i* individuals, including covariates representing borrowers' social, political, and institutional contexts as well as their values (Perry, 2000).

Borrowers' loan repayment success, defined as being in repayment as opposed to any other loan status, was explored using a dichotomous variable in which 1 represents active repayment status and 0 represents any other repayment status.

$$\mathrm{P}(Repayment_i) = \frac{\exp\{\beta_0 + \beta_k x_i\}}{1 + \exp\{\beta_0 + \beta_k x_i\}}$$

In this equation, *x* represents *k* number of covariates for *i* individuals, including covariates representing degree completion, institutional characteristics, borrowers' race and ethnicity, age, family structure, parental education level, income, and academic preparation for college (Gross, Cekic, Hossler, & Hillman, 2009; Cunningham & Kienzl, 2011).

Third, the total amounts that borrowers repay over time were estimated using a several models. To begin, I calculated the total payment amount for borrowers in

standard loan repayment. The total payment amount for these borrowers is equal to the total amount of debt they owe, plus interest, discounted by 3% a year (Baum & Johnson, 2016).

To calculate the total payment amount for borrowers in IDR, I first estimated their expected earnings during repayment. I estimated that borrowers' income would grow 4% annually (Delisle & Holt, 2012). I assumed that borrowers would pay 15% of their discretionary income every year and remain continuously employed during repayment. To test the sensitivity of these estimates, I also estimated separate growth rates by gender, race, and ethnicity to determine how labor market discrimination influences total payment amounts for borrowers in IBR.

Methods to Address Bias

Propensity score analysis separates the design and analysis stages of the study (Stuart, 2010). Propensity score analysis, which can be thought of as a form of data preprocessing, decreases model dependence in the analysis stage (Ho et al., 2007). By using propensity scores to create treatment and control groups with similar distributions on a set of covariates, propensity score analysis renders the form and covariates in regression models far less important (Stuart & Rubin, 2007). Propensity score analysis can help convince technical and non-technical readers alike that the study results presented are not simply those that were most favorable to the researcher by demonstrating the extent to which treatment and control groups are comparable based on the observed covariates and are unbiased by selection into treatment (Ho et al., 2007; Rosenbaum & Rubin, 1985).

In addition to providing evidence for causal arguments, propensity score analysis offers several statistical benefits. This method separates the design and analysis stages of the study, because separate models are used to determine how likely a borrower is to choose to enroll in IDR and the relationship between borrowers and other outcomes, such as loan repayment status and career choice. This separation of the models decreases model dependence, or in other words, the possibility that the model changes dramatically depending on its form and covariates (Ho et al., 2007; Stuart, 2010). Furthermore, using propensity score analysis ensures that the average estimated borrower outcomes are not based on a sample of borrowers with vastly different characteristics. By weighting borrowers based on the likelihood that they will choose IDR before estimating each group's outcomes, the outcomes estimates are based on borrowers who are very similar. Focusing on borrowers with substantial overlap in their pre-treatment characteristics, or within the region of common support, reduces the bias of an estimated effect (Ho et al., 2007; Titus, 2007).

Propensity scores allow treatment and control group members to be compared according to multiple characteristics by "compressing the relevant factors into a single score" (Garrido et al., 2014, p. 1703) and balancing the distribution of that score across both groups. Balance, or a similar distribution of a variable across the treatment and control groups, is an important property in propensity score models. If the treatment and control groups are balanced on the propensity score, we can assume they are also balanced on the covariates related to treatment selection (Caliendo & Kopeinig, 2005). When estimating propensity scores using logistic regression, common measures of model fit, such as parameter estimates, are not helpful (Stuart, 2010). This is because propensity

score analysis is not estimation, but rather "nonparametric preprocessing" to decrease selection bias before applying parametric models to the sample (Ho et al., 2007, p. 199). The best way to determine whether propensity analysis has been done correctly is to examine the balance of propensity scores and covariates of the propensity score model across treatment and control groups.

Stuart (2010) notes that because propensity scores summarize multidimensional relationships, it is helpful to conduct multiple tests to determine how well the propensity scores balance out the differences between the treatment and control groups. However, these tests cannot determine whether unmeasured covariates will ultimately bias the treatment effect estimates. Propensity score analysis can only balance measured covariates across treatment and control groups. Unlike randomized experiments, propensity score analysis cannot control for unmeasured, or latent, variables (Rubin, 2002).

By weighting borrowers based on the likelihood that they will choose IDR before estimating each group's outcomes, the outcomes estimates are based on borrowers who are very similar. I did not, however, discard observations outside the region of common support, as is often recommended when using propensity scores for matching (Ho et al., 2007; Stuart, 2010; Stuart & Rubin, 2007; Caliendo & Kopeinig, 2005; Cook, 2009). The goal of propensity score analysis is to reduce selection bias, or the difference between the treatment and control group due to borrowers' ability to choose their repayment plan. However, propensity score analysis does not always accurately deal with interpolation and extrapolation bias. Interpolation bias "results from incorrect adjustment in regions of interpolation," or the region of common support, whereas extrapolation bias "results from

improperly adjusting for controls where data are needed but do not exist" (King & Zeng, 2006, p. 148). Simply ignoring propensity scores outside the region of common support will not always eliminate extrapolation bias, because it is impossible to know for certain whether or not the propensity score is a reliable estimate of borrowers' propensity to select one repayment plan over another. This problem is sometimes referred to as the Propensity Score Tautology. As Ho and colleagues (2007) explain,

The estimated propensity score is a balancing score when we have a consistent estimate of the true propensity score. We know we have a true estimate of the propensity score when matching on the propensity score balances the raw covariates. Of course, once we have balance on the covariates, we are done and do not need to look back. That is, it works when it works, and when it does not work, it does not work (and when it does not work, keep working at it) (p. 219).

To escape the "infinite regress" (King & Zeng, 2006, p. 152) of the Propensity Score Tautology and ensure that I did not influence the outcomes of the analysis by selecting the propensity score specification that best suited my purpose, I did not use propensity scores to drop observations. Instead, I weighted observations by their propensity scores to give more importance to the observations that fall within the region of common support and less to those outside it. The propensity score weighting method, as I describe in Step 4, is not subject to the problem of the Propensity Score Tautology (King & Neilsen, 2016).

Before I could estimate treatment effects, I first had to estimate propensity scores and use them to create similar treatment and control groups. First, I modeled borrowers' propensity to select IDR using logistic regression including covariates that represent

borrowers' academic backgrounds, demographics, parental experiences with education, and institutional characteristics, all measured prior to treatment. I assessed the balance between treatment and control groups using standardized estimates of means, scatterplots of the distribution of propensity scores, and quantile-quantile plots. Second, I weighted borrowers according to their propensity scores to create two similar groups. I describe the weighting method in more detail later in this chapter.

I used Taylor linearized weights to account for the variance in the survey data. B&B data is weighted based on institutional participation in NPSAS:06/07 as well as individual student participation in all three waves of B&B data collection (2007-2008, 2008-2009, and 2011-2012). I included survey weights and strata in the regression model in order to estimate outcomes that are generalizable to the national population of student loan borrowers (Cook, McGuire, Meara, & Zaslavsky, 2009; Stuart & Rubin, 2007). The results of the analysis are thus generalizable to bachelor's degree recipients with student loans nationwide.

Analysis Steps

To answer the remaining research questions, I used propensity score analysis and regression models. I followed the six-step process outlined in Garrido et al. (2014):

1. Choose covariates for the propensity score model
2. Balance propensity scores between treatment and control groups
3. Balance covariates between treatment and control groups
4. Choose a matching or weighting strategy
5. Balance covariates between treatment and control groups after matching
6. Estimate treatment effects

Step 1: Choice of covariates for propensity score model.

I took a liberal approach to selecting covariates for the propensity model. There is little cost, in terms of bias, in including covariates that are not strongly related to treatment assignment, as long as the covariates could not have been affected by the treatment and the control group is much larger than the treatment group (Ho et al., 2007; Stuart 2010). Any variable that could affect the treatment assignment, according to theory and previous research, should be included in the propensity score model as a covariate in order to meet the ignorable treatment assignment assumption (Caliendo & Kopeinig, 2005; Stuart, 2010). In order to minimize bias, I selected a large number of covariates that could have affected borrowers' decisions to select IDR. I did not include any variables that were collected after borrowers selected their repayment plans in order to avoid selecting covariates that could be influenced by the treatment (Caliendo & Kopeinig, 2005; Ho et al., 2007).

The covariates in the final propensity score model represent the variety of factors that might influence whether or not a borrower chooses IDR. Based on the conceptual model in Chapter Two, I selected covariates representing each of the four contexts that shape how borrowers perceive postsecondary education, finance, and their future careers. To represent observable family influences, I selected covariates related to family finances, including parents' income or independents' income, EFC, TRIO program eligibility, and SAT and ACT test scores. I also selected covariates related to how college knowledge is passed down through families, including parents' education level and a composite variable, *LoanHelp*, indicating whether borrowers discussed financial aid

decisions with family or friends, did research about financial aid online, and/or expected anyone other than their spouse to help them repay their loans.

To represent observable community influences on loan repayment decisions, I selected covariates related to geography and civic life. Unfortunately, B&B08/12 does not contain zip codes or any other information that might identify borrowers' locations more precisely than the state in which they lived. Thus, these covariates were mostly measured at the state-level. Community context covariates included Census division, high school type, and urbanicity of borrowers' residences.

To represent institutional influences on loan repayment decisions, I included covariates for institutional sector, distance from the borrower's home, and Carnegie classification. I also included covariates for price and financial aid, such as tuition and fees, student budget less grant aid, whether or not borrowers spoke with a financial aid officer, and cumulative amount borrowed. I included a covariate indicating whether or not borrowers compared lenders, something that institutions influence by creating and disseminating preferred lender lists.

Finally, to measure the influence of policy and social contexts on borrowers' decisions, I included a number of covariates for borrowers' identities, such as gender, age, race, ethnicity, and disability status. Though these demographic characteristics are unique to each borrower, they carry meaning because of the way that policy and society view them. For example, demographic characteristics are given meaning through policies that make it easy for companies to unfairly pay women and people of color for their work and through the social isolation of people with disabilities. Political and social oppression and advantage can be measured in many ways. The model includes interaction terms for

race and gender, as well as interactions between race and socioeconomic status. I defined socioeconomic status based on family income and parents' educational attainment. I also included variables representing the effects of policies on borrowers' environments, such as median family income and unemployment in the state.

Table 2. Summary statistics for propensity score model covariates

	Non-IDR borrowers	IDR borrowers
Highest education level of either parent in 2007-08 (logged)	1.540	1.296
	(0.026)	(0.077)
LoanHelp	1.437	1.706
	(0.044)	(0.094)
Student budget minus federal and state grants as of 2007-08	16,202.500	14,942.320
	(418.307)	(1214.065)
Institution sector and full-time status in 2007-08	3.948	4.147
	(0.057)	(0.103)
Talked with staff about financial decisions as of 2007-08	0.486	0.538
	(0.019)	(0.045)
Cumulative loan amount borrowed for undergraduate through 2007-08 (cubed)	12,700,000,000,000.000	15,700,000,000,000.000
	(895,000,000,000)	(1,880,000,000,000)
Portion of state with associate's degree or higher in 2005	0.346	0.349
	(0.002)	(0.005)
Median household income in state in 2005	56,360.940	56,567.350
	(332.077)	(812.08)
Unemployment rate in state in 2005	5.069	5.112
	(0.044)	(0.074)
SES	1.183	0.947
	(0.048)	(0.086)
Amount of state need-based grants in 2007-08	476.778	299.888
	(41.244)	(66.562)
Income (dependents' parents and independents) in 2006	62,192.950	51,168.400
	(2,237.932)	(6,308.146)
Tuition and fees paid in 2007-08	8,008.623	7,198.041

	(321.394)	(955.195)
Carnegie Classification (2005 basic) for institution in 2007-08	2.230	2.209
	(0.025)	(0.049)
Race*gender	10.381	9.966
	(0.109)	(0.270)
Race/ethnicity (multiple) without foreign students	1.548	1.632
	(0.049)	(0.104)
Gender	1.520	1.614
	(0.019)	(0.042)
Age as of 12/31/07	26.059	27.168
	(0.323)	(0.797)
Job in 2009 related to bachelor's degree major	1.388	1.273
	(0.045)	(0.109)

Step 2: Balance of propensity scores across treatment and control groups.

I estimated the propensity scores for the treatment and control groups using a logistic model in which the outcome was the binary choice of treatment (IDR or any other repayment plan) and the independent variables are based on theory. To examine the balance of propensity scores, I created graphs comparing the frequency of propensity scores between treatment and control groups. I compared the distribution of propensity scores between treatment and control groups to ensure that there is sufficient overlap between the distributions, or a region of common support (Caliendo & Kopeinig, 2005; Garrido et al.; 2014, Stuart, 2010; Stuart & Rubin, 2007).

Figure 4. Distribution of propensity scores by treatment group

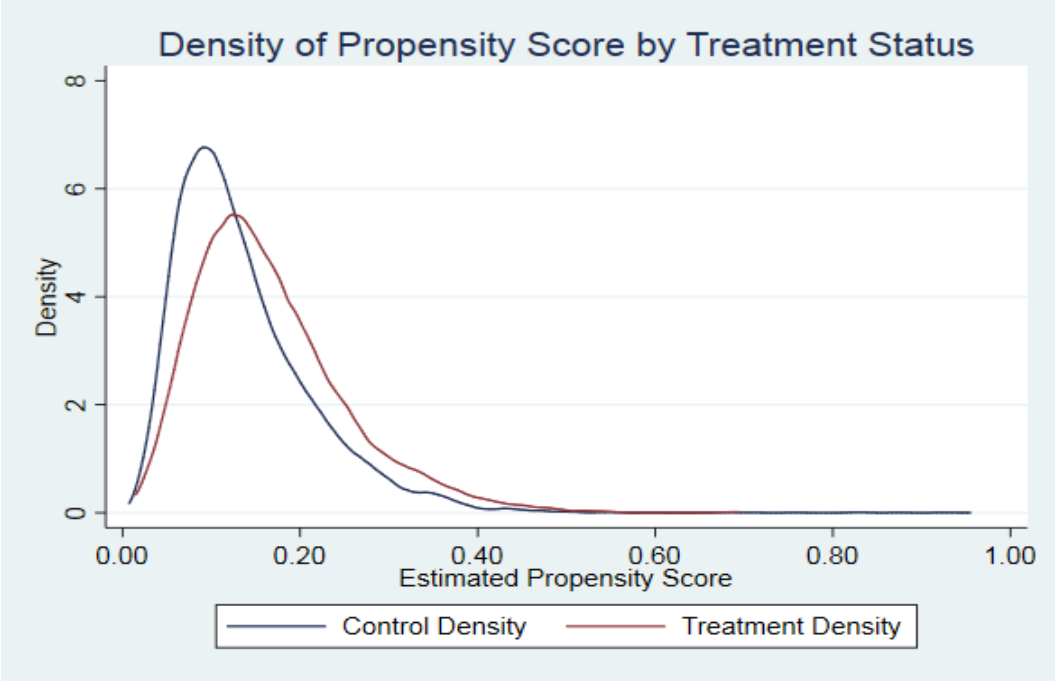


Figure 4 shows similar distributions for the treated and untreated groups, indicating a good deal of overlap at lower propensity score values from 0.05 to 0.40, with

treated observations skewing slightly higher. This is evidence in support of a substantial region of common support. Most observations in the sample fall within this region.

The QQ plot in Figure 5 shows that there is a strong positive relationship between the groups based on the propensity scores. The treated group (labeled treated_phat on the Y-axis) has slightly higher propensity scores than the untreated group. However, the relationship between the treated and untreated groups is positive and approximates a 45-degree angle, as it should be in a well-specified propensity score model (Ho et al., 2007). This is further evidence for the balance of propensity score distributions by treatment status.

Figure 5. QQ plot of propensity scores by treatment status

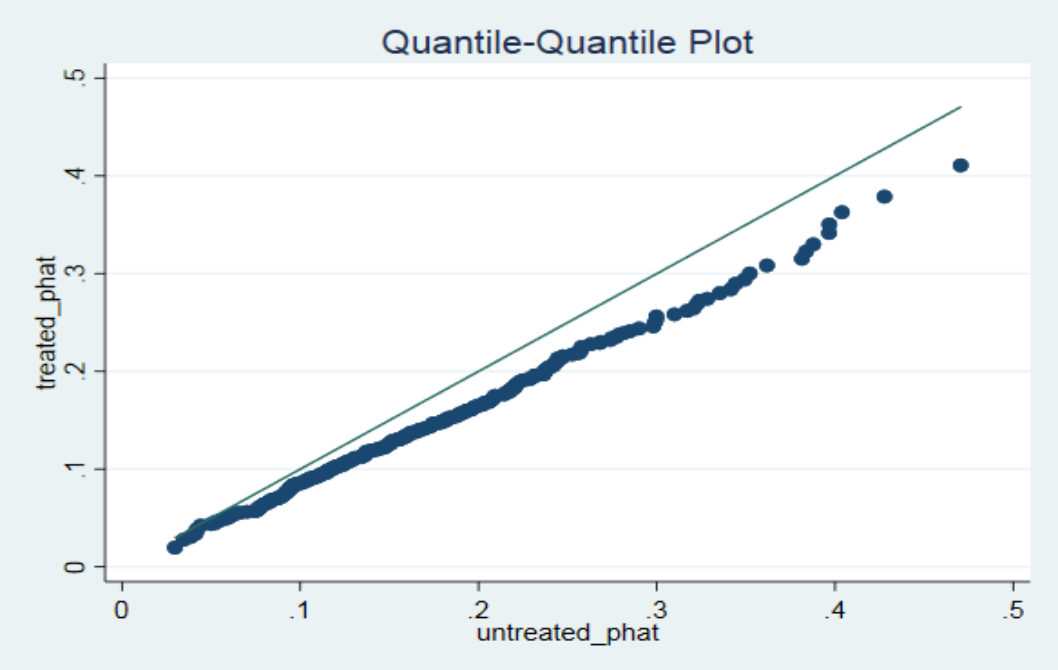


Table 3 below demonstrates that the propensity score distribution of untreated observations is heavily concentrated around zero and has a slightly longer tail, with a small number higher value propensity scores around 0.4.

Table 3. Cumulative distribution of propensity scores by treatment status

	Non-IDR Borrowers			IDR borrowers		
Propensity score	Freq.	Percent	Cum.	Freq.	Percent	Cum.
0.01	4	0.23	0.23	0	0	0
0.02	10	0.57	0.8	1	0.37	0.37
0.03	29	1.67	2.47	3	1.11	1.48
0.04	41	2.36	4.83	5	1.85	3.32
0.05	71	4.08	8.91	5	1.85	5.17
0.06	73	4.2	13.1	5	1.85	7.01
0.07	90	5.17	18.28	10	3.69	10.7
0.08	102	5.86	24.14	11	4.06	14.76
0.09	122	7.01	31.15	18	6.64	21.4
0.1	130	7.47	38.62	13	4.8	26.2
0.11	115	6.61	45.23	18	6.64	32.84
0.12	107	6.15	51.38	16	5.9	38.75
0.13	99	5.69	57.07	16	5.9	44.65
0.14	90	5.17	62.24	11	4.06	48.71
0.15	84	4.83	67.07	15	5.54	54.24
0.16	68	3.91	70.98	10	3.69	57.93
0.17	54	3.1	74.08	10	3.69	61.62
0.18	45	2.59	76.67	10	3.69	65.31
0.19	60	3.45	80.11	10	3.69	69
0.2	58	3.33	83.45	10	3.69	72.69
0.21	48	2.76	86.21	6	2.21	74.91
0.22	37	2.13	88.33	8	2.95	77.86
0.23	37	2.13	90.46	10	3.69	81.55
0.24	21	1.21	91.67	10	3.69	85.24
0.25	24	1.38	93.05	6	2.21	87.45
0.26	17	0.98	94.02	3	1.11	88.56
0.27	17	0.98	95	4	1.48	90.04
0.28	13	0.75	95.75	2	0.74	90.77
0.29	13	0.75	96.49	5	1.85	92.62
0.3	9	0.52	97.01	1	0.37	92.99

0.31	7	0.4	97.41	1	0.37	93.36
0.32	10	0.57	97.99	4	1.48	94.83
0.33	6	0.34	98.33	1	0.37	95.2
0.34	5	0.29	98.62	3	1.11	96.31
0.35	7	0.4	99.02	1	0.37	96.68
0.36	5	0.29	99.31	1	0.37	97.05
0.37	3	0.17	99.48	0	0	97.05
0.38	2	0.11	99.6	3	1.11	98.15
0.39	1	0.06	99.66	2	0.74	98.89
0.4	3	0.17	99.83	1	0.37	99.26
0.41	1	0.06	99.89	0	0	99.26
0.42	0	0	99.89	1	0.37	99.63
0.43	1	0.06	99.94	0	0	99.63
0.44	0	0	99.94	0	0	99.63
0.45	0	0	99.94	0	0	99.63
0.46	1	0.06	100	0	0	99.63
0.47	0	0	100	1	0.37	100
Total	1,740	100		271	100	

Step 3: Balance of covariates across treatment and control groups.

I also examined the standardized means for the covariates by treatment status. Stuart (2010) defines covariate balance as the similarity of the empirical distributions of the full set of covariates in the treated and control groups. The standardized difference in means is equal to "the difference in means of each covariate, divided by the standard deviation in the full treated group" (Stuart, 2010, p. 11). Garrido et al. (2014) recommends comparing the standardized means of covariates within blocks of the propensity score, rather than across the entire distribution. The standard deviation used in these calculations is the same before and after matching.

According to Rubin (2002), the standardized difference in means should not exceed 0.25, while Garrido et al. (2014) notes that recommendations range from 10 to 25 percent. Even in a randomized control trial, some difference in means may be expected,

particularly in small samples (Garrido et al., 2014). One of the goals of propensity score matching is to reduce the observable differences in the treatment and control groups as much as possible so that we can attribute any differences in outcomes to the treatment rather than to systematic differences in the groups themselves. Thus, the difference in means should be as small as possible in order to compare outcomes between the groups.

After comparing covariate means for variables included in the propensity model as well as those that were not, I made some adjustments to the propensity model. Some of these adjustments included adding interaction terms, re-categorizing variables, and adding higher order terms (Garrido et al., 2014; Ho et al., 2007). The standardized means for the final propensity model are shown in Figure 6 in comparison with the means after weighting.

Step 4: Choice of propensity score analysis strategy.

There are many ways to use propensity scores to balance treatment and control groups. I chose to use propensity weights for two reasons. First, as I mentioned in Step 2, the propensity weighting strategy does not have the same problems that propensity score matching strategies do. Propensity score weighting is not subject to the Propensity Score Tautology because the propensity scores are not used in a way that leads to model dependence, researcher influence, or added bias (King & Neilsen, 2016).

The second reason I chose to use propensity weights is that this method allowed me to take advantage of the survey elements in my dataset. Weighting is one of the best matching strategies to use with sample survey data when the goal is to generalize findings to the general population and when estimating the average treatment effect (ATE) or the average treatment effect on the treated (ATT) (DuGoff et al., 2014; Garrido et al., 2014).

By multiplying the propensity weights with the survey weights, I can estimate treatment effects that are generalizable and thus useful beyond this study.

Propensity score weighting uses a weighted average of the control group to compare with the treatment group, thus decreasing variance by including a larger number of observations in the sample (Caliendo & Kopeinig, 2005; Garrido et al, 2014). This means that the control group observations with propensity scores close to the treated group observations are weighted more heavily than the control group observations whose propensity scores are farther away from the treated group observations. Under kernel weighting, which is used to estimate the average treatment effect on the treated (ATT), all treatment group members are weighted 1, and all control group members are weighted $p/(1-p)$, where p represents the propensity score (DuGoff et al, 2014). When estimating the average treatment effect for the entire sample, the control group members are weighted according to the inverse proportion of their distance from the treatment group, but only within a certain range. This is called inverse propensity treatment weighting (IPTW). All treatment group members are weighted $1/p$, and all control group members are weighted $1/(1-p)$, where p represents the propensity score (Garrido et al., 2014; Titus, 2007). In order to incorporate the complex survey design into the estimates, the propensity weights must be multiplied by the survey weights. The composite weight is then used in the regression model to estimate population-level effects (DuGoff et al., 2014; Ridgeway, Kovalchik, Griffin, & Kabeto, 2015). I stabilized the weights by the marginal probability of treatment assignment in order to reduce the variance of the weights (Brookhart et al., 2013). Because there is no best practice using stabilized

weights to find ATTs, I focus on ATEs in this study. As such, the results may slightly overestimate the specific effects of IDR on program participants.

Step 5: Balance of covariates across treatment and control groups after matching.

After weighting, I checked again to make sure that the balance of covariates across treatment and control groups was similar, including covariates that were not selected as part of the propensity score model (Ho et al., 2007). Stuart and Rubin (2007) recommend comparing the mean values of covariates across treatment and comparison groups, standardized by the standard deviation of the full sample, as well as the variances, correlations, and interactions between covariates. They note that the standardized difference in means after weighting should generally be less than 0.25, while the variance ratios should be close to one and fall between 0.5 and 2.0 (Rubin, 2002). As shown in Figure 6, the standardized differences in means improved for most covariates after weighting, though the variance of most covariates does not.

Figure 6. Ratio of standardized mean differences to variance before and after adding stabilized inverse propensity weights

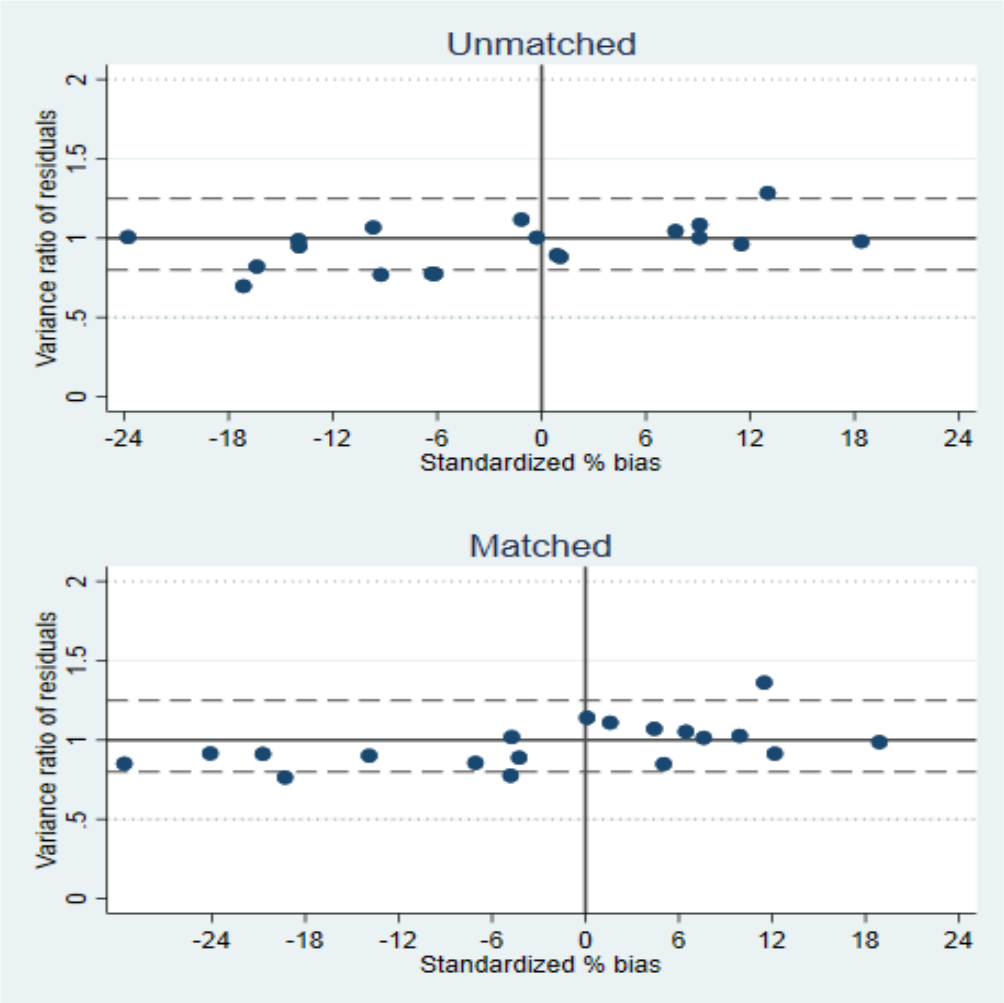
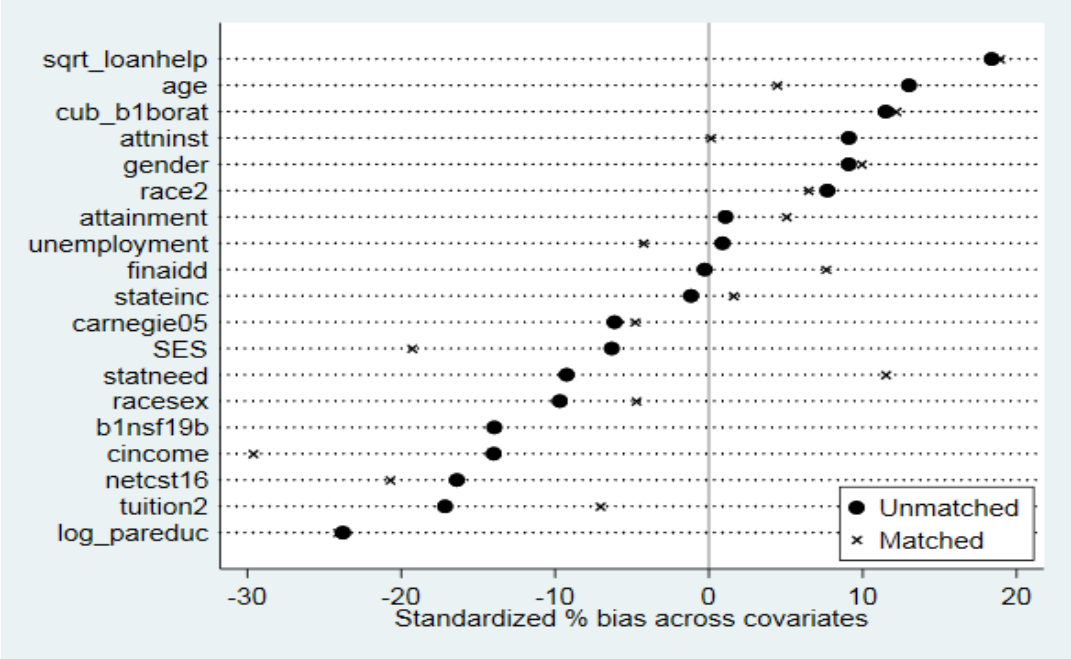


Figure 7. Standardized mean differences before and after adding stabilized inverse propensity weights



Step 6: Estimation of treatment effects.

To answer my research questions about the effects of IDR on borrowers' loan repayment success and public service career choices, I ran logistic regressions separately for the treatment and control group and compared their outcomes. As I mentioned before, I included a composite weight made up of the survey weight and the propensity score weight as a covariate in each of the regression models. The outcome models incorporate the strata and primary sampling unit survey format using the *svy* Stata command. I also included covariates measured after borrowers selected a repayment plan. These additional covariates help explain the influence of early work and life experiences on borrowers' outcomes.

In the regression models examining borrowers' loan repayment success, defined as being in repayment on their debts as opposed to being in any other status, such as default or forbearance, I controlled for borrowers' work and family experiences since graduation, as well as their experiences with loan deferment or forbearance. I also included covariates related to borrowers' satisfaction with their undergraduate education and their reasons for not enrolling in graduate school. I controlled for the same covariates in the regression model examining borrowers' decisions to enter public service careers and added controls for volunteer work, job satisfaction, and employer benefits. In this regression model, public service careers were defined based on the borrower's employer. Borrowers were counted as in public service occupations if their employers were the school where the borrower was enrolled in 2009, a nonprofit organization, a local, state, or federal government, or the military, including civilian employees.

One of the goals of propensity score analysis is to reduce model dependence, or the influence of model specification on the outcome of interest. In regression analysis, researchers can influence the outcome of the analysis simply by trying different combinations of covariates and different model specifications over and over until they find a model that meets their expectations. Propensity score analysis reduces the possibility of model dependence and offers a more objective method of analyzing outcomes because the propensity score model is assessed based on balance instead of measures of model fit. Still, it is important to conduct sensitivity analysis for the regression model used after matching (Ho et al., 2007). I tested the sensitivity of my regressions using Mantel-Haenszel bounds, an extension of Rosenbaum bounds that tests for bias from unobserved variables. Propensity score analysis reduces model dependence

by balancing observed variables, but it cannot account for heterogeneity from variables that are not included in the outcomes model. Mantel-Haenszel bounds measure the size of the impact of unobserved variables by determining the odds that a confounding variable could predict assignment into treatment. The bias is measured in terms of bounded odds ratios (DiPrete & Gangl, 2004). Mantel-Haenszel bounds are intended to test bias in dichotomous outcomes (Alvarado & Muniz, 2018).

To answer my research question about the predicted value of borrowers' total payment amounts, I calculated borrowers' monthly payment amounts over the length of their repayment plan. I began by taking the average monthly loan payment and income level for borrowers in the treatment and control groups, weighted by the propensity score and survey weight. My calculations assume 4% annual income growth (Delisle & Holt, 2012) and a discount rate of 3% (Baum & Johnson, 2016). I added the value of these payments and compared them for borrowers in the treatment and control groups to find the average difference in total payment amounts over time. The results of these calculations are shown in Chapter Four.

Chapter Four

RQ1. What are the characteristics of borrowers who enroll in IDR?

Tables 4 through 8 show the distribution of key demographic, academic, institutional, financial aid, and postgraduate variables by treatment group. Because these statistics represent each group before balancing on the propensity score, statistically significant differences indicate which variables affect selection into IDR. Overall, IDR borrowers are more likely than non-IDR borrowers to be women, Hispanic, or from lower-income families. They are also more likely to hold associate's degrees prior to finishing their bachelor's, have lower GPAs, and major in different fields, including social sciences and education. IDR borrowers are statistically significantly more likely to be unmarried parents after graduation. They are also more likely to pursue careers in education, communication, and non-management positions.

Table 4. Demographic characteristics of survey sample

Weighted sample (WTE000)		Non-IDR borrowers in sample *n*=213,160.3	IDR borrowers in sample *n*=36,238.5	Total sample size *n*=249,397.8
Age at time of BA/BS	18 to 33	86.13%	85.51%	86.04%
Uncorrected χ^2: 49.538, df:3	34 to 49	10.17%	9.47%	10.07%
Design-based	50 to 65	3.69%	4.64%	3.83%
F(2.44, 1167.31)=0.985,				
p=0.387	66 to 80	0.00%	0.38%	0.06%
% female		51.96%	61.44%	53.34%
Uncorrected χ^2: 61.922, df:1				
Design-based				
F(1, 479)=4.223, *p*=0.040				
% Hispanic		8.13%	14.71%	9.09%
Uncorrected χ^2: 89.741, df:1				
Design-based				
F(1, 479)=4.425, *p*=0.036				
Race	Black or African American	8.27%	12.30%	8.85%
Uncorrected χ^2: 78.484, df:6	Asian American	5.32%	6.21%	5.45%
	American Indian or Native			
Design-based	Hawaiian	0.63%	0.54%	0.62%
F(48, 2626.73)=0.915,				
p=0.477	White	81.58%	74.89%	80.61%

Household income in 2006 Uncorrected χ^2: 259.948, df:25 Design-based F(4.33, 2073.68)=3.999, *p*=0.002	$0	0.03%	0.43%	0.09%
	Less than $25,000	30.82%	44.51%	32.81%
	$25,001 to $50,000	20.23%	22.52%	20.56%
	$50,001 to $75,000	17.70%	15.39%	17.37%
	$75,001 to $100,000	14.67%	9.28%	13.88%
	More than $100,001	16.55%	7.86%	15.29%
Highest Parental education level Uncorrected χ^2: 285.682, df:10 Design-based F(8.69, 4164.80)=1.759, *p*=0.074	Less than high school	5.24%	11.01%	6.08%
	High school diploma or equivalent	16.89%	22.82%	17.75%
	Some college, vocational, or trade school	18.58%	17.19%	18.39%
	Associate's degree	6.49%	6.17%	6.44%
	Bachelor's degree	26.11%	21.43%	25.43%
	Master's degree	18.79%	12.48%	17.88%
	Professional degree (e.g., JD, LLB, MD, DDS)	3.67%	2.97%	3.57%
	Doctorate	3.13%	2.41%	3.02%

Census divisions	New England (CT, ME, MA, NH, RI, VT)	4.57%	9.35%	5.26%
Uncorrected χ^2: 225.008, df:8	Middle Atlantic (NJ, NY, PA)	14.29%	9.19%	13.55%
	East North Central (IN, IL, MI,			
Design-based	OH, WI)	16.20%	11.99%	15.59%
F(6.91, 3310.13)=1.318,	West North Central (IA, KS, MN,			
p=0.238	MO, NE, ND, SD)	11.35%	10.76%	11.26%
	South Atlantic (DE, DC, FL, GA,			
	MD, NC, SC, VA)	17.74%	14.06%	17.20%
	East South Central (AL, KY, MS,			
	TN)	5.01%	6.77%	5.27%
	West South Central (AR, LA, OK,			
	TX)	11.01%	12.48%	11.22%
	Mountain (AZ, CO, ID, NM, MT,			
	UT, NV, WY)	7.17%	5.57%	6.94%
	Pacific (AK, CA, HI, OR, WA)	12.66%	19.82%	13.70%

Note: Column totals may not sum to 100 due to rounding or omission of small categories. Design based F statistics are shown to account for the complex survey design of the B&B:08/12. This sample is not weighted according to propensity scores. The sample is weighted using the B&B: 08/12 panel weight, WTE000.

IDR borrowers are more likely to be women (p=0.040), of Hispanic origin (p=0.036), and lower income (p=0.002) than borrowers in other loan repayment programs. There is a nearly 10 percentage point difference in the portion of women borrowers between IDR borrowers (61.44%) and non-IDR borrowers (51.96%). IDR borrowers are almost twice as likely to be Hispanic as non-IDR borrowers, 14.71% and 8.13% respectively. A substantial portion of IDR borrowers, 44.51%, come from families who earned less than $25,000 in 2006, compared to 30.82% of non-IDR borrowers.

On all other demographic variables, borrowers were statistically significantly the same. Borrowers were comparable ages when finishing their bachelor's degrees (p=0.387) and had similar racial backgrounds (p=0.477). A larger portion of IDR borrowers have parents who did not finish high school (11.01%, compared to 5.24% of non-IDR borrowers) or who have a high school diploma or equivalent (22.82%, compared to 16.89% of non-IDR borrowers), but these differences were not significant (p=0.074). There is no statistical difference in geographic location for IDR and non-IDR borrowers.

Table 5. Academic characteristics of survey sample

Weighted sample (WTE000)		Non-IDR borrowers in sample *n*=213,160.3	IDR borrowers in sample *n*=36,238.5	Total sample size *n*=249,397.8
Field of study	Computer and information sciences	6.27%	6.64%	7.16%
Uncorrected χ^2: 350.901, df:9	Engineering and engineering technology	9.90%	8.12%	10.99%
Design-based	Biology and physical sciences	12.63%	10.33%	14.02%
F(8.28, 3968.41)=2.744, *p*=.005	General studies and other	1.79%	2.95%	2.19%
	Social sciences	8.11%	13.65%	9.95%
	Humanities	7.11%	5.17%	7.81%
	Health care fields	6.66%	3.69%	7.16%
	Business	14.07%	14.02%	15.96%
	Education	6.36%	12.18%	8.01%
	Other applied	13.63%	23.25%	16.76%
GPA	1.50 to 1.99	0.34%	0.00%	0.29%
Uncorrected χ^2: 204.945, df:4	2.00 to 2.49	0.63%	8.86%	6.95%
Design-based	2.50 to 2.99	24.26%	37.91%	26.24%
F(3.79, 1803.56)=3.073, *p*=.016	3.00 to 3.49	35.75%	28.27%	34.66%
	3.50 to 3.99	33.02%	24.96%	31.85%

Note: Column totals may not sum to 100 due to rounding or omission of small categories. Design based F statistics are shown to account for the complex survey design of the B&B:08/12. This sample is not weighted according to propensity scores. The sample is weighted using the B&B: 08/12 panel weight, WTE000.

IDR borrowers' academic backgrounds differ in two important ways from non-IDR borrowers. First, IDR borrowers major in different fields from their non-IDR counterparts (p=0.005). In particular, non-IDR borrowers are more likely to major in health care (6.66%) than IDR borrowers (3.69%). Twice as many IDR borrowers majored in education, 12.18%, compared to just 6.36% of non-IDR borrowers. IDR borrowers are more likely to be found in general studies or in other applied majors than non-IDR borrowers.

Second, IDR borrowers had significantly lower GPAs than non-IDR borrowers (p=0.016). 8.86% of IDR borrowers had GPAs between 2.00 and 2.49, compared to just 0.63% of non-IDR borrowers, and 37.91% of IDR borrowers had GPAs between 2.50 and 2.99, compared to 24.26% of non-IDR borrowers. Nearly 69% of non-IDR borrowers had GPAs of 3.00 or above, in contrast to just 53.23% of IDR borrowers.

Table 6. Institutional characteristics of survey sample

Weighted sample (WTE000)		Non-IDR borrowers in sample *n*=213,160.3	IDR borrowers in sample *n*=36,238.5	Total sample size *n*=249,397.8
Carnegie and Control	Public two-year	0.53%	1.05%	0.60%
	Public four-year, doctoral/research extensive	25.71%	20.71%	25.00%
Uncorrected χ^2: 114.417, df:6	Public four-year, all other	37.33%	46.92%	38.68%
Design-based	Private nonprofit four-year, doctoral/research extensive	6.14%	7.47%	6.32%
F(5.38, 2532.93)=0.987, *p*=0.427	Private nonprofit four-year, all other	20.00%	13.56%	19.09%
	For-profit	6.25%	4.97%	6.07%
	Other	4.04%	5.32%	4.22%
Selectivity	Not public or private nonprofit four-year	6.24%	5.60%	6.15%
Uncorrected χ^2: 170.225, df:4	Very selective	24.97%	14.96%	23.52%
Design-based	Moderately selective	51.81%	66.32%	53.92%
F(3.51, 1677.82)=2.349, *p*=0.061	Minimally selective	12.49%	8.10%	11.85%
	Open admission	4.49%	5.02%	4.57%
Net cost (tuition and fees minus grants)	$0 to $9,999	85.78%	87.54%	86.03%
	$10,000 to $19,999	11.19%	6.55%	10.54%
Uncorrected χ^2: 74.831, df:3	$20,000 to $29,999	2.30%	4.31%	2.59%
Design-based	$30,000 to $39,999	0.73%	1.60%	0.85%
F(2.22, 1045.88)=1.249, *p*=0.289				

Note: Column totals may not sum to 100 due to rounding or omission of small categories. Design based F statistics are shown to account for the complex survey design of the B&B:08/12. This sample is not weighted according to propensity scores. The sample is weighted using the B&B: 08/12 panel weight, WTE000.

There are no statistically significant differences in institutional characteristics between IDR and non-IDR borrowers. IDR borrowers are no more likely to attend public or private institutions (p=0.427) or institutions with higher net costs, as measured by tuition and fees minus grants (p=0.289). IDR borrowers are clustered at moderately selective institutions (66.32% compared to 51.81%), but this difference is not statistically significant (p=0.061). While three variables are not sufficient to say that institutional characteristics have no influence over IDR selection, there is no indication in this data that IDR borrowers attend certain kinds of institutions. Similarly, as seen in Figure 13 below, financial aid characteristics have no statistical relationship to IDR selection.

Table 7. Financial aid characteristics of survey sample

Weighted sample (WTE000)		Non-IDR borrowers in sample *n*=213,160.3	IDR borrowers in sample *n*=36,238.5	Total sample size *n*=249,397.8
Sources of financial aid	No aid received	17.50%	16.90%	17.41%
Uncorrected χ^2: 63.403, df:4	Grants only	12.12%	9.69%	11.77%
Design-based	Loans only	14.92%	10.88%	14.34%
F(3.94, 1885.3)=0.903, *p*=0.460	Grants and loans	33.65%	41.92%	34.86%
	Other	21.80%	20.61%	21.63%
Cumulative undergraduate debt	$0 to $9,999	26.30%	17.42%	25.01%
Uncorrected χ^2: 135.374, df:5	$10,000 to $19,999	38.77%	36.15%	38.39%
Design-based	$20,000 to $29,999	20.73%	26.86%	21.62%
F(4.64, 2220.98)=1.704, *p*=0.139	$30,000 to $39,999	9.36%	14.68%	10.13%
	$40,000 to $49,999	4.36%	4.50%	4.38%
	$50,000 to $59,999	0.47%	0.38%	0.46%
Monthly payment as a % of income	0% to 9%	76.60%	68.78%	75.32%
Uncorrected χ^2: 98.785, df:9	10% to 19%	17.68%	23.20%	18.58%
Design-based	20% to 29%	2.32%	3.90%	2.58%
F(5.86, 2653.16)=0.824, *p*=0.549	30% to 39%	1.64%	1.52%	1.62%
	40% to 49%	1.08%	2.48%	1.31%
	50% to 59%	0.04%	0.00%	0.04%
	60% to 69%	0.32%	0.00%	0.27%
	70% to 79%	0.00%	0.00%	0.00%
	80% to 89%	0.26%	0.12%	0.24%
	90% to 99%	0.06%	0.00%	0.05%

Note: Column totals may not sum to 100 due to rounding or omission of small categories. Design based F statistics are shown to account for the complex survey design of the B&B:08/12. This sample is not weighted according to propensity scores. The sample is weighted using the B&B: 08/12 panel weight, WTE000.

There was no statistically significant relationship between IDR selection and multiple financial aid variables, including types of financial aid (p=0.460), cumulative undergraduate debt (p=0.139), and the percentage of a borrower's income used to pay for student loans (p=0.549). IDR borrowers, like non-IDR borrowers, mostly use a combination of grants and loans to pay for college. (Note that although all members of the sample did have student loans, some still responded that they only used grants to pay for college. I verified that all of these borrowers answered affirmatively to other questions about their loan status in all waves of the survey. This calls into question the accuracy of this survey item.) Both types of borrowers are more likely to leave their bachelor's degrees with less than $30,000 in student loan debt. Additionally, the vast majority of both types of borrowers pay less than 20% of their monthly income towards their student loan debt.

Table 8. Post-graduation characteristics of survey sample

Weighted sample (WTE000)		Non-IDR borrowers in sample *n*=213,160.3	IDR borrowers in sample *n*=36,238.5	Total sample size *n*=249,397.8
% meeting partial financial hardship Uncorrected χ^2: 10.855, df:21 Design-based F(1, 479)=0.742, *p*=0.390		42.26%	38.34%	41.69%
% paid through loan forgiveness program Uncorrected χ^2: 49.606, df:1 Design-based F(1, 429)=2.04, *p*=0.154		10.79%	6.29%	9.80%
% job related to major in 2012 Uncorrected χ^2: 69.101, df:2 Design-based F(1.96, 919.00)=1.853, *p*=0.158		46.61%	40.40%	45.70%
Number of months unemployed 2008 to 2012 Uncorrected χ^2: 46.555, df:4 Design-based F(3.67, 1758.89)=1.018, *p*=0.393	0	65.83%	66.55%	65.93%
	1 to 5	20.49%	21.36%	20.61%
	6 to 11	6.33%	3.15%	5.87%
	12 to 23	4.85%	6.93%	5.15%
	More than 24	2.51%	2.02%	2.44%

Occupational category in 2012	Agriculture occupations	0.25%	0.05%	0.22%
Uncorrected χ^2: 670.607, df:32	Air transportation professionals	0.18%	0.00%	0.16%
Design-based	Artists and designers	3.33%	2.94%	3.28%
F(21.55, 10191.32)=1.632, *p*=0.033	Business managers	14.39%	14.84%	14.46%
	Business occupations (non-management)	11.50%	14.77%	11.97%
	Business/legal support (non-secretarial)	11.00%	7.28%	10.47%
	Communication professionals	1.74%	5.60%	2.29%
	Computer/information systems occupation	7.27%	2.92%	6.65%
	Construction/mining occupations	0.67%	0.28%	0.62%
	Engineering technicians	0.79%	0.05%	0.68%
	Engineer	5.17%	0.16%	4.46%
	Fitters, tradesmen, and mechanics	1.85%	4.81%	2.28%
	Food service occupations	2.38%	0.97%	2.18%
	Healthcare professionals (non-nurses)	0.51%	0.85%	0.56%
	Information professionals	0.24%	0.65%	0.30%
	Legal professionals	0.07%	0.56%	0.14%
	Life scientists	0.68%	0.25%	0.62%
	Math-related occupations	0.20%	0.00%	0.17%
	Military	0.81%	0.79%	0.81%
	Nurses	5.13%	2.29%	4.72%
	Other educators	1.99%	3.45%	2.20%
	Other healthcare occupations	2.78%	5.56%	3.18%
	Personal care occupations	2.23%	3.78%	2.45%
	Physical scientists	0.99%	1.16%	1.02%
	PK-12 educators	5.69%	10.19%	6.34%
	Postsecondary educators	1.17%	1.25%	1.18%
	Protective service occupations	2.03%	1.71%	1.99%

	Sales occupations	5.84%	3.89%	5.56%
	Secretaries and administrative assistants	2.69%	1.20%	2.48%
	Social scientists	0.93%	1.68%	1.04%
	Social service professionals	4.03%	5.97%	4.31%
	Sports occupations	0.60%	0.00%	0.51%
	Transport support occupations	0.84%	0.11%	0.74%
Marital status in 2012 Uncorrected χ^2: 138.666, df:3 Design-based F(2.90, 1391.13)=2.862, *p*=0.037	Unmarried, no dependent children	46.56%	44.91%	46.32%
	Unmarried with dependent children	6.03%	13.11%	7.06%
	Married, no dependent children	21.34%	17.16%	20.73%
	Married with dependent children	26.07%	24.82%	25.89%

Note: Column totals may not sum to 100 due to rounding or omission of small categories. Design based F statistics are shown to account for the complex survey design of the B&B:08/12. This sample is not weighted according to propensity scores. The sample is weighted using the B&B: 08/12 panel weight, WTE000.

IDR borrowers differ from standard repayment plan borrowers in several ways after graduation. First, IDR borrowers are statistically significantly more likely to pursue different occupations after graduation (p=0.033). 10.19% of IDR borrowers are PK-12 educators, compared with 5.69% of standard plan borrowers. IDR borrowers are also more likely to be fitters, tradesmen, and mechanics (4.81%) than standard plan borrowers (1.85%), and they are more likely to be non-management business professionals (14.77%) than borrowers in other plans (11.50%). IDR borrowers are more likely to work in communications (5.60%) than borrowers in other plans (1.74%). In contrast, standard plan borrowers are more likely to be engineers (5.17%) than IDR borrowers (0.16%), and they are more likely to work in non-secretarial business and legal support positions (11.00%) than IDR borrowers (7.28%). Non-IDR borrowers are more likely to work in computer and information systems (7.27%) than IDR borrowers (2.92%).

In addition to occupational differences, IDR borrowers make different choices about starting families after graduation (p=0.037). IDR borrowers are statistically significantly more likely to be unmarried with dependent children (13.11%) than borrowers in other plans (6.03%). Borrowers in other plans are more likely to be married without dependent children (21.34%) than IDR borrowers (17.16%).

There are no statistically significant differences between IDR borrowers and borrowers in other plans in terms of the portion meeting partial financial hardship (p=0.390), the percentage planning to repay their loans through a forgiveness program (p=0.154), the percentage whose jobs were closely related to their college majors (p=0.158), or the number of months spent unemployed since graduation (p=0.393).

RQ2. Are IDR borrowers more likely than borrowers in standard loan repayment to work in public service organizations?

For the purpose of answering this research question, I defined public service careers as careers in nonprofit organizations; local, state, or federal government; the military (including civilian employees); and employment at the school where the student was enrolled. To understand the relationship between IDR and public service careers, I conducted a logistic regression analysis on a weighted sample of borrowers. The sample weights were calculated by multiplying the inverse propensity of treatment weights by the survey sample weights in order to estimate average treatment effects that can be generalized to the survey population (DuGoff et al., 2014; Ridgeway, Kovalchik, Griffin, & Kabeto, 2015).

I ran two models: the first controls only for the covariates included in the propensity score model to control for any remaining variation in the means after weighting, and the second model includes additional covariates measured after repayment plan and theoretically related to public service work. These covariates include major lifestyle choices (whether the borrower is considering graduate school, family status, military service); student loan characteristics (cumulative amount owed, status of latest federal loan, cumulative amount consolidated); and work characteristics (working for pay, employer offering benefits, number of jobs, months unemployed, occupation, full-time or part-time status, hours worked per week, satisfaction with work-life balance, satisfaction with importance of work).

The regression models demonstrate that there is no statistically significant difference in public service outcomes for treatment and control groups. IDR selection

was associated with slightly increased odds of working in public service careers in Model 2, but the relationship was not significant (OR=1.093, SE=0.230, p=0.672).

A Hosmer and Lemeshow goodness-of-fit test showed no significant differences between the predicted values in Model 2 (p=0.350) and the observed values in the dataset, although the test for Model 1 was statistically significant (p=0.01). This means that Model 2 is a better model of the data as observed in the sample.

A Mantel-Haenzel sensitivity analysis demonstrates the magnitude of an unobserved confounder's effect on selection into IDR that would change the significance of the average treatment effect (Alvarado & Muniz, 2018). In the case of public service careers, an unobserved covariate would have to increase the odds of selecting IDR by at least 1.4 in order to change the effect estimate. Though this test does not indicate which unobserved covariates might influence treatment selection, it does suggest that the outcomes shown below are somewhat sensitive to unobserved factors.

Table 9. Mantel-Haenzel sensitivity analysis for public service outcome model

Assumed Bias Level	p-value of treatment effect estimate
1	0.443
1.1	0.349
1.2	0.158
1.3	0.058
1.4	**0.018**
1.5	**0.005**
1.6	**0.001**
1.7	**0.000**
1.8	**0.000**
1.9	**0.000**
2	**0.000**

Table 10. Public service career outcome models

	Model 1		Model 2	
	Odds Ratio	Std Error	Odds Ratio	Std Error
Treatment	1.106	0.202	1.093	0.230
Parents' highest education level (logged)	1.034	0.178	1.123	0.202
Family and staff support	1.043	0.091	1.057	0.099
Student budget minus federal and state grants	1.000	0.000	1.000	0.000
Institution sector and full-time status	1.004	0.078	1.045	0.089
Talked with counselor or financial aid staff	0.885	0.168	0.801	0.164
Cumulative undergraduate loan debt (cubed)	1.000	0.000	1.000	0.000
Portion of state with associate's degree or higher (2005)	0.024	0.099	0.029	0.137
State median income (2005)	1.000	0.000	1.000	0.000
State unemployment rate (2005)	0.855	0.112	0.861	0.120
Parents' highest education level*Family income level	0.914	0.095	0.954	0.112
Amount of state need-based grants received (2007-08)	1.000	0.000	1.000	0.000
Family income (2006)	1.000+	0.000	1.000*	0.000
Tuition and fees paid in 2007-08	1.000	0.000	1.000	0.000
Basic 2005 Carnegie code for institution	1.529*	0.271	1.672*	0.318
Race*gender	1.024	0.042	1.038	0.049
Race/ethnicity	0.918	0.074	0.928	0.089
Gender	1.763*	0.340	1.789*	0.393
Age	1.028+	0.016	1.019	0.017
Job related to major (2009)	1.203	0.142	1.265	0.163
Likelihood of enrolling in a graduate program (2012)	-	-	1.129	0.114
Family status (2012)	-	-	1.113	0.097
Cumulative loan amount owed (2012)	-	-	1.000	0.000
Status of latest federal loan (2012)	-	-	0.935	0.095
Cumulative amount of federal loans consolidated (2012)	-	-	1.000	0.000
Working for pay (2012)	-	-	0.593	0.418
Employer offered any benefits (2012)	-	-	1.388	0.470

Number of jobs (2012)	-	-	1.098	0.547
Months unemployed since graduation (2012)	-	-	1.002	0.021
Occupation (2012)	-	-	1.057**	0.011
Full-time or part-time job status (2012)	-	-	1.001	0.298
Served in the military as of 2012	-	-	2.436	1.121
Hours worked per week in all jobs (2012)	-	-	1.020	0.013
Work-life balance (2012)	-	-	1.064	0.098
Importance of work (2012)	-	-	1.381*	0.144
Constant	0.143	0.201	0.001*	0.004

Note: + p<0.1, *p<0.05, **p<0.001.

RQ3. Are IDR borrowers less likely than borrowers in standard loan repayment to be in repayment?

For the purpose of answering this research question, I defined repayment as a binary variable where 1 represented borrowers whose latest federal loan was in repayment in 2012, and 0 represented all other loan statuses. To understand the relationship between IDR and repayment status, I conducted a logistic regression analysis on a weighted sample of borrowers. The sample weights were calculated by multiplying the inverse propensity of treatment weights by the survey sample weights in order to estimate average treatment effects that can be generalized to the survey population (DuGoff et al., 2014; Ridgeway, Kovalchik, Griffin, & Kabeto, 2015).

I ran two models: the first controls only for the covariates included in the propensity score model to control for any remaining variation in the means after weighting, and the second model includes additional covariates measured after repayment plan and theoretically related to public service work. These covariates include major lifestyle choices (whether the borrower is considering graduate school, family status, military service); student loan characteristics (ratio of loans to salary, cumulative amount consolidated); and work characteristics (number of jobs, hours worked per week). Unlike the public service models, the repayment model also included an indicator for attending a two-year institution, as institution type has been shown to be related to loan repayment success.

The regression models demonstrate that there is a statistically significant difference in repayment outcomes for treatment and control groups. IDR selection was associated with increased odds of loan repayment in Model 2 (OR=2.556, SE=0.532,

$p<0.001$). A Hosmer and Lemeshow goodness-of-fit test showed no significant differences between the predicted values in Model 1 ($p= 0.355$) and Model 2 ($p=0.299$) and the observed values in the dataset. This demonstrates that both models reliably represent the observed outcomes in the sample.

A Mantel-Haenzel sensitivity analysis demonstrates the magnitude of an unobserved confounder's effect on selection into IDR that would change the significance of the average treatment effect (Alvarado & Muniz, 2018). In the case of repayment, an unobserved covariate would have to increase the odds of selecting IDR by at least 1.7 in order to change the effect estimate. Though this test does not indicate which unobserved covariates might influence treatment selection, it does suggest that the outcomes shown below are somewhat sensitive to unobserved factors.

Table 11. Mantel-Haenzel sensitivity analysis for repayment model

Assumed Bias Level	p-value of treatment effect estimate
1	0.000
1.1	0.000
1.2	0.000
1.3	0.000
1.4	0.002
1.5	0.008
1.6	0.026
1.7	**0.063**
1.8	**0.127**
1.9	**0.220**
2	**0.337**
2.1	**0.466**
2.2	**0.466**
2.3	**0.349**
2.4	**0.249**
2.5	**0.170**
2.6	**0.111**
2.7	**0.069**

2.8	0.042
2.9	0.024
3	0.014

To further test the sensitivity of this result, I ran both models with an alternative definition of repayment. In this analysis, repayment was defined as a binary variable where 1 represented borrowers whose latest federal loan was either in repayment in 2012, in-school or 6-month grace deferment period, paid in full, or borrowers whose loans had been canceled due to death, disability, bankruptcy, or closed school; and 0 represented all other loan statuses. A Hosmer and Lemeshow goodness-of-fit test showed no significant differences between the predicted values in Model 1 (p= 0.213) and Model 2 (p=0.499) and the observed values in the dataset. This demonstrates that both models reliably represent the observed outcomes in the sample. In both models, there was no statistically significant relationship between IDR and repayment. Thus, the relationship between repayment and IDR, shown in Table 12, should be considered only in the context of a narrow definition of repayment.

Table 12. Student loan repayment outcome models

	Model 1		Model 2	
	Odds Ratio	Std Error	Odds Ratio	Std Error
Treatment	2.122**	0.357	2.064**	0.367
Parents' highest education level (logged)	0.982	0.136	0.959	0.138
Family and staff support	1.152+	0.085	1.135+	0.087
Student budget minus federal and state grants	1.000	0.000	1.000	0.000
Institution sector and full-time status	0.935	0.060	0.920	0.062
Talked with counselor or financial aid staff	1.050	0.167	0.982	0.163
Cumulative undergraduate loan debt (cubed)	1.000*	0.000	1.000	0.000
Portion of state with associate's degree or higher (2005)	0.160	0.552	0.551	1.986
State median income (2005)	1.000	0.000	1.000	0.000
State unemployment rate (2005)	1.115	0.120	1.062	0.120
Parents' highest education level*Family income level	1.093	0.094	1.080	0.095
Amount of state need-based grants received (2007-08)	1.000	0.000	1.000	0.000
Family income (2006)	1.000	0.000	1.000	0.000
Tuition and fees paid in 2007-08	1.000	0.000	1.000	0.000
Basic 2005 Carnegie code for institution	1.183	0.179	1.179	0.190
Race*gender	1.043	0.035	1.046	0.037
Race/ethnicity	0.940	0.060	0.943	0.064
Gender	0.713*	0.117	0.684*	0.117
Age	1.005	0.013	1.017	0.014
Job related to major (2009)	0.962	0.066	0.969	0.070
Likelihood of enrolling in a graduate program (2012)	-	-	0.837*	0.070
Did not attend a two year institution as of 2009	-	-	1.044	0.191
Number of dependent children (2012)	-	-	0.880	0.081
Living with children or dependents (2012)	-	-	1.339	0.262
Ratio of federal loans to annual salary (2012)	-	-	1.019**	0.003
Amount of federal loans consolidated (2012)	-	-	1.000	0.000
Number of jobs (2012)	-	-	1.180	0.358

Served in the military as of 2012	-	-	0.760	0.364
Hours worked per week (2012)	-	-	1.027*	0.009
Constant	0.292	0.361	0.067+	0.094

Note: + p<0.1, *p<0.05, **p<0.001.

RQ4. How do the present values of the estimated total amounts repaid differ between IDR borrowers and borrowers in other repayment plans?

To answer this research question, I calculated how much borrowers in IDR and in other payment plans would likely pay over time. To understand how total payments might differ for IDR borrowers and non-IDR borrowers, I did a simplified calculation of total payment amounts: I estimated an income stream for each IDR borrower using their early career wages as a starting point. Then I calculated 15% of discretionary income for each year based on my simulated wages. The sum of these payments over 25 years, discounted to the present value, represents the total estimated payment amount for IDR borrowers. For non-IDR borrowers, I simply added a 6.8% interest rate to the total amount the borrower owed and discounted the payments. Because interest is not capitalized for IBR borrowers if the borrower remains continuously enrolled in the plan, and I assumed that all IDR borrowers remained continuously enrolled, interest would not be relevant for borrowers who pay for the full 25 years of the amortization period. I did not incorporate loan forgiveness into this repayment model because I assumed that the total costs incurred from the borrowers' perspectives would include only whatever payments they made during the amortization period. Loan forgiveness is a cost to the government, not to the borrower. Finally, inflation was not included in these calculations. All dollar amounts are shown in 2012 dollars.

Comparing these two groups, IDR borrowers' predicted total payment amounts were more varied and also higher than non-IDR borrowers. I found that IDR borrowers are statistically significantly more likely to pay more than non-IDR borrowers ($p<0.001$), as are borrowers of color ($p=0.005$). 56.27% of non-IDR borrowers will make total

payments between $0 and $10,000, compared to just 5.03% of IDR borrowers. Among IDR borrowers, borrowers of color pay more than White borrowers. The most common total payment amounts by race were -$89,999 to -$80,000 for White students, -$49,999 to -$40,000 for Black or African American students, -$29,999 to -$20,000 for Hispanic students, and $20,001 to $30,000 for Asian American students. There were no statistically significant differences in payment amounts between men and women, nor between Hispanic and non-Hispanic students.

Table 13. Net present value of estimated total payments by treatment group

IPTW weighted sample (WTE000*phat)		Non-IDR borrowers	IDR borrowers	Total sample
NPV total expected payments	-$100,000 to -$90,000	0.00%	7.55%	0.01%
Uncorrected *X^2*: 1.06E+04, df:19	-$89,999 to -$80,000	0.00%	10.06%	0.01%
Design-based	-$79,999 to -$70,000	0.00%	8.18%	0.01%
F(12.37, 2.0e+05)=4037.069, *p*<0.001	-$69,999 to -$60,000	0.00%	5.66%	0.01%
	-$59,999 to $50,000	0.00%	7.55%	0.01%
	-$49,999 to -$40,000	0.00%	6.92%	0.01%
	-$39,999 to -$30,000	0.00%	4.40%	0.00%
	-$29,999 to -$20,000	0.00%	5.66%	0.01%
	-$19,999 to -$10,000	0.00%	2.52%	0.00%
	-$9,999 to $0	0.00%	5.66%	0.01%
	$1 to $10,000	56.27%	5.03%	56.22%
	$10,001 to $20,000	21.21%	6.29%	21.19%
	$20,001 to $30,000	9.07%	3.14%	9.06%
	$30,001 to $40,000	5.50%	4.40%	5.50%
	$40,001 to $50,000	3.76%	3.77%	3.76%
	$50,001 to $60,000	1.69%	1.26%	1.69%
	$60,001 to $70,000	1.53%	3.14%	1.53%
	$70,001 to $80,000	0.60%	2.52%	0.60%
	$80,001 to $90,000	0.06%	3.14%	0.07%
	$90,001 to $100,000	0.31%	3.14%	0.32%

Table 14. Net present value of estimated total payments by gender, IDR borrowers only

IPTW weighted sample (WTE000*phat)		Male	Female	Total sample
NPV total expected payments	-$100,000 to -$90,000	9.86%	5.68%	7.55%
Uncorrected *X*^2: 24.239, df:19	-$89,999 to -$80,000	9.86%	10.23%	10.06%
Design-based	-$79,999 to -$70,000	11.27%	5.68%	8.18%
F(19.00, 3001.85)=1.270, *p*=0.192	-$69,999 to -$60,000	4.23%	6.82%	5.66%
	-$59,999 to $50,000	8.45%	6.82%	7.55%
	-$49,999 to -$40,000	8.45%	5.68%	6.92%
	-$39,999 to -$30,000	7.04%	2.27%	4.40%
	-$29,999 to -$20,000	9.86%	2.27%	5.66%
	-$19,999 to -$10,000	1.41%	3.41%	2.52%
	-$9,999 to $0	7.04%	4.55%	5.66%
	$1 to $10,000	5.63%	4.55%	5.03%
	$10,001 to $20,000	2.82%	9.09%	6.29%
	$20,001 to $30,000	4.23%	2.27%	3.14%
	$30,001 to $40,000	1.41%	6.82%	4.40%
	$40,001 to $50,000	1.41%	5.68%	3.77%
	$50,001 to $60,000	0.00%	2.27%	1.26%
	$60,001 to $70,000	1.41%	4.55%	3.14%
	$70,001 to $80,000	0.00%	4.55%	2.52%
	$80,001 to $90,000	2.82%	3.41%	3.14%
	$90,001 to $100,000	2.82%	3.41%	3.14%

Table 15. Net present value of estimated total payments by ethnicity, IDR borrowers only

IPTW weighted sample (WTE000*phat)		Not Hispanic or Latino	Hispanic or Latino	Total sample
NPV total expected payments	-$100,000 to -$90,000	8.03%	4.55%	7.55%
Uncorrected *X^2*: 15.105, df:19	-$89,999 to -$80,000	10.95%	4.55%	10.06%
Design-based	-$79,999 to -$70,000	8.03%	9.09%	8.18%
F(19.00, 3001.26)=0.793, *p*=0.718	-$69,999 to -$60,000	5.11%	9.09%	5.66%
	-$59,999 to $50,000	8.76%	0.00%	7.55%
	-$49,999 to -$40,000	7.30%	4.55%	6.92%
	-$39,999 to -$30,000	4.38%	4.55%	4.40%
	-$29,999 to -$20,000	4.38%	13.64%	5.66%
	-$19,999 to -$10,000	2.19%	4.55%	2.52%
	-$9,999 to $0	6.57%	0.00%	5.66%
	$1 to $10,000	5.11%	4.55%	5.03%
	$10,001 to $20,000	5.84%	9.09%	6.29%
	$20,001 to $30,000	3.65%	0.00%	3.14%
	$30,001 to $40,000	3.65%	9.09%	4.40%
	$40,001 to $50,000	3.65%	4.55%	3.77%
	$50,001 to $60,000	1.46%	0.00%	1.26%
	$60,001 to $70,000	2.19%	9.09%	3.14%
	$70,001 to $80,000	2.92%	0.00%	2.52%
	$80,001 to $90,000	2.92%	4.55%	3.14%
	$90,001 to $100,000	2.92%	4.55%	3.14%

Table 16. Net present value of estimated total payments by race and ethnicity, IDR borrowers only

IPTW weighted sample (WTE000*phat)		White	Black or African American	Hispanic	Asian American	Other	Total sample
NPV total expected payments	-$100,000 to -$90,000	10.00%	4.55%	4.55%	0.00%	0.00%	7.55%
Uncorrected *X^2*: 90.042, df:76	-$89,999 to -$80,000	11.00%	9.09%	4.55%	11.11%	16.67%	10.06%
Design-based F(57.87, 9143.620)=1.550, *p*=0.005	-$79,999 to -$70,000	9.00%	0.00%	9.09%	22.22%	0.00%	8.18%
	-$69,999 to -$60,000	3.00%	13.64%	9.09%	11.11%	0.00%	5.66%
	-$59,999 to $50,000	9.00%	4.55%	0.00%	11.11%	16.67%	7.55%
	-$49,999 to -$40,000	6.00%	18.18%	4.55%	0.00%	0.00%	6.92%
	-$39,999 to -$30,000	6.00%	0.00%	4.55%	0.00%	0.00%	4.40%
	-$29,999 to -$20,000	5.00%	0.00%	13.64%	0.00%	16.67%	5.66%
	-$19,999 to -$10,000	3.00%	0.00%	4.55%	0.00%	0.00%	2.52%
	-$9,999 to $0	5.00%	9.09%	0.00%	11.11%	16.67%	5.66%
	$1 to $10,000	7.00%	0.00%	4.55%	0.00%	0.00%	5.03%
	$10,001 to $20,000	6.00%	9.09%	9.09%	0.00%	0.00%	6.29%
	$20,001 to $30,000	1.00%	4.55%	0.00%	33.33%	0.00%	3.14%
	$30,001 to $40,000	4.00%	4.55%	9.09%	0.00%	0.00%	4.40%
	$40,001 to $50,000	4.00%	4.55%	4.55%	0.00%	0.00%	3.77%
	$50,001 to $60,000	1.00%	4.55%	0.00%	0.00%	0.00%	1.26%
	$60,001 to $70,000	1.00%	9.09%	9.09%	0.00%	0.00%	3.14%
	$70,001 to $80,000	4.00%	0.00%	0.00%	0.00%	0.00%	2.52%
	$80,001 to $90,000	3.00%	0.00%	4.55%	0.00%	16.67%	3.14%
	$90,001 to $100,000	2.00%	4.55%	4.55%	0.00%	16.67%	3.14%

Summary of Results

The key findings for this study are as follows:

Differences between IDR borrowers and borrowers in other repayment plans

- IDR borrowers are more likely to be women, of Hispanic origin, and lower income than their peers in other loan repayment programs.
- Borrowers in general studies, education, and other majors, as well as borrowers with lower GPAs are statistically significantly more likely to select IDR over other repayment plans.
- IDR borrowers are no more likely to attend public or private institutions, more selective institutions, or institutions with higher net costs, as measured by tuition and fees minus grants.
- There is no statistically significant relationship between IDR selection and types of financial aid, cumulative undergraduate debt, or the percentage of a borrower's income used to pay for student loans.
- IDR borrowers are more likely to work in non-management business occupations, communications, PK-12 education, and work as fitters, tradesmen, and mechanics. Non-IDR borrowers are more likely to work in non-secretarial business and legal support positions, computer or information systems, and as nurses or engineers.
- IDR borrowers are statistically significantly more likely to be unmarried with dependent children than borrowers in other plans. Borrowers in other plans are more likely to be married without dependent children than IDR borrowers.

Relationship between IDR and public service careers

- There is no statistically significant difference in mean public service outcomes for IDR borrowers and borrowers in other repayment plans.

Relationship between IDR and loan repayment success

- IDR borrowers are statistically significantly more likely to be repaying their loans than borrowers in other plans.

Comparison of future estimated payments

- IDR borrowers are statistically significantly more likely to pay more than non-IDR borrowers, as are students of color.

Chapter Five

In this study, I used data from the Baccalaureate and Beyond (08:12) longitudinal study to model differences in outcomes for borrowers in IDR and in other loan repayment plans. After weighting each group by propensity to choose IDR, I found that IDR borrowers are statistically significantly more likely to be in repayment, but they are no more likely than other borrowers to pursue careers in public service. IDR is more costly than other repayment plans, especially for borrowers of color. These results indicate that IDR is achieving at least one of its intended goals: serving as a safety net that helps borrowers successfully repay their debts. Researchers and policymakers can use these results to explore options for better aligning Public Service Loan Forgiveness (PSLF) and IDR. This study can also inform discussions on the equitable distribution of costs and subsidies in IDR plans.

Implications

The descriptive data presented in research question one indicate that IDR borrowers are a unique group, comprised disproportionately of women, borrowers of Hispanic origin, and low-income students. These students may have chosen IDR in anticipation of below average earnings due to labor market discrimination. Because women, Hispanics, and low-income people do not have the economic, social, or cultural capital of their peers who are men, White, or high-income, they may choose repayment plans that will protect them against worse labor market outcomes that would influence loan repayment. In other ways, students who expect to fare worse on the labor market make educational decisions that hedge against risk. Hillman (2013) notes that high-achieving, low-income students tend to undermatch, or attend less selective institutions

than they qualify for, due to loan aversion. This could be due to fear about their ability to repay their debts, in spite of their apparent qualifications for the labor market. The ways that IDR borrowers and loan averse, undermatching students react to anticipated labor market difficulties indicate that information about labor market outcomes is available in some form, either anecdotal or data-based. With this information, students are making decisions with long-term effects on their financial well being.

Not only are IDR borrowers from different backgrounds, they also have different educational experiences. IDR borrowers were statistically significantly more likely to have lower GPAs than non-IDR borrowers, as well as major in social sciences and education. Borrowers with lower expected earnings due to their majors and academic achievement, as measured by GPA, may logically select a repayment plan that would protect them against worse labor market opportunities. Similar to the demographic characteristics that are related to IDR selection, GPA and anticipated differences in earnings by major seem to encourage borrowers to hedge against the risk of future difficulties repaying their student loans.

These data demonstrate that the characteristics of postsecondary institutions themselves are not related to IDR selection. IDR selection is also unrelated to loan amounts, which are driven by institutionally-determined costs of attendance. The lack of difference between IDR borrowers and non-IDR borrowers' institutional characteristics indicates that IDR borrowers are not concentrated in low-selectivity, low-cost institutions, as might be expected if IDR borrowers were clustered at institutions with less desirable post-graduation opportunities. Instead, borrowers with lower expected earnings across all kinds of institutions respond to anticipated financial difficulties by selecting

IDR. Institutional characteristics may be unimportant because student loan servicers, not financial aid officers, help borrowers select repayment plans.

Interestingly, there is no significant difference in the portion of income that both kinds of borrowers dedicate to student loan payments. This indicates that borrowers experience the burden of loan repayment in similar ways, despite possible differences in expectations of loan repayment difficulty. Since I do not have data on the portion of income that borrowers would dedicated to their loan payments under a different repayment plan, it is impossible to say whether and to what extent IDR helps borrowers manage their payments. It is possible that some borrowers selected IDR because their payments under the standard plan would have been unreasonably burdensome. It is also possible that borrowers selected IDR but ended up paying the amount that they would have under a standard repayment plan, as borrowers in the original Income-Based Repayment program pay the lesser of 15% of their income or the standard 10-year repayment amount.

IDR is also unrelated to the types of financial aid borrowers use, the cumulative amount they borrow, and the ratio of their student loans to their income after graduation. It is surprising that IDR borrowers are not more likely to have higher debt amounts, given that IDR is generally assumed to be more attractive to borrowers with greater debts. This could be because this study does not include borrowers who attended graduate school, who would have larger debt amounts. This could also be an indication that borrowers are not strictly rational in their assessment of the risks of repayment. Houle and Warner (2017) found that borrowers with lower debt amounts were actually more likely than borrowers with larger debts to return home to live with their parents, though that study

included non-completers and this study includes only bachelor's degree graduates. Nonetheless, total amount owed is not a statistically significant driver of IDR uptake. IDR borrowers are using other information to assess the likely risks of repayment.

After graduation, there are significant differences in borrowers' occupational choices: IDR borrowers are more likely to work in non-management business positions, communications, PK-12 education, or in a trade; while non-IDR borrowers are more likely to work in non-secretarial business and legal support positions, computer or information systems, as nurses or as engineers. There is a clear difference in the prestige associated with each group's occupational outcomes. Many IDR borrowers end up in lower status occupations that are generally associated with lower pay. In contrast to the engineers and computer specialists in the non-IDR borrower group, IDR borrowers who are teachers or work in blue collar jobs may anticipate greater fluctuations in their wages and seek to protect themselves against adverse labor market outcomes using IDR. Similarly, IDR borrowers are more likely to be unmarried parents after graduation than non-IDR borrowers. Without the financial support of a second income, these parents may seek to protect themselves against income fluctuations by selecting IDR.

Regression analysis on the weighted data shows that there is not a causal relationship between IDR and public service loan forgiveness. This indicates that IDR borrowers are not responding to the built-in incentive to choose lower paying careers in order to benefit more quickly from loan forgiveness. For federal policy makers and for others who are interested in the financial sustainability of income-driven financing models, like income share agreement providers, this may be interpreted as a positive result insofar as it indicates a lack of moral hazard. Essentially, borrowers are not

choosing IDR because in order to make lower payments. However, if IDR and PSLF are intended to benefit one another, this result indicates that borrowers do not make the connection between the two policies. Either the benefit of enrolling in IDR is not explicit enough or large enough to incent borrowers to change the sector in which they are employed, or loan forgiveness alone is not sufficient to induce borrowers to change their career paths. Given the large number of federal loan forgiveness programs tied to career choice (Hegji, Smole, & Heisler, 2014), policymakers may prefer to assume that the former is the case. Though the evidence on the effectiveness of loan forgiveness programs is mixed, larger financial incentives do appear to be more influential in workers' career choices (Liou & Lawrenz, 2011).

This study found that there is a causal relationship between IDR and loan repayment success. Using propensity score analysis allows us to rule out observable selection bias. In short, IDR is working as intended. Descriptive studies by the U.S. Government Accountability Office (GAO) and ED indicated that IDR likely made borrowers' payments are more manageable, and this study confirms their arguments. GAO found that between 81% and 87% of IDR borrowers who entered repayment between 2010 and 2014 were in active repayment, compared to just 56% of borrowers in standard repayment plans (U.S. Government Accountability Office, 2015). Additionally, at a servicing summit in 2014, ED presented delinquency rates by repayment plan showing that IDR borrowers typically have lower delinquency rates than borrowers in standard repayment plans. IDR delinquency rates ranged from 2.1% to 15.4%, depending on the IDR plan, compared to 14.9% for borrowers in standard repayment plans (U.S. Department of Education, Federal Student Aid, 2014).

Notably, the relationship between repayment and IDR is sensitive to the definition of repayment used. I found a statistically significant relationship between IDR and repayment when repayment was narrowly defined, but no statistically significant relationship between IDR and repayment when repayment was expanded to include borrowers who had paid in full, were in school or in a six month grace deferment period, and who had had their loans canceled due to death, disability, bankruptcy, or a closed school. I selected these additional loan statuses to include in the repayment variable because they are similar to those included in the federal government's loan repayment rate (Janice & Voight, 2016). However, it appears that borrowers in these other loan statuses are meaningfully different from borrowers in repayment. Borrowers who had repaid in full by 2012 and borrowers in in-school deferment, for example, would not be using IDR any longer. (More information would be needed to ascertain how borrowers could be in in-school or in a six month grace deferment if they indicated that they had not enrolled in an additional degree program since their bachelor's degree as of 2012.)

This study found that IDR is more expensive than other repayment plans, because it extends payments over a longer period of time. If borrowers do not end up paying their loans for the full amortization period, then IDR may not be as costly. My calculations assume that borrowers will pay based on their income each year of the full repayment period. Borrowers who chose to defer to go to graduate school or who take time out of the workforce for unpaid caregiving would have different repayment patterns and thus different total payment amounts. More detailed information about borrowers' lives during repayment, as well as many more years of income data, would be necessary to create accurate models of the total costs associated with different repayment plans.

Among IDR borrowers, based on the assumptions described above, I found that borrowers of color pay more than their White peers. This is due to the fact that my calculations assumed a steady increase of income based on borrowers' initial income. Any racial gaps in income were thus assumed to continue over the repayment period.

Contributions to the Literature

Previous research had mixed findings on the relationship between repayment plan choice, repayment success and public service work. My study corroborates the findings in Herbst (2018), which found that IDR borrowers are significantly and substantially more likely to be in repayment than delinquent. Unlike Field (2009) and Chen (2017), I found no significant relationship between repayment plan choice and public service work. Dissimilarities in student populations and data sources may explain some of the differences in these findings.

The present study contributes to the literature on loan repayment by confirming that repayment plans are related to loan repayment success. Herbst (2018) demonstrated that student loan borrowers who chose IDR were 21 percentage points less likely to be delinquent on their payments than borrowers in other payment plans, in addition to being more successful on other financial measures, like credit scores and loan payment amounts. Given that the borrowers in Herbst's study had already been delinquent on their loan payments, this finding is particularly encouraging. Taken together, our studies provide evidence for the theory that IDR reduces the burden of loan payments for borrowers who are struggling, making it easier for them to manage the loan repayment process. IDR allows borrowers to overcome difficulties and get back on track in their payments.

In contrast to previous studies on loan repayment, my research does not find any significant relationship with public service work. Chen (2017) found a small but significant relationship between IDR and public sector career choice. His simulation based on NSLY97 data found that IDR borrowers are one percentage point more likely to pursue public service careers. One critical difference between Chen's study and mine is that NSLY97 does not indicate whether borrowers actually enrolled in income-driven repayment programs, but assumes that all borrowers would have selected IDR if it were available. Field (2009) does have data on IDR enrollment and found much larger effects. Field (2009) found that law students who were randomly assigned to an income-driven loan option were 36 to 45 percent more likely to pursue public service careers. The statistically insignificant finding in my study may point to differences in the populations studied. Public service loan forgiveness was originally designed as an incentive for graduate and professional students, not for undergraduates. As such, the law students in the Field study were the target audience for PSLF.

Contributions to Theory

This study provides some evidence for the risk aversion theory underlying IDR. IDR borrowers have characteristics that suggest they might anticipate worse labor market outcomes, including being women, Hispanic, from lower income families, and having lower GPAs. In response to these risks, they seek out the protections that IDR offers. In their assessment of IDR plans' policy goals, Hillman and Gross (2014) argue that IDR is designed to address both liquidity risks, or the risk of not being able to make monthly payments, and income risks, or the risk of not being able to pay off the total loan. The results of this study provide evidence that IDR is successfully meeting both goals. The

primary beneficiaries of IDR, those who anticipate adverse labor market outcomes, appear to view IDR as a protection against income risks. Once in the process of repaying their loans, IDR helps borrowers protect against liquidity risk and makes it more likely that they will remain in repayment. These risk protections come at a cost for some borrowers, especially borrowers of color.

In contrast to the conceptual model outlined in Chapter 2, I did not find that institutional variables were related to IDR selection or to repayment success. In the course of my analyses, I found there was a statistically significant relationship between borrowers talking to family or friends and choosing IDR, but no relationship for borrowers talking to counselors or financial aid office staff. This lends support for the theory that family and friends, not institutions, are key sources of navigational capital. This finding is unsurprising given that nothing is required of institutions beyond exit counseling. This finding points to an opportunity for institutions to do more to make students aware of their loan repayment options.

Supporting the conceptual model in Chapter 2, I found that institutional and demographic variables were related to public service career choice. Carnegie classification, gender, and family income in 2006 were all statistically significantly related to public service.

Limitations

There are several notable limitations to the present study. First, the population of borrowers includes only bachelor's degree graduates and cannot be generalized to other borrowers. Non-completion is a strong predictor of loan default (Looney & Yannelis, 2015) but was not a factor that I could include in my models. The relationship between

IDR and loan repayment success is likely different for non-completers. Furthermore, the results in this study do not indicate that IDR borrowers are less likely to default. There is not enough data to investigate default patterns for these borrowers, so I can only tell whether they are more likely to be in repayment than in any other status. It is worth noting that borrowers must contact loan servicers in order to enroll in IDR, which could influence the likelihood of being in repayment. However, borrowers who contact their servicers could also choose to go into deferment or forbearance instead of going into IDR, so navigational capital is not necessarily a confounding factor here. To understand whether or not borrowers are likely to default, it would be necessary to have a larger sample of defaulters, possibly over a longer time frame. There is no reliable data available, to my knowledge, on the timeframe in which default is most likely to occur, though Looney and Yannelis (2015) assert that default is more prevalent in the early years of repayment. It may be that default is more likely more than four years after graduation.

In addition, my sample excluded borrowers who pursued graduate school. As graduate and professional students were the intended beneficiaries for PSLF, this study cannot inform policy debates about whether this program is well targeted for those borrowers. My study can only indicate that PSLF does not appear to change the occupational choices of undergraduate borrowers.

Furthermore, the study did not take unobservable covariates into account. Propensity score analysis can only take into account the observable factors that may have biased the treatment effects. Unobservable factors, such as borrowers' desires to repay their student loan debt more quickly, their desire to pursue a career in a family business,

or financial troubles unrelated to student loan debt might have influenced their willingness to enroll in IDR and thus the outcomes I observed. Further research would need to address this shortcoming through random assignment or latent variable analysis.

Implications for Future Research

To understand how IDR is related to public service careers for graduate and professional student borrowers, and to observe longer-term repayment outcomes, this study would need to be replicated on a different dataset. This B&B study will include one more wave of data collection in 2018. By following up with participants ten years after graduation, B&B will collect rich data on students' long-term outcomes, including their lives after graduate school. With this data, future research can examine repayment patterns for IDR borrowers after graduate school and further into their careers. Using this data, a sequence analysis could show how repayment patterns vary over time. Future work could also determine whether borrowers who complete graduate school are more responsive to the PSLF incentive in IDR than borrowers with bachelor's degrees.

Additionally, more complex simulations of future loan payments could be used to examine how interest capitalization is related to total costs. Borrowers who change repayment statuses and repayment plans have their interest capitalized, or added to the principal owed on their loans. Borrowers who go into deferment for graduate school or into forbearance for financial difficulties, for example, would have their interest added to the amount owed. The same would be true for borrowers who switch into IDR after beginning repayment in another plan or after being automatically enrolled in standard repayment if they no longer qualified for an older IDR plan. This means that total payment amounts are likely to be higher for borrowers who have difficulty continuously

repaying their loans. If these borrowers do not remain enrolled in IDR, their costs may be higher over time. Women borrowers, who are more often responsible for unpaid caregiving at home, may experience more of these fluctuations and thus have higher total payment amounts if they are not in IDR. Future simulations could disaggregate by gender to determine how interest capitalization is related to gender-based differences in labor market participation.

Implications for Policy

This study offers evidence that can inform two proposals that are currently part of the Higher Education Act reauthorization debate: proposals to expand IDR to more borrowers and to eliminate Public Service Loan Forgiveness.

Expanding IDR to more borrowers.

IDR occupies a unique place in higher education policy debates: policy proposals to expand IDR have bipartisan support, although for very different reasons. On the left, advocates argue that IDR will expand access and equity, while on the right, IDR is seen as an efficient means of cost sharing between students and funders.

The descriptive data presented in RQ1 indicates who might respond well to policy proposals to expand IDR: women, Hispanic students, and low-income students, who are the primary beneficiaries currently. Borrowers who have reason to be concerned about adverse labor market outcomes are attracted to the safety net provided by IDR and might support further expansion of these programs. Already, left-leaning advocates have come out in favor of universal or opt-out enrollment in IDR (New America et al., 2014), with the exception of The Institute for College Access and Success (2017), who has expressed concerns about inequitable total costs. Right-leaning advocates have expressed support

for expanding privately-funded IDR in the form of income share agreements, occasionally arguing that they will promote access to higher education for loan averse and underserved students (James, 2016).

If IDR borrowers are more likely to be in repayment, as demonstrated by the present study, then expanding IDR will also help more borrowers avoid default and other negative loan outcomes. This in turn increases the efficiency of the federal loan programs, which will see greater revenues if more borrowers repay through IDR. Conservative policy analysts have suggested that the consequences of an income-driven financing system, though smaller overall, should be shared between students and institutions to further increase efficiency. The American Enterprise Institute has argued that income share agreements, which function similarly to IDR, could be used to more efficiently share the risk of higher education financing between students and institutions to create incentives for better post-graduation outcomes (Kelly, Palacios, & DeSorrento, 2014).

The benefits of expanding IDR are not entirely positive, however. While underserved borrowers may be drawn to these programs, and while these programs may increase the likelihood of repayment success, IDR comes at a cost for some borrowers. Borrowers of color have lower starting salaries, and thus take longer to repay under IDR. As interest accrues over the span of repayment, these borrowers may also end up paying more than their White peers. IDR cannot help borrowers overcome the problem of wage discrimination in the labor market, and it may in fact exacerbate income inequalities. Any policy debate about expanding IDR should consider whether these additional costs are outweighed by the benefit borrowers receive through decreased likelihood of default.

Eliminating Public Service Loan Forgiveness.

Proponents of eliminating PSLF will find evidence to support their position in this study. IDR does not appear to effectively steer undergraduate borrowers to careers in public service. It is possible that graduate student borrowers may benefit from PSLF, but these borrowers were not included in the present study. If policymakers wish undergraduate students to pursue careers in public service, more explicit links to IDR may be necessary. For example, the PSLF and IDR forms could be merged, or ED could expand the number of servicers who offer PSLF.

References

Adnett, N., & Tlupova, D. (2008). Informed choice? The new English student funding system and widening participation. *London Review of Education*, *6*(3), 243–254.

Akers, B., & Chingos, M. M. (2014). *Student loan safety nets : Estimating the costs and benefits of Income-Based Repayment.* Washington, DC: Brown Center on Education Policy at Brookings. Retrieved from https://www.brookings.edu/wp-content/uploads/2016/06/IBR_online.pdf

Alvarado, S.E. & Muniz, P. (2018). Racial and ethnic heterogeneity in the effect of MESA on AP STEM coursework and college STEM major aspirations. *Research in Higher Education.* doi: 10.1007/s11162-018-9493-3.

American Institutes for Research. (2016). *Searching for the best deal: How students and their families view income share agreements*. Washington, DC: American Institutes for Research. Retrieved from http://www.air.org/sites/default/files/downloads/report/Income-Share-Agreements-ISAs-Searching-for-the-Best-Deal-Jan-2016.pdf

Asher, L., Cheng, D., & Thompson, J. (2014, April). *Should all student loan payments be income-driven? Trade-offs and challenges.* Washington, DC: The Institute for College Access and Success. Retrieved from http://ticas.org/sites/default/files/pub_files/TICAS_IDR_White_Paper.pdf

Bärnighausen, T., & Bloom, D. E. (2009). Financial incentives for return of service in underserved areas: a systematic review. *BMC Health Services Research*, *9*(86). https://doi.org/10.1186/1472-6963-9-86

Barr, N. (2004). Higher education funding. *Oxford Review of Economic Policy*, *20*(2), 264–283. doi: 10.1093/oxrep/grh015

Barr, N. (2009). Financing higher education: Lessons from economic theory and reform in England. *Higher Education in Europe*, *34*(2), 201–209. doi: 10.1080/03797720902867419

Baum, S., & Johnson, M.C. (2016). *Strengthening federal student aid: An assessment of proposals for reforming federal student loan repayment and federal education tax benefits.* Washington, DC: The Urban Institute. Retrieved from https://www.urban.org/research/publication/strengthening-federal-student-aid-assessment-proposals-reforming-federal-student-loan-repayment-and-federal-education-tax-benefits/view/full_report

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *The Journal of Political Economy*, *70*(5), 9–49. Retrieved from https://www.journals.uchicago.edu/toc/jpe/current

Becker, G. S. (1993). Nobel lecture: The economic way of looking at behavior. *The Journal of Political Economy*, *101*(3), 385–409. Retrieved from https://www.journals.uchicago.edu/toc/jpe/current

Beer, G., & Chapman, B. (2004). HECS system changes: impact on students, *Agenda, 11*(2), 157–174. Retrieved from https://tandfonline.com/toc/ragn20/current

Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2009). *The FAFSA project : Results from the H & R Block FAFSA experiment and next steps*. (NBER Working Paper Series No. 15361). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w15361

Brookhart, M.A., Wyss, R., Layton, J.B., & Stürmer, T. (2013). Propensity score methods for confounding control in non-experimental research. *Circ Cardiovasc Qual Outcomes, 6*(5). 604-611. doi: 10.1161/CIRCOUTCOMES.113.000359

Boatman, A., Evans, B., & Soliz, A. (2014). Applying the lessons of behavioral economics to improve the federal student loan programs : Six policy recommendations. Washington, DC: Lumina Foundation. Retrieved from https://www.luminafoundation.org/files/publications/ideas_summit/Applying_the_Lessons_of_Behavioral_Economics_to_improve_the_Federal_Policy_Loan_Programs.pdf

Caetano, G., Patrinos, H. A., & Palacios, M. (2011, July). *Measuring aversion to debt: An experiment among student loan candidates*. (Policy Research Working Paper No. 5737.) World Bank: Washington DC. https://doi.org/10.1596/1813-9450-5737

Caliendo, M., & Kopeinig, S. (2005). *Some practical guidance for the implementation of propensity score matching.* (IZ Discussion Papers No. 1588). Retrieved from http://hdl.handle.net/10419/33478

Callender, C. (2002). The costs of widening participation: Contradictions in New Labour's student funding policies. *Social Policy and Society*, *1*(02), 83-94. doi: 10.1017/S1474746402000210

Callender, C. (2003). Attitudes to debt: School leavers and further education students' attitudes to debt and their impact on participation in higher education. London: Universities UK. Retrieved from http://www.universitiesuk.ac.uk/Pages/home.aspx

Canadian Federation of Students. (2012). *Income-contingent loan repayment schemes*. Ottawa, ON: Canadian Federation of Students Ontario. Retrieved from http://cfsontario.ca/

Cellini, S. R. (2008). Causal inference and omitted variable bias in financial aid research: Assessing solutions. *The Review of Higher Education*, *31*(3), 329–354. http://doi.org/10.1353/rhe.2008.0010

Chapman, B. (1997). Conceptional issues and the Australian experience with income-contingent charges for higher education. *The Economic Journal*, *107*(442), 738–751. Retrieved from http://www.res.org.uk/view/economichome.html

Chapman, B. (2006). Chapter 25: Income contingent loans for higher education: International reforms. *Handbook of the Economics of Education*, *2*(06), 1435–1503. http://doi.org/10.1016/S1574-0692(06)02025-3

Chapman, B., & Greenaway, D. (2006). Learning to live with loans? International policy transfer and the funding of higher education. *World Economy*, *29*(8), 1057–1075. http://doi.org/10.1111/j.1467-9701.2006.00822.x

Chapman, B., & Ryan, C. (2005). The access implications of income-contingent charges for higher education: Lessons from Australia. *Economics of Education Review*, *24*(5), 491–512. http://doi.org/10.1016/j.econedurev.2004.08.009

Chen, H. (2017). The impact of student loan repayment reform on schooling, work, and borrowing decisions. Unpublished manuscript, Department of Economics, The Ohio State University, Columbus, OH. Retrieved from https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=NASM2017&paper_id=562

Cook, B. L., McGuire, T. G., Meara, E., & Zaslavsky, A. M. (2009). Adjusting for health status in non-linear models of health care disparities. *Health Services Outcomes Research Methodolology*, *9*(1), 1–21. http://doi.org/10.1007/s10742-008-0039-6

Cunningham, A. F., & Kienzl, G. S. (2011). *Delinquency : The untold story of student loan borrowing*. Washington, DC: Institute for Higher Education Policy. Retrieved from http://www.ihep.org/sites/default/files/uploads/docs/pubs/delinquency-the_untold_story_final_march_2011.pdf

Dearden, L., Fitzsimons, E., Goodman, A., & Kaplan, G. (2008). Higher education funding reforms in England: The distributional effects and the shifting balance of costs. *The Economic Journal*, *118*(February), 100–125. http://doi.org/10.1111/j.1468-0297.2007.02118.x

Debande, O., & Vandenberghe, V. (2007). *Refinancing Europe's higher education through deferred and income-contingent higher education fees: An empirical assessment using Belgian, German & UK data*. (LSE STICERD Research Paper No. 124). London: LSE Suntory and Toyota International Centre for Economics and Related Disciplines. Retrieved from http://sticerd.lse.ac.uk/dps/case/cp/CASEpaper124.pdf

Delaney, J. A., & Dharmapala, D. (2015). *"Pay It Forward" and higher education subsidies.* Unpublished manuscript, College of Education, University of Illinois at Urbana-Champaign, Champaign, IL. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2518248

Delisle, J. (2016, February 10). Another student loan bombshell in Obama's latest budget. [Blog post]. Retrieved from http://www.edcentral.org/another-student-loan-bombshell-obamas-latest-budget/

Delisle, J., & Holt, A. (2012). *Safety net or windfall?: Examining changes to income-based repayment for federal student loans.* Washington, DC: New America. Retrieved from https://static.newamerica.org/attachments/2332-safety-net-or-windfall/NAF_Income_Based_Repayment.18c8a688f03c4c628b6063755ff5dbaa.pdf

DiPrete, T.A., & Gangl, M. (2004). Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and and instrumental variables estimation with imperfect estimates. *Sociological Methodology, 34.* 271-310. Retrieved from http://journals.sagepub.com/home/smxa

Douglas-Gabriel, D. (2014, December 29). There's a way to dramatically lower student debt payments, but hardly anyone uses it. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/get-there/wp/2014/12/29/theres-a-way-to-dramatically-lower-student-debt-payments-but-hardly-anyone-uses-it/

Dowd, A. C. (2008). Dynamic interactions and intersubjectivity: Challenges to causal modeling in studies of college student debt. *Review of Educational Research, 78*(2), 232–259. http://doi.org/10.3102/0034654308317252

DuGoff, E. H., Schuler, M., & Stuart, E. A. (2014). Generalizing observational study results: Applying propensity score methods to complex surveys. *Health Services Research, 49*(1), 284–303. http://doi.org/10.1111/1475-6773.12090

Dynarski, S., & Kreisman, D. (2013). *Loans for educational opportunity: Making borrowing work for today's students*. Washington, DC: Brookings Institution. Retrieved from http://www.hamiltonproject.org/files/downloads_and_links/THP_DynarskiDiscPaper_Final.pdf

Federal Perkins Loan Program, Federal Family Education Loan Program, and William D. Ford Federal Direct Loan Program; Final Rule. 34 C.F.R. § 674, 682, and 685. (2008).

Field, E. (2009). Educational debt burden and career choice: Evidence from a financial aid experiment at NYU Law School. *American Economic Journal: Applied Economics*, *1*(1), 1–21. Retrieved from https://www.aeaweb.org/journals/app

Field, K. (2014, January 24). TurboTax helps spread the word on student-loan repayment options. *The Chronicle of Higher Education*. Retrieved from http://chronicle.com/article/TurboTax-Helps-Spread-the-Word/144163

Flannery, D., & O'Donoghue, C. (2011). Education finance systems in Ireland. *The Economic and Social Review*, *42*(3), 237–270. Retrieved from http://www.esr.ie/

Friedman, M. (1955). The role of government in education. In R. A. Solo (Ed.), *Economics and the Public Interest*. New Brunswick, NJ: Rutgers University Press.

Garrido, M.M., Kelley, A.S., Paris, J., Roza, K., Meier, D.E., Morrison, R.S., & Aldridge, M.D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research*, *49*(5). 1701-1720. doi: 10.1111/1475-6773.12182

Ginder, S.A., Kelly-Reid, J.E., & Mann, F.B. (2017, December). *Graduation rates for selected cohorts: 2008-13; outcome measures for cohort year 2008; student financial aid, academic year 2015-16; and admissions in postsecondary institutions, Fall 2016: First look (provisional data).* U.S. Department of Education, National Center for Education Statistics: Washington, DC. Retrieved from https://nces.ed.gov/pubs2017/2017150rev.pdf

Glazerman, S., & Seftor, N. (2005). *The NIH Intramural Research Loan Repayment program: Career outcomes of participants and nonparticipants*. Washington, DC: Mathematica Policy Research. Retrieved from https://www.mathematica-mpr.com/our-publications-and-findings/publications/the-nih-intramural-research-loan-repayment-program-career-outcomes-of-participants-and-nonparticipants

Goldrick-Rab, S., Harris, D. N., & Trostel, P. A. (2009). Why financial aid matters (or does not) for college success: Toward a new interdisciplinary perspective. In *Higher education: Handbook of theory and research, volume XXIV* (pp. 1–45). http://doi.org/10.1007/978-1-4020-9628-0_1

Gross, J. P. K., Cekic, O., Hossler, D., & Hillman, N. (2009). What matters in student loan default: A review of the research literature. *Journal of Student Financial Aid*, *39*(1), 19–29. http://doi.org/10.2307/302397

Halpert, J. (2014, November 18). 7 ways to stop student loans from ruining your life. *CNBC.com*. Retrieved from http://www.cnbc.com/2014/11/17/7-ways-to-stop-student-loans-from-ruining-your-life.html

Hansen, J. S. (2008). Rethinking student aid: Learning from international experience. In S. Baum, M. Mcpherson, & P. Steele (Eds.), *The effectiveness of student aid*

policies: What the research tells us. (pp. 189-212). Washington, DC: The College Board. Retrieved from http://advocacy.collegeboard.org/sites/default/files/rethinking-stu-aid-effectiveness-of-stu-aid-policies.pdf#page=15

Hegji, A., Smole, D. P., & Heisler, E. J. (2014). *Federal student loan forgiveness and loan repayment programs*. (CRS Report No. R43571). Washington, DC: Congressional Research Service. Retrieved from https://fas.org/sgp/crs/misc/R43571.pdf

Herbst, D. (2018). Liquidity and insurance in student loan contracts: Estimating the effects of income-driven repayment on default and consumption. Unpublished manuscript, Department of Economics, Princeton University, Princeton, NJ. Retrieved from https://drive.google.com/file/d/1A-gq_LIqffY6r2gDTcUK9-Y3ZV8Go6SU/view

Hess, A. (2017, October 27). "How this millenial paid off over $100,000 of debt in 6 years." *CNBC.com.* Retrieved from https://www.cnbc.com/2017/10/26/how-this-millennial-paid-off-over-100000-of-debt-in-6-years.html

Hillman, N. W. (2013). *Reforming repayment : Using income-related loans to reduce default*. Washington, DC: American Enterprise Institute. Retrieved from http://www.aei.org/wp-content/uploads/2013/06/-kelly-hillmannick-conference_085349110877.pdf

Hillman, N. W., & Gross, J. P. K. (2014). *Can income-driven repayment policies be efficient, effective, and equitable?* Washington, DC: Lumina Foundation. Retrieved from

https://www.luminafoundation.org/files/publications/ideas_summit/Can_Income-Driven_Repayment_Policies_Be_Efficient_Effective_and_Equitable.pdf

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*(3), 199–236. http://doi.org/10.1093/pan/mpl013

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*(396), 945–960. Retrieved from http://www.amstat.org/asa/publications/Journals.aspx

Houle, J.N., & Warner, C. (2017). Into the red and back to the nest? Student debt, college completion, and returning to the parental home among young adults. *Sociology of Education, 90*(1). 89-108. http://doi.org/10.1177/0038040716685873

Jacobs, B., & van der Ploeg, F. (2006). Guide to reform of higher education: A European perspective. *Economic Policy*, *21*(47), 535+537–592. Retrieved from https://academic.oup.com/economicpolicy

James, K.J. (2016, July). *New funding, aligned incentives: How private financing options can foster higher education innovation.* Washington, DC: American Enterprise Institute. Retrieved from http://www.aei.org/publication/new-funding-aligned-incentives-how-private-financing-options-can-foster-higher-education-innovation/

Janice, A., & Voight, M. (2016). *Making sense of student loan outcomes: How using repayment rates can improve student success*. Washington, DC: Institute for Higher Education Policy. Retrieved from http://www.ihep.org/sites/default/files/uploads/docs/pubs/making_sense_of_student_loan_outcomes_paper.pdf

Ji, Y. (2017). Job search under debt: Aggregate implications of student loans. Unpublished manuscript, Hong Kong University of Science and Technology. Retreived from https://ssrn.com/abstract=2976040

Johnson, O. (2010). Assessing neighborhood racial segregation and macroeconomic effects in the education of African Americans. *Review of Educational Research, 80*(4). 527-575. http://doi.org/10.3102/0034654310377210

Kelly, A., Palacios, M., & DeSorrento, T. (2014). *Investing in value, sharing in risk: Financing higher education through income share agreements.* Washington, DC: American Enterprise Institute. Retrieved from http://www.aei.org/publication/investing-in-value-sharing-risk-financing-higher-education-through-income-share-agreements/

King, G., & Neilsen, R. (2016). Why propensity scores should not be used for matching. Unpublished manuscript, Institute for Quantiative Social Science, Harvard University, Cambridge, MA. Retrieved from https://gking.harvard.edu/publications/why-propensity-scores-should-not-be-used-formatching

King, G., & Zeng, L. (2006). The dangers of extreme counterfactuals. *Political Analysis, 14*(2). 131-159. Retrieved from https://www.cambridge.org/core/journals/political-analysis

Kobliner, B. (2017, June 29). "7 ways to make paying down your student loans a little easier." *Refinery29.* Retreived from http://www.refinery29.com/2017/06/157539/pay-down-student-loans-beth-kobliner-advice

Kolbe, T., & Strunk, K. O. (2012). Economic incentives as a strategy for responding to Teacher staffing problems: A typology of policies and practices. *Educational Administration Quarterly*, *45*(5), 779–813. http://doi.org/10.1177/0013161X12441011

Liou, P. Y., & Lawrenz, F. (2011). Optimizing teacher preparation loan forgiveness programs: Variables related to perceived influence. *Science Education, 95*(1), 121–144. http://doi.org/10.1002/sce.20409

Looney, A., & Yannelis, C. (2015). *A crisis in student loans? How changes in the characteristics of borrowers and the institutions they attended contributed to rising loan defaults*. Washington, DC: Brookings Institution. Retrieved from https://www.brookings.edu/wp-content/uploads/2016/07/PDFLooneyTextFallBPEA.pdf

Marks, G.N. (2009). The social effects of the Australian Higher Education Contribution Scheme (HECS). *Higher Education, 57.* 71-84. Retrieved from https://link.springer.com/journal/10734

McFadden, D. (1999). Rationality for economists? *Journal of Risk and Uncertainty*, *19*(1), 73–105. http://doi.org/10.1023/A:1007863007855

Mincer, J. A. (1974). Schooling and earnings. In J. A. Mincer (Ed.), *Schooling, Experience, and Earnings* (pp. 41–63). Cambridge, MA: National Bureau of Economic Research. http://doi.org/10.1017/CBO9781107415324.004

Mueller, T. (2013). Changes to the student loan experience: Psychological predictors and outcomes. *Journal of Student Financial Aid*, *43*(3). Retrieved from https://publications.nasfaa.org/jsfa/

Nelson, L. A. (2012, October 23). An underused lifeline. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/news/2012/10/23/despite-student-debt-concern-income-based-repayment-lags#sthash.weRkJm7B.8H2CImut.dpbs.

New America, Young Invincibles, & National Association of Student Financial Aid Administrators. (2014). *The case for payroll withholding: Preventing student loan defaults with automatic Income-Based Repayment*. Reimagining Aid Design and Delivery (RADD) Paper Series. Retrieved from http://www.luminafoundation.org/files/resources/payroll-withholding.pdf

Oosterbeek, H. (1998). An economic analysis of student financial aid schemes. *European Journal of Education*, *33*(1), 21–29. Retrieved from https://onlinelibrary.wiley.com/journal/14653435

Palacios, M. (2003, March). *Options for financing lifelong learning*. (World Bank Policy Research Working Paper No. 2994). Washington, DC: World Bank. http://doi.org/10.1596/1813-9450-2994

Peek, A., Mason, J., & Soldner, M. (2016). *How loan-averse young adults view income share agreements*. Washington, DC: American Institutes for Research. Retrieved from https://www.air.org/resource/how-loan-averse-young-adults-view-income-share-agreements

Perna, L. W. (2006). Understanding the relationship between information about college prices and financial aid and students' college-related behaviors. *American Behavioral Scientist*, *49*(12), 1620–1635. http://doi.org/10.1177/0002764206289144

Perry, J. L. (2000). Bringing society in: Toward a theory of public service motivation. *Journal of Public Administration Research and Theory, 10*(2), 471–488. Retrieved from https://academic.oup.com/jpart

Phillips, R. L., Dodoo, M. S., Petterson, S., Xierali, I., Bazemore, A., Teevan, B., Bennett, K., Legagneur, C., Rudd, J., & Phillips, J. (2009). *Specialty and geographic distribution of the physician workforce: What influences medical student & resident choices?* Washington, DC: The Robert Graham Center. Retrieved from http://www.graham-center.org/content/dam/rgc/documents/publications-reports/monographs-books/Specialty-geography-compressed.pdf

Rasmussen, C. J. (2006). Effective cost-sharing models in higher education: Insights from low-income students in Australian universities. *Higher Education, 51*(1), 1–25. http://doi.org/10.1007/s10734-004-6373-x

Ridgeway, G., Kovalchik, S.A., Griffin, B.A., & Kabeto, M.U. (2015). Propensity score analysis with survey weighted data. *Journal of Causal Inference, 3*(2). 237-249. http://doi.org/10.1515/jci-2014-0039

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39*(1), 33–38. http://doi.org/10.1080/00031305.1985.10479383

Rubin, D.B. (2002). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology, 21.* 169-188. https://doi.org/10.1023/A:1020363010465

Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating Causal Effects: Using Experimental and Observational Designs*. Washington, D.C.: American Educational Research Association. Retrieved from http://www.aera.net/Publications/Books/Estimating-Causal-Effects-Using-Experimental-and-Observational-Designs

Shen, H., & Ziderman, A. (2009). Student loans repayment and recovery: International comparisons. *Higher Education, 57*(3). 315-333. Retrieved from http://www.jstor.org/stable/40269125

Simpson, W. B. (1987). Income-contingent student loans: Context, potential and limits. *Higher Education, 16*, 699–721. Retrieved from https://link.springer.com/journal/10734

Smith, K., Gibson, J., & Burbank, J. (2016). *Policymakers' guide to Pay It Forward*. Seattle, WA: Economic Opportunity Institute. Retrieved from http://www.eoionline.org/wp/wp-content/uploads/Policymakers-Guide-to-Pay-It-Forward.pdf

St. John, E. P., Paulsen, M. B., & Starkey, J. B. (1996). The nexus between college choice and persistence. *Research in Higher Education, 37*(2), 175–220. http://doi.org/10.1007/BF01730115

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*(1), 1–21. http://doi.org/10.1214/09-STS313

Stuart, E. A., & Rubin, D. B. (2007). Matching methods for causal inference: Designing observational studies. *Best Practices in Quantitative Methods*, 608. Retreived from

https://www.researchgate.net/publication/228519896_Matching_methods_for_causal_inference_Designing_observational_studies

Stratford, M. (2014, January 24). "Modest uptick for income-based repayment." *Inside Higher Ed.* Retrieved from https://www.insidehighered.com/quicktakes/2014/01/24/modest-uptick-income-based-repayment

Teranishi, R., & Briscoe, K. (2006). Social capital and the racial stratification of college opportunity. In J.C. Smart (Ed.) *Higher Education Handbook of Theory and Research, XXI.* (pp. 591-614). Thousand Oaks, CA: Springer.

Cheng, D., & Thompson, J. (2017, May). *Make it simple, keep it fair: A proposal to streamline and improve income-driven repayment of federal student loans.* Washington, DC: The Institute for College Access and Success. Retrieved from https://ticas.org/sites/default/files/pub_files/make_it_simple_keep_it_fair.pdf

Tierney, W. G., & Venegas, K. M. (2009). Finding money on the table: Information, financial aid, and access to college. *The Journal of Higher Education*, *80*(4), 363–388. http://doi.org/10.1353/jhe.0.0059

Titus, M. A. (2007). Detecting selection bias, using propensity score matching, and estimating treatment effects: An application to the private returns to a master's degree. *Research in Higher Education*, *48*(4), 487–521. http://doi.org/10.1007/sl

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232. http://doi.org/10.1016/0010-0285(73)90033-9

U.S. Department of Education, Federal Student Aid. (n.d.). Repayment plans. Retrieved from https://studentaid.ed.gov/sa/repay-loans/understand/plans

U.S. Department of Education, Federal Student Aid. (2014, December 1). *Servicer summit: Loan portfolio briefing document.* Retrieved from https://fsaconferences.ed.gov/conferences/library/2014/servicing/2014ServicingSummitPortfolioOverview.pdf

U.S. Department of Education, Federal Student Aid. (2016a). *Income-driven repayment plans for federal student loans*. Retrieved from https://studentaid.ed.gov/sa/sites/default/files/income-driven-repayment.pdf

U.S. Department of Education, Federal Student Aid. (2016b). *Income-driven repayment plans: Questions and answers*. Retrieved from https://studentaid.ed.gov/sa/sites/default/files/income-driven-repayment-q-and-a.pdf

U.S. Department of Education, Federal Student Aid. (2017a). Direct Loan portfolio by loan repayment plan. Retrieved from https://studentaid.ed.gov/sa/about/data-center/student/portfolio

U.S. Department of Education, Federal Student Aid. (2017b). *Session 21: Income-Driven Repayment (IDR) plans & Public Service Loan Forgiveness (PSLF).* 2017 FSA Training Conference for Financial Aid Professionals. Retrieved from https://fsaconferences.ed.gov/2017sessionlist.html

U.S. Government Accountability Office. (2015). *Federal student education could do more to help ensure borrowers are aware of repayment and forgiveness options*.

Washington, DC: U.S. Government Accountability Office. Retrieved from http://www.gao.gov/assets/680/672136.pdf

Woodhall, M. (1983). *Student loans as a means of financing higher education: Lessons from international experience.* (World Bank Staff Working Paper No. 599). Washington, DC: World Bank. Retrieved from https://eric.ed.gov/?id=ED242231

Yosso, T.J. (2005). Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race, Ethnicity, and Education, 8*(1). 69-91.