Essays on the Economics of Student Loans and Economic Demography

by

Michael John Murto

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2023

Doctoral Committee:

Professor John Bound, Chair
Professor Kevin M. Stange
Professor Melvin Stephens, Jr.
Professor Basit Zafar

Michael J. Murto
mjmurto@umich.edu
ORCID ID 0000-0002-7625-5881

To Emily & Alba

ACKNOWLEDGEMENTS

I am deeply indebted to many people who have supported me–academically and personally–during the process of completing this dissertation. It would have been impossible to finish this work without their guidance, empathy, suggestions, occasional necessary diversions, and patience. I truly feel that I cannot express to these people how much their support has meant to me, nor can I fully thank them. Hopefully, these words will offer a small glimpse into the depth of my feeling for them.

First, I want to thank my committee members: John Bound, Kevin Stange, Mel Stephens, and Basit Zafar. Kevin was an early supporter of my research–his enthusiasm to help and genuine interest in my ideas was inspiration for me to develop this research. Basit was quick to offer support and his eye for detail helped sharpen my work–I hope to emulate his capacity and work ethic as a researcher. Mel's expertise as an applied microeconomist and willingness to advise my research greatly improved the final product. The immense value of these people's time, energy, and guidance is not lost on me, and I am deeply appreciative of them.

I owe special thanks to John. He advised my Summer Research Assistantship after my first year–a project that examined current literature in higher education finance–and encouraged me to develop my interest in this area. His early support and guidance as well as his continuous investment in my ideas nurtured this research project. He later, in addition to being a mentor and supervisor, became a coauthor when he invited me to work on a project on competing risks in demographic analysis, which eventually developed into Chapter III of this dissertation. For the many hours of his time spent in discussions, his endless patience, detailed feedback, insightful suggestions, and push to improve this work, I am incredibly thankful. Even more so, I am thankful for his engagement with me as a person. His kindness, empathy, and his willingness to see and appreciate the individual is admirable. His credentials and accomplishments as a researcher are impeccable and act as inspiration for me. And yet his warmth as a person is an even better reflection on him.

I owe thanks to many other faculty at the University of Michigan. Along with John, Susan Dynarski acted as a reader for my third year paper, which eventually developed into Chapter II of this dissertation. I am very grateful for her feedback, suggestions, and guidance. Charlie Brown lived up to his reputation as an insightful, thoughtful contributor. This work is immensely improved by engaging with his questions and comments. I deeply appreciate Ash Craig's interest and input. Discussions with Andreas Hagemann on applied econometrics improved my empirical analysis. Zach Brown, Sara Heller, Mike Mueller-Smith, and Edward Norton graciously offered input and suggestions as well. This work has greatly benefited from the collaborative scholarly environment at Michigan.

Martha Bailey was instrumental in the development of this dissertation, especially the empirical methodology. She was always willing to lend an ear and offer advice. I am deeply appreciative for the opportunity to work with her as a research assistant and collaborator and for her contributions to my growth as an economist.

Thanks, too, to my co-authors: Tim Waidmann and Arline Geronimus, who along with John co-authored Chapter III of this dissertation. And thanks to Martha, Andrew Goodman-Bacon, and Valentina Duque, who have been wonderful collaborators on additional research.

During this time, I was supported by a National Institute on Aging grant to the Population Studies Center at the Institute for Social Research (T32AG000221) as a pre-doctoral trainee in formal demographic methods and economic demography under the supervision of Martha Bailey and later John Bound. My work also benefited from the financial support of the Marshall Weinberg Endowments, the Population Studies Center Research Grants, the Population Studies Center Director's Research Fund (via Professor Paula Fomby), the University of Michigan Economics Departmental Fellowship, the University of Michigan Economics Department Research Grant, the Rackham Graduate School Research Grant, and the William Haber Graduate Fellowship. Furthermore, I'd like to thank Brenda Avoletta, Allen Fisher, Cathy Kelmar, and Steve Lonn for their help in data access and preparation. Many thanks to the helpful and attentive staff at Michigan, especially Laura Flak, Julie Esch, Julie Heintz, and Lauren Pulay in the Economics department and to Miriam Rahl at the ISR.

Prior to Michigan, I was lucky to have a number of excellent, giving mentors and professors who helped build my interest in research and encouraged me to continue my studies. I especially want to thank Bob Cumby, Susan Fleck, Bruce Meyer, and Gene Amromin. Each of these people acted as excellent advisors in steps along my development as an economist.

I also owe a deep debt of gratitude to my fellow graduate students at Michigan, their significant others, and my friends from other stages in life–their support, both academic and personal, has been integral in completing this project. I especially want to thank James and Joye Allen, Jon Denton-Schneider and Zach Yetmar, Jenny Mayo and Luke Burns, Ben Pyle and Gillian Wener, Elird Haxhiu, and Thomas Helgerman, whose friendship has helped sustain me during my time at Michigan and who continually inspire me to be a better researcher and person. Shawn Martin, Paul Kindsgrab, Dylan Moore, and Evelyn Smith formed an early research group along with Jenny and Ben. Avery Calkins, Mike Ricks, Andrew Simon, and Brenden Timpe offered feedback and support as I developed my research. Matt Nafziger and Twila Albrecht, Matthew Gorey, Erik and Julia Johnson, Mike Schoppmann and Sonia Chopra, Rachel Carnahan and Will McConnell, Scott Houghton, Dylan Lukes, Phil Sung, and Leigh Monahan were constants in my life over this time, and their friendship and support helped me get through the most difficult times.

My family has offered their unwavering support, without which I couldn't have finished this project. To my parents, Tom and Lauren Murto, sister and her family, Jenny, Graham, Lincoln and Lilly Clark, brothers and their families, Andrew, David and Bethany, Eric and Emmy, I owe you more than I can every repay. To my in-laws, Tom and Joanne Bertsche, Brian and Alexa, Katie and Clark, and Joe, thank you for welcoming me into your family. And to my grandmother, Alice Murto, and my grandparents who I lost, Bob Murto and Donald and Loraine McFarland, I love and cherish you all.

To Alba, you came into my life partway through my time at Michigan, but you have made me a complete person. The depth of my love for you is something I couldn't have imagined before you were born.

Above all, to my wife, Emily. You have shown me limitless patience and love. You truly are the best thing that could have ever happened to me. Without your love and support, I could not have done this, and I cannot express to you how you much mean to me. I love you.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

This dissertation studies causes of key human capital outcomes in the United States. Beginning with education, I examine how the student loan repayment plan structure changes borrowers' behavior, plausibly affecting both major and occupational selection. In health, I examine how the opioid epidemic led to more deaths than suggested by overdoses alone, but caution that methods leveraging geographic variation to identify the total impact of opioids should be exercised with attention paid modelling assumptions.

Driven by concerns over borrowers' ability to repay student loans, policymakers in the United States have turned to Income Driven Repayment (IDR) plans to offer flexible terms of repayment. Chapter I examines borrowers' educational investments when an IDR plan becomes a viable alternative. I construct a model of borrower behavior, highlighting the central role of repayment plan structure in students' optimal decision-making. I link academic records from a major public university to consumer credit bureau data to assess the empirical evidence for shifting major selection. After an expansion in IDR policies, male borrowers are more likely to select majors with worse initial labor market outcomes but higher wage growth, consistent with theoretical predictions: selecting majors associated with 3.2%-4.7% higher poverty rates but also 4%-5.7% higher earnings growth during their early career. Given proposed expansions to IDRs, this is relevant to understand how students' educational investments will respond.

Chapter II turns to the labor market, examining a directed search model with educational investment and loans. Repayment plans alter search behavior: under Fixed Repayment, higher payments push high ability people into higher-wage submarkets whereas lower ability people move to lower-wage jobs. These heterogeneous responses mirror two strands of current literature within one model. Under an IDR plan, individuals search in higher wage submarkets relative to the no-loan case, with the difference depending plan parameters, including time remaining in repayment, the income disregard, and limits on maximum repayments. As in existing literature, under a graduate tax, search behavior coincides with the no-loan case. However, this is true if borrowers cannot save. By introducing features

more common to IDR plans or allowing saving, this no longer holds. I outline conditions on efficient loans, which mitigate differences in search behavior between borrowers and non-borrowers and encourage optimal college attendance.

With John Bound, Timothy A. Waidmann, and Arline T. Geronimus, Chapter III examines how opioid use in the United States increased other-cause mortality. In recent decades, life expectancy among U.S. adults stagnated or fell. Dramatic increases in drug and alcohol-related deaths, especially opioid overdoses, have been identified as a key driver of those trends. Estimating the total contribution of opioid use to population mortality using standard demographic techniques assumes independence between competing risks. Alternative methods have been developed that rely on spatiotemporal correlations among causes of death to identify the total impact of specific behaviors. However, these methods require assumptions about causal pathways. We relax several of these assumptions by introducing economic distress as a potential common driver of opioid use and other types of mortality and disaggregate mortality trends by time and type of opioid involved. We find evidence that correlations driven by economic conditions affect prior estimates of opioids' total impact on mortality and sensitivity to time period and type of drug studied, substantially overstating recent contributions of opioid use to life expectancy trends.

CHAPTER I

Student Loans and Human Capital Investments: The Role of Repayment Plan Structure

1.1 Introduction

As of 2022, 43 million borrowers owe over $1.6 trillion in federal student loans.[1] Loans are the dominant form of federal aid for higher education today: in constant 2020 dollars, aggregate federal loan issuance increased from $7.5 million in 1970–about the same size as federal grants–to $21.9 million in 1992–about 50% larger than federal grants–before experiencing rapid growth over the 1990s. At the turn of the century, over $50 million new federal loans were issued, more than three times the size of the federal grant program, and the loan program continued to grow over the next decade, doubling again in size.[2] The staggering growth in the federal loans program and evidence of borrowers' difficulty repaying their debts has triggered policymakers to rethink the repayment system for loans.

On August 24, 2022, the Biden Administration announced major policy changes to federal student loans. In addition to means-tested forgiveness for current borrowers, the Administration announced a new *Income Driven Repayment* (IDR) plan that ties repayment to income rather than the amount borrowed. Theoretical work argues that such plans can reduce default for borrowers, efficiently solve the hold-up problem in investment, and ease distortions in labor market behavior by insuring against low wage realizations on the labor market. On the other hand, IDRs could potentially cause their own distortions by effectively subsidizing low-monetary return human capital investments. However, there is *no empirical evidence* on how students' human capital investments respond to repayment plan structure.

[1]National Student Loan Data System (NSLDS) (2022)
[2]Ma and Pender (2021)

In this paper, I present evidence on how some students adjust their human capital investment decisions when an IDR becomes a viable alternative. Forward-looking agents should anticipate the insurance value offered by IDRs while in school. In particular, as IDR offers both insurance and subsidy value to particular labor market returns, agents should adjust their behavior while in school and select majors that have increased relative returns after the IDR introduction: majors with lower remunerative or more variable returns over early career. To empirically assess the importance of this margin, I construct a dataset that links rich academic records from a major public research university–from which I can observe academic decisions on the part of students over time, but not financial aid data–to credit bureau data–from which I can observe borrowing while in school. These data allow me to observe actual academic decisions and borrowing status of students as loan policies change. I examine an earlier IDR policy expansion that led to a dramatic shift in repayment patterns in the United States, which sheds light on the extent to which students respond to IDR availability.

I implement a differences-in-differences analysis to identify causal effects of the IDR expansion, using non-borrowers as a "control" group as the policy does not affect their relative returns. I find evidence that male borrowers are more likely to select majors with initially lower remunerative returns, but higher income growth, consistent with the theoretical predictions. In particular, men select majors with 0.2-0.3 percentage points higher poverty rates in early career labor market outcomes and 1-2 percentage points higher income growth between early- and mid-career after the IDR expansion and subsequent growth in usage. These results are robust to specifications that account for non-random selection into treatment and control groups as well as potential endogenous selection into treatment induced by the policy change. Recognizing that borrowers and non-borrowers are not randomly selected groups, and moreover the *composition* of the groups may differ over time, I extend the reweighting approach of Abadie (2005) to explicitly account for compositional changes in treatment and control groups to recover the average treatment on the treated (ATT).[3] I find no shifts in behavior for women–given Zafar (2013)'s finding that men are more likely to emphasize pecuniary labor market returns in their selection of major relative to women, this is consistent with the hypothesis that men are more responsive to pecuniary returns on the labor market *net of loan repayments*.

Over the past two decades, researchers have identified considerable distortions in

[3]This contributes to the nascent literature in applied econometrics on difference-in-differences with compositional shifts (Hong (2013)).

borrower behavior associated with student loans *in general*, though there has been little empirical work examining repayment plan structure specifically and none that examines human capital investments and repayment plan structure. Research suggests that student loans succeed in promoting access to higher education (Black et al. (2020); Solis (2012); Card and Solis (2022)) and historically credit constraints are not a major constraint in the decision to *attend* school (Carneiro and Heckman (2002); Keane and Wolpin (2001); Johnson (2013))–though the combination of increasing costs and stagnating borrowing limits may increase the importance of this constraint (Lochner and Monge-Naranjo (2016)). However, default rates on student loans–while declining over the 1990s–remain high (Looney and Yannelis (2015)). Furthermore, student loan delinquency rates (90+ days past due) are higher than corresponding rates for mortgage debt, auto debt, and in recent years, even credit card debt (Federal Reserve Bank of New York, Center for Microeconomic Data (2022)). Difficulties in repaying loans are only one undesirable outcome from education debt. There is some evidence that borrowers pursue higher income occupations as a result of their debt (Rothstein and Rouse (2011)), have lower likelihood of attending graduate school (Akers (2012); Millet (2003)), lower likelihood of purchasing a home (Cooper and Wang (2014)), decreased entrepreneurship (Krishnan and Wang (2019)), and delay family formation (Addo (2014)).

In related work, Hampole (2023) documents that the introduction of no-loans policies by university financial aid departments *increases* the probability that students select majors with lower initial earnings but higher lifetime earnings, suggesting that financial frictions impact optimal human capital investments, especially when considering a particular major's exposure to income variability over time. She also documents increases in initial earnings variability, consistent with an increased appetite for risk. Here, I present evidence that the effect of these financial frictions on student behavior depends on the *structure* of repayment plans, especially considering a major's exposure to risk of low-earnings as well as income variability over time. This work suggests that student loans–and the specifics of student loan repayment plans–are an important consideration in students' optimal human capital investments, with implications for the earnings distribution and trajectories of college graduates.

Much of this empirical work examines the *historical* setting of loan policy in the United States.[4] Historically, student loans in the United States follow an *installment*

[4]Solis (2012) and Card and Solis (2022) consider Chilean higher education loans, but are some of the best evidence on the effect of loans on attendance.

repayment plan, with fixed repayment amount determined by the principal and interest rate due over a set period of time–10 years under the standard plan. This inflexible repayment plan comes at a time when borrowers face the most uncertainty in the labor market (Dynarski (2014)). This mismatch between repayment requirements and labor market uncertainty is reflected in the *repayment burden* of student loans: while the *repayment amount* is known to the borrower, the share of their disposable income required to service the debt–the *repayment burden*–is unknown (Barr et al. (2019)). An alternative approach instead links repayment to labor market realizations, through Income Driven Repayment. IDR plans insure against bad labor market outcomes, reversing the relationship between repayment amount and repayment burden–while the exact repayment amount is unknown to the borrower ex ante, the repayment burden is known (Stiglitz (2014); Barr et al. (2019)). Theoretical models suggest that optimally designed loan systems should consider income (Lochner and Monge-Naranjo (2016), Quiggin (2014)), and that IDR plans could reduce default (Chapman (2014); Quiggin (2014)), efficiently solve hold-up investment problems (Moen (1998)), and ease distortionary behavior on the labor market (Ji (2021); Kaas and Zink (2011)).[5] On the other hand, IDR plans may also lead to moral hazard, effectively subsidizing low-return human capital investments and less labor market effort (Looney (2022); Quiggin (2014)).[6]

In recent empirical work, Herbst (2023) documents post-graduation outcomes, highlighting a reduction in default rate and increase in consumption for borrowers enrolled in IDR plans, consistent with the insurance benefits of IDR. Mueller and Yannelis (2019) also document a reduction in default rates among borrowers enrolled in IDR and highlight their

[5]In separate work, I highlight that labor market distortions under fixed repayment differ for different types of borrowers, and hence the introduction of IDR plans can lead some workers to search in higher wage sub-markets and others in lower wage sub-markets relative to the fixed repayment system. Furthermore, IDR can introduce a different form of time-varying distortionary behavior, as the subsidy value of forgiveness becomes more relevant in optimal labor market behavior, see Chapter II of this work.

[6]Quiggin (2014) argues that borrowers who expect to repay loans with probability 1 do not distort labor market behavior, since distortions only delay repayment–only those who expect to partially repay their debt would suffer from moral hazard. However, Karamcheva et al. (2020) project that typical borrowers enrolled in IDRs will experience substantial forgiveness and find that increased IDR enrollment explains a significant amount of the decline in student loan repayment rates over the last decade. The latest Department of Education projections are that 43% of borrowers entering repayment on IDR in 2022 will have debt forgiven under the IDR program, suggesting this is a non-negligible effect (Department of Education (2022)). Furthermore, Mueller and Yannelis (2022) highlight that such comparisons of timing and present value of repayments depend on comparisons between the underlying interest rate on the debt and the discount rate. For interest rates that are higher (lower) than the discount rate, the delayed repayments under IDR may increase (decrease) present value of aggregate repayments, even if the dollar amounts repaid are the same. Finally, these projections are under current law and do not account for any changes caused by the Biden Administration's new proposed IDR plan, which significantly increases generosity of IDR terms (Biden (2022)). This suggests that an even larger fraction of borrowers may exhibit distortionary behavior due to moral hazard.

role in insulating borrowers from wealth shocks. In an experimental setting, Mueller and Yannelis (2022) identify significant reductions in repayment amounts and default probabilities as well as increases in consumption among borrowers induced to switch into IDR plans. Furthermore, using simulation exercises, they suggest reductions in present value of aggregate repayments among their sample of borrowers by enrolling in IDR plans, driven in part by eventual forgiveness of remaining balances. However, no empirical work has documented any effects on *human capital investments* among borrowers in the United States.[7] Yet, forward-looking agents should anticipate the insurance value of IDR plans and change their optimal investments accordingly.

Major declaration is a relevant margin to consider for how IDR availability affects investment behavior. Recent work has highlighted that these choices *within* college lead to dramatically different labor market returns (Altonji et al. (2016); Hershbein and Kearny (2014)), income trajectories (Martin (2022)), as well as exposure to earnings variability both across graduates with the same major and within graduates over time (Andrews et al. (2022)). Furthermore, expected labor market returns are an important consideration in students' major selection (Beffy et al. (2012); Wiswall and Zafar (2015)) and similar discrete choices–over occupations rather than major–play an important role in managing exposure to risk (Dillon (2018)). This literature suggests that major selection has important implications for graduates' expected labor market returns *and* exposure to labor market risk. Finally, Looney (2022) highlights that a particular concern with IDR plans is potential subsidization of unproductive majors.

While theoretical models suggest that repayment plan structure should affect students' optimal human capital investments, the magnitude of such effects is not clear. Additionally, students may have strong preferences for particular majors independent of labor market returns (Wiswall and Zafar (2015)) or aversion to debt in general (Field (2009), Caetano et al. (2019)) that may dampen these effects. Finally, the burden of enrolling in such plans (Cox et al. (2020); Mueller and Yannelis (2022)) and lack of information about such plans may also lead to small effects. In all, empirical measurements to assess the extent to which repayment plan structure matters for students' human capital investment decision will improve our

[7] IDR funding schemes have existed in other nations–notably Australia and the UK–for a considerable time. Research on the effects of these systems typically centers on questions of access, repayment, and certain other post-graduation outcomes, see Chapman (2006) and de Gayardon et al. (2018) for reviews. Comparison between system types, however, is constrained by the lack of policy variation in repayment structure *within* nations. Furthermore, to the best of my knowledge, there is no empirical evidence on the effect of these systems on human capital investments *within* higher education by borrowers.

understanding of the effects of these policies and can help inform policymaking moving forward.

In the following section, I briefly describe the policy background in student loans. Section 1.3 presents a theoretical model that highlights the importance of repayment plan in borrowers' human capital investment decisions, particularly major selection. Section 1.4 describes the data I use and Section 1.5 describes my empirical approach. Section 1.6 presents results from my empirical analysis and Section 1.7 concludes.

1.2 Policy Background

As part of Title IV of the Higher Education Act (HEA) of 1965, loans were part of a broader strategy to promote access to higher education for families of modest means (Presidential Task Force on Education (1964); Presidential Task Force on Education (1967)).[8] Initially, loan disbursements were roughly equal to federal grants. However, a combination of conditions–including stagnating real appropriations for grants (Cervantes et al. (2005)), declining direct public funding for universities and corresponding increasing reliance on private, tuition-based financing (McPherson and Schapiro (2006)), and amendments to the HEA in 1992 that opened lending to middle income families–led to federal loans being the dominant form of student aid by the turn of the century. Figure 1.1 documents trends in Federal student aid over the last three decades.

As the student loans program expanded, additional lending programs were added. This resulted in a constellation of programs geared towards different borrowers, administered by different sources, with different eligibility requirements and loan terms. For the purposes of this paper, I focus on the Stafford loan program, the largest federal loan program.[9]

Under the Stafford loan program, students can qualify for means-tested Subsidized loans, under which no interest would accrue while the student was in school, and any student can take on Unsubsidized loans, under which interest would accrue, but no payments would be due while the student was in school. Interest rates are set statutorily and fixed for the lifetime of the loan–that is, the interest rate in effect *at time of disbursement* is

[8] The first Federal student loan program was implemented in 1958 by the National Defense Education Act. This loans program–what would eventually become the Perkins loan program–was folded into the HEA in 1965.

[9] Approximately 75% of all Federal loans disbursed in 2020-2021 were Stafford loans, the majority of which–55%–were directed towards undergraduate borrowers. Both of these are slightly lower than recent averages, which were approximately 80% and 65% between 2008-2019, respectively (Ma and Pender (2021)).

Figure 1.1: Aggregate Federal Student Aid, by Type

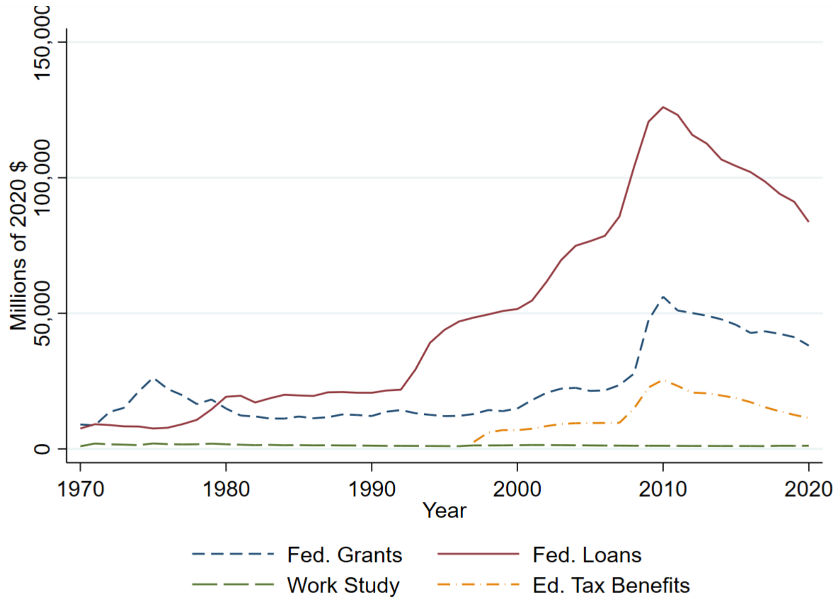


Figure Notes: Data are for all federal loans, regardless of level of study (i.e. include loans made to undergraduate and graduate students). The 1992 Higher Education Act extension created the unsubsidized loan program, which opened up borrowing to all students, regardless of family income. Beginning in 2007, graduate students could borrow under PLUS loan program with no statutory cap. 2008 was the last time the federal loan limits were raised.
Source: Ma and Pender (2021)

fixed for the lifetime of the loan.[10] Because students typically borrow for multiple years, their effective interest rate on their debt is the weighted average of individual loans.[11] Similarly, there are annual and lifetime limits on the loans, which differ based on the dependency status for students. Tables 1.1 and 1.2 report these loan terms for dependent students in disbursement years 2006-2019. Additional information is reported in Appendix A.1.

In order to receive Federal loans, students fill out the Free Application for Federal Student Aid (FAFSA). One application is necessary for each academic year, with applications due prior to the beginning of the academic year.[12] The Department of Education uses FAFSA information to determine grant eligibility, work-study eligibility, and loan eligibility. If a

[10]Between 1988 and 2006, student loans had variable interest rates, recalculated annually. However, borrowers could "lock in" rates through consolidation onto a fixed rate loan.

[11]While students can consolidate multiple loans into a single loan, the interest rate on the new loan is again a weighted average of the consolidated loans.

[12]States and individual institutions also use FAFSA information to determine aid. As such, the due date can be State- and institution-specific. Typically, though, they are required during the Summer prior to initial enrollment for new students or at the end of the academic year for continuing students, e.g. at the end of Spring term for students continuing their studies the following Fall term.

student accepts their loans, disbursements are generally made directly to the institution on behalf of the student.[13] After student-borrowers drop below half-time enrollment or complete their studies, they undertake "exit counseling," which gives information on interest rates, initial grace period, forbearance and deferral, default, and repayment options. The student selects an initial repayment option and, after a 6-month grace period, enters repayment.

Under the standard, 10-year fixed repayment program, monthly payments are determined by the borrower's original principal and effective interest rate according to the formula:

$$PaymentAmount_{FR} = P_0 \frac{R(1+R)^{120}}{(1+R)^{120}-1}$$

Where P_0 is the original principal[14] and R is the monthly effective interest rate.

Here, it is immediately obvious that the standard loan repayment is insensitive to labor market realizations–repayment is determined *only* by the initial principal and interest rate. In general, these repayment amounts are due monthly. If a borrower misses three consecutive payments (90 days past due), their delinquency is reported to credit bureaus, which could adversely affect credit access. In addition, if a borrower misses 9 consecutive payments (270 days past due), they default on their loans. Loan default potentially results in (i) loan acceleration, where the entire balance is immediately due; (ii) loss of benefits, such as loan deferral, forbearance, or access to alternative repayment plans (including IDR plans); (iii) wage garnishment and seizure of tax returns and other federal benefits to cover debts; and (iv) additional fees added onto the loan (CFR (2021)). Because federal student loans are not automatically dischargeable in bankruptcy, borrowers bear these costly default penalties, in general.[15]

In contrast to fixed repayment, Income Driven Repayment (IDR) amounts do not explicitly depend on the principal and interest rate. There are two key parameters in any IDR plan: the

[13]If the total aid–inclusive of loans–exceeds the funds necessary to pay for tuition, fees, and other expenses due to the institution, excess funds are refunded to the student. These funds must be used for "qualified education expenses," which includes room and board. If the student decides to return their loan money within 120 days of disbursement, associated fees and any accrued interest in that time are waived. Disbursements occur at the beginning of academic terms while the student is enrolled.

[14]For loans that have any accrued interest at the time the borrower enters repayment–e.g. if a borrower has unsubsidized loans–accrued interest is capitalized (added to principal) in determining P_0.

[15]In fact, since 2005, all student loans, including private loans, are not automatically dischargeable in bankruptcy. It is true that loans *can* be discharged–either in full or part–through bankruptcy via a separate petition to the court showing "undue hardship," but this is an extremely rare occurrence: Taylor and Sheffner (2016) report that 0.1% of Chapter 7 bankruptcy claims petitioned the court to discharge student debt. Interestingly, approximately half of those who *did* attempt received some form of relief. Nevertheless, discharge in bankruptcy is not the norm for student loans.

Table 1.1: Borrowing Limits on Federal Student Loans

Academic Year	Year of Study First Year	 Second Year	 Third Year and above	Aggregate Limit
2006-2007	$2,625	$3,500	$5,500	$23,000
2007-2008	$3,500	$4,500	$5,500	$23,000
2008-2019	$5,500	$6,500	$7,500	$31,000
	no more than $3,500 sub.	no more than $4,500 sub.	no more than $5,500 sub.	no more than $23,000 sub.

For 2006-2008, the total limit for Federal loans, inclusive of subsidized and unsubsidized loans, coincided with the limit on subsidized loans alone for dependent students. Beginning in 2008, the total limit for loans was increased, but the cap for subsidized loans remained at 2007-2008 levels. For all years, the limit for subsidized loans did not differ between dependent and independent students, though independent students had larger total loan maximums. Only limits for dependent students are reported here. Limits for independent students are reported in Appendix A.1. Source, CFR (2021) and Hegji (2021)

income disregard–a level of income below which no payments are required–and the *repayment rate*–the share of income above the income disregard devoted to repaying the debt. In addition, most IDR plans–and all of them in the U.S.–have a finite window of time over which the loan must be repaid. Any outstanding balance remaining at the end of the repayment window is forgiven.[16] While the interest rate and principal do not affect the monthly repayment amount, they *do* affect the maximum amount that can be repaid, and in turn the amount of time a borrower is in repayment up to the maximum repayment window. If at any time the borrower fully repays their outstanding principal and accrued interest, then they exit repayment.[17] Under a general IDR, the monthly repayment amount is:

$$PaymentAmount_{IDR,t} = \begin{cases} 0 & \text{if } Y_t < b \text{ Or } t > T \text{ Or } P_t + A_t = 0 \\ \iota(Y_t - b) & \text{if } Y_t \geq b \text{ and } \iota(Y_t - b) < (P_t + A_t) \text{ and } t \leq T \\ P_t + A_t & \text{if } \iota(Y_t - b) \geq P_t + A_t \text{ and } t \leq T \end{cases}$$

Where Y_t is monthly income in month t, b is the income disregard, ι is the repayment rate, T is the maximum repayment window, P_t is outstanding principal in month t, and A_t is accrued interest in month t. For simple interest loans–as student loans are in the United States–the

[16] In the United States, the IRS considers forgiven debt as income, in general.

[17] Higher education financing can also take the form of fully taxpayer funded or a "graduate tax." The latter is similar to IDR in that only those who attend college–or in some cases, only those who complete college–finance the program, with contributions dependent on income realizations. However, in contrast to an IDR, the graduate tax does not end if the payer contributes enough to cover their own expenses plus interest, in general. That is, a graduate tax is an additional tax levied only on those who attended (completed) higher education.

Table 1.2: Interest Rate on Federal Student Loans

	Interest Rate	
Academic Year	Subsidized	Unsubsidized
2006-2007	6.80%	6.80%
2007-2008	6.80%	6.80%
2008-2009	6.00%	6.80%
2009-2010	5.60%	6.80%
2010-2011	4.50%	6.80%
2011-2012	3.40%	6.80%
2012-2013	3.40%	6.80%
2013-2014	3.86%	3.86%
2014-2015	4.66%	4.66%
2015-2016	4.29%	4.29%
2016-2017	3.76%	3.76%
2017-2018	4.45%	4.45%
2018-2019	5.05%	5.05%
2019-2020	4.53%	4.53%

Beginning in 2013, the interest rate for the academic year is determined by the high yield on the 10-year treasury at the final auction held prior to June 1, plus 2.05 percentage points. Source, CFR (2021) and Hegji (2021).

outstanding principal and accrued interest evolve according to the equations:

$$P_t = P_{t-1} - \max\{RepaymentAmount_{IDR,t} - A_t, 0\}$$
$$A_t = \max\{A_{t-1} + RP_{t-1} - RepaymentAmount_{IDR,t}, 0\}$$

Where R is the underlying interest rate,[18] and P_0 is some initial principal and $A_0 = 0$.

Under the IDR, the repayment amount can change as the labor market realizations–income Y_t–changes for the borrower. This allows for insurance against particularly bad labor market conditions, by adjusting repayment in response to bad draws.[19] Furthermore, the income disregard dramatically reduces default risk, since payments are not due for very low incomes.[20] On the other hand, the underlying principal and interest limit the total amount a

[18] Certain IDRs in the United States subsidize interest if the repayment amount is insufficient to cover accruing interest. This limits the amount that A_t can grow, but would not affect the speed at which P_t is paid down.

[19] In the United States, the repayment amount is recalculated annually, after a borrower recertifies their income using data from the previous year. In the event that a borrower experiences and adverse income shock mid-year, they can notify the Department of Education, but this adjustment would not be automatic.

[20] Herbst (2023) documents a 22 percentage point reduction in delinquency rates for distressed borrowers on IDR plans.

borrower can repay on the loan, offering "upside" protection from paying more than the underlying loan, accounting for interest.[21]

While suggestions for an Income-Driven Repayment have existed since at least the 1950s (see, e.g., Shireman (2017)), there was no IDR program for federal loans for the first three decades of federal involvement in postsecondary loans. In 1994, the Federal Government began offering a repayment plan called Income Contingent Repayment (ICR). This plan limited the repayment amount to the lesser of 20% of any income over the Federal Poverty Line or an alternative 12-year installment plan. The repayment window was a maximum of 25 years.[22]

As with any IDR plan, the underlying interest rate and principal are still key elements for the plan, though they no longer directly relate to the monthly repayment amount. That is, if repayments under the ICR plan were too low, interest would accrue on the loan increasing the total debt the borrower held. Under ICR, this interest was capitalized into the underlying principal up to a maximum limit of 10% the original principal, increasing future interest obligations. After this, interest would continue to accrue, but would not be added to the underlying principal.

The terms of this plan and its eligibility requirements made it an unattractive option for student loan repayment–the repayment amount as a share of income was relatively high, at 20%, and the income disregard was relatively low, set at the federal poverty line. Furthermore, the interest capitalization increased future obligations. Additionally, the plan originally did not explicitly account for family size in calculating the income disregard, penalizing borrowers who married or started families (Chapman (2006)).[23] Shireman (2017) argues that the shift in political landscape in the mid 1990s led the Department of Education to underemphasize the benefits of the ICR, limiting its effectiveness and role in repayment.

[21] In Quiggin (2014), it is this upside protection that limits moral hazard.

[22] Notably, the ICR plan could not be used for FFEL loans, though a borrower could consolidate FFEL loans into a new Direct Loan. See Appendix A.1 for additional information. The 1998 amendments to the Higher Education Act created the option for "Income Sensitive Repayment" on FFEL loans but (i) the details were left up to individual loan servicers and (ii) it did not allow payments to be less than accruing interest or extend beyond the standard 10-year repayment window, limiting its usefulness as an insurance mechanism against bad labor market outcomes for borrowers.

[23] The terms of the ICR proposal developed on December 1, 1994 included a provision to account for family size—see Government Publishing Office (1994)—but the terms of the plan published in the Code of Federal Regulations read "[i]f a borrower provides documentation acceptable to the Secretary that the borrower has more than one person in the borrower's family, the Secretary applies the HHS Poverty Guidelines for the borrower's family size." This language shifted to automatically apply the family-size-adjusted poverty line with the July 1, 2013 update of the Code of Federal Regulations (CFR (2021)).

In 2007, Congress passed legislation that introduced a new IDR plan, called the Income Based Repayment (IBR) plan. This program was implemented in 2009, lowering the repayment amount to 15% of income over 150% of the federal poverty line. During the first three years of repayment, any unpaid interest on subsidized loans did not accrue on the borrower's debt. This plan had an eligibility requirement, called a *partial financial hardship*–essentially, if the income based calculation exceeded the payments under the standard, 10-year repayment plan, the borrower was ineligible. If the borrower initially qualified, but later lost their partial financial hardship, any accrued interest would be capitalized into the principal.

While the Income Based Repayment offered some advantages over the older Income Contingent Repayment plan, there was relatively little increase in the aggregate share of borrowers enrolled in IDR plans. In 2013, the Pay As You Earn (PAYE) plan began. This program lowered the repayment amount to 10% of discretionary income, still defined as income above 150% the federal poverty level. This plan was initially only available to "new" borrowers: those who had no federal student loan debt prior to October 2007 and borrowed on or after October 2011. Like the IBR plan, it had an interest subsidy for subsidized loans for the first three years of repayment and an income eligibility requirement; however, in the event that a borrower lost their partial financial hardship, their interest capitalization was capped at 10% of their initial debt, limiting the impact of accruing interest. Additionally, it shortened the repayment time from 25 years to 20 years.

Prior to the introduction of PAYE, a relatively small and constant fraction of borrowers and total balances were enrolled in IDR repayment plans–a little less than 10% and 20%, respectively. After PAYE's implementation, IDR usage steadily grew. As of 2019, over a third of all borrowers holding over half of outstanding balances were enrolled in an IDR plan. Figure 1.2 displays this dramatic shift in repayment in the United States. Unfortunately, due to policy changes that led to changes in data storage at the Department of Education, data are not available for 2011-2013 (Avoletta (2020)). However, note that there is little movement prior to 2011 in IDR usage and at the beginning of 2013 usage statistics were in line with those from 2011. Despite these potential benefits of IDR plans and their recent growth, there are significant barriers to enrollment in the United States that likely discourage their use (Cox et al. (2020)).

In 2016, the Revised Pay As You Earn (RePAYE) plan was introduced. This plan was available to all borrowers with no income requirement. Like PAYE, it required 10% of

Figure 1.2: Share of Outstanding Balance/Borrowers by Repayment Type

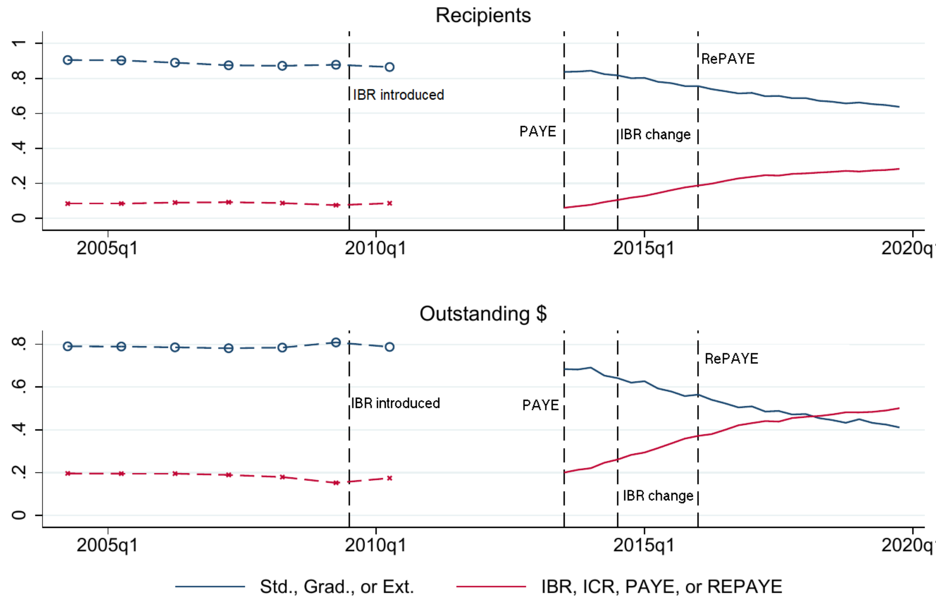


Figure notes: Standard, Graduated, and Extended repayment plans reported together. Income Contingent Repayment (ICR); Income Based Repayment (IBR); Pay As You Earn (PAYE); and Revised Pay As You Earn (RePAYE) plans reported together. Introduction dates of IBR, PAYE, RePAYE as well as change in IBR are marked, see Table 1.3 for details.
Sources: after 2013, National Student Loan Data System (NSLDS) (2022). Prior to 2011, data are from the decommissioned Common Services for Borrowers (CSB) Data Mart, retrieved via personal communication with Department of Education (Avoletta (2020)). CSB Data Mart data uses number of loans to determine share of borrowers, rather than individual borrowers. Data are unavailable between 2011 and 2013 due to migration of data system reporting and storage at Department of Education.

income over 150% the poverty line, and subsidized 100% of accruing interest on subsidized loans for the first three years. In addition, RePAYE subsidized 50% of any accruing interest on subsidized loans for all years after that and subsidized 50% of any accruing interest on unsubsidized loans for all years. Finally, because there was no partial financial hardship requirement, the only event that could induce interest capitalization was voluntarily leaving the RePAYE plan or failing to recertify.

Table 1.3 details the differences between the IDR policies in the United States. In this paper, I will specifically examine changes in borrower behavior around the introduction of PAYE in 2013. This choice is motivated by (i) potential confounding factors with the Great Recession and the introduction of Public Student Loan Forgiveness (PSLF) concurrent with IBR in 2009 and (ii) the fact that, empirically, it is after PAYE in 2013 that IDR usage began to steadily increase. Because this study asks how IDR expansion changes human capital investments, I focus on a policy change that preceded actual shifts in repayment behavior.[24]

1.3 Motivating Model

I build on the model of Altonji et al. (2016), which captures the students' college-major decision by modeling uncertainty over abilities and discrete choice of major. Majors have different labor market returns, depending on major-specific ability. To their model, I add (i) pecuniary costs of education, (ii) labor market dynamics, and (iii) risk averse agents. In an extended version of this model presented in Appendix A.2, I consider labor market uncertainty, default risk, and particular policy parameters that limit lifetime IDR payments. The main takeaways from this model are that loans with fixed rate repayment discourage investment in majors with "risky" wage-growth profiles, relative to either grant-based education or IDR loans. This is especially true when imperfect credit markets in the labor market prevent consumption smoothing. In contrast, IDR loans increase appetite for risk, by limiting the cost of low wage draws on the labor market. The IDR loans have two effects on behavior: (i) IDR automatically provides partial consumption smoothing for low-wage draws–including wage draws that would trigger default under the fixed repayment regime, making them especially valuable for investments with higher probability of very low wage draws–and (ii) forgiveness of any remaining balance at the end of the repayment window subsidizes risky behavior, by passing costs for low wage draws away from the borrower.[25] If the labor market has perfect credit markets, then the consumption smoothing effect of IDR loans is redundant, and the IDR acts as a pure subsidy for risk (moral hazard).[26] However, if there are imperfect credit markets, then both effects are relevant.

[24]The statutory authority for both PAYE and RePAYE is the same authority that led to the creation of ICR in 1994. Hence, *only* Direct Loans could be enrolled in these plans. As noted, privately-held FFEL loans could be consolidated into a new loan, though this could result in interest capitalization, increasing total obligations. For my empirical section, I use data from a University that was enrolled in the Direct Loans program for the entirety of the timeframe I examine, mitigating concerns that the end of the FFEL program in 2010 could itself induce additional IDR usage and thus shifts in borrower behavior. See Appendix A.1 for additional details.

[25]The model only considers a simple partial equilibrium, and does not model the government's decision, nor labor demand nor education supply. Only the student/worker decisions are considered.

[26]It is not necessarily true that this is an inefficient outcome. For example, if the "risky" investments are also socially valuable, then this subsidy may be efficient. It is worth noting that both President Clinton and President Obama explicitly cited selection into teaching, nursing, and social work as rationale for implementing IDR plans (Clinton (1993); Obama (2014)). In addition, the law signed by President Bush in 2007 included a separate program–designed to work in conjunction with IBR–to promote sorting into public service oriented careers, the Public Student Loan Forgiveness (PSLF) program. In this sense, the subsidies offered by IDRs may be a policy choice to promote investments in "risky" majors that also have positive externalities.

Table 1.3: IDR Policy comparison

	ICR	IBR	PAYE	RePAYE
Year Introduced:	1994	2009	2013	2016
Rate	20%	15% (10% post 2014)	10%	10%
Repayment Time	25 years	25 years (20 post 2014)	20 years	20 years for undergrad. 25 for grad.
Income eligibility	none	PFH	PFH	none
Income disregard	FPL[a]	150% FPL	150% FPL	150% FPL
Interest subsidy	none	100% for sub. loans, 3 years		+ 50% after +50% for unsub. loans
Capitalized int.	monthly capped 10% principal	No longer PFH no cap	No longer PFH capped 10% principal	none n/a

[a]The ICR initially did not account for family composition in determining the income disregard. This changed explicitly with the 2013 update to the Code of Federal Regulations. Prior to this, a borrower could request adjustments based on family size. “FPL” stands for Federal Poverty Line

Here, I present a simplified model that highlights key differences in borrower responses under fixed repayment and income-driven repayment, which motivates the use of a difference-in-differences approach in my empirical analysis. I shut down labor market uncertainty and consider a simplified version of the IDR that links repayment to income, but does not consider repayment caps. The full version of the model is presented in Appendix A.2.

1.3.1 Simple Model

Consider a simple model where agents make the following sequential decisions: (i) they choose whether or not to invest in human capital; (ii) they choose how to fund their investment; (iii) they select a major. The agents' utility is determined by expected labor market returns to their selected major net of any repayments on their investment financing. When they make their financing decision, agents choose between loans with known repayment plan and an alternative, costly financing scheme, e.g. working while in school.

Agents begin with wealth $y_i \in \mathbb{R}_+$, which can be used to pay for pecuniary costs of human capital investment.[27] Let l_i determine the loans agent i borrows and z_i be the alternative funds. Their total financing must pay for exogenously determined pecuniary costs of investment, $y_i + l_i + z_i = x_i$ where $l_i, z_i \geq 0$. Agents select a major j from set J and recieve two periods of income from this major, which is a function of major-specific ability, A_i^j and wage w_j^t for periods $t = \{1, t\}$. For simplicity, assume $J = \{1, 2\}$. Loans are repaid in each period by repayment plan R_t which may be a function of total loan balance (under standard repayment) or income (under IDR). Total utility is given by $u(w_j^1 A_i^j - R_1) + u(w_j^2 A_i^j - R_2)$. In this model, I assume that the alternative form of financing, z_i, is costly in that it reduces the effectiveness of human capital investments; that is, ability is a decreasing function of alternate financing: $A_i^j(z_i)$.[28] If the agent does not invest in human capital, they receive $u(c_0)$ in both periods.[29]

[27]Here, I assume that initial wealth can *only* be used for human capital investment. If there is an alternative use for wealth–e.g. consumption–agents will choose between consumption and investment in order to maximize lifetime utility.

[28]If there were no cost to z, then there is no reason to finance human capital investments through lending. Alternatively, one could consider a case where some level of z is "free," after which it becomes costly. Under any repayment program, agents will exhaust this "free" financing prior to borrowing or using costly alternative financing.

[29]My empirical approach examines students at a selective institution who have already chosen to enroll and, as highlighted in Section 1.4, this population does not change significantly over my time frame. While the decision to invest or not–including the choice of institution–is a crucial consideration, and one I believe future research on the effects of IDRs should examine as a margin of interest, it is beyond the scope of this paper.

First, consider the agents' major selection, conditional on their selection of l_i and z_i. Suppose $w_1^1 = w_1^2$ and $w_2^1 < w_2^2$ such that $w_2^1 < w_1^1 < w_2^2$–in words, one major gives higher initial wage income but no growth, while the other offers an intertemporal tradeoff via lower initial wages but higher growth. An agent, i, will select major 1 if and only if

$$u(w_1^1 A_i^1 - R_1) + u(w_1^1 A_i^1 - R_2) \geq u(w_2^1 A_i^2 - R_1) + u(w_2^2 A_i^2 - R_2)$$

Consider how fixed repayment loans affect this decision: let $R_1 = (1+R)l_i$ and $R_2 = 0$. All else equal, the presence of fixed repayment loans will reduce both sides of this inequality, but will reduce the right hand side by a larger magnitude for risk averse agents. Hence, loans with fixed repayment plans reduce the relative attractiveness of majors that have more variable monetary returns relative to majors that have higher initial wages but no growth.

Consider how IDR loans affect this decision: let $R_1 = w_j^1 A_i^j \iota$ and $R_2 = w_j^2 A_i^j \iota$ where $\iota \in (0,1)$ is the repayment rate. If preferences have constant relative risk aversion, then the optimal choice of major no longer depends on loan balances.[30] That is, the comparison between majors is identical to

$$u(w_1^1 A_i^1) + u(w_1^1 A_i^1) \geq u(w_2^1 A_i^2) + \ u(w_2^2 A_i^2)$$

Hence, agents with this repayment plan make their major decision independently of borrowing status, all else equal.

This motivates the use of a difference-in-differences analysis in my empirical approach: the difference in major selection between borrowers and non-borrowers is determined by the repayment plan structure. Let IDR and $NoIDR$ denote whether there has been an IDR expansion or not, which corresponds to treatment, and let $T = \{0,1\}$ denote pre- or post-IDR expansion, respectively. Ideally, to measure the impact of the IDR expansion, I would measure the change in probability of selecting a given major among borrowers after the expansion of IDR:

$$P[J = j | A, y, l > 0, IDR, T = 1] - P[J = j | A, y, l > 0, NoIDR, T = 1]$$

[30]Intertemporally homothetic preferences ensure this is the case. CRRA preferences are one example that have this property.

However, the latter term is unobservable. By using comparisons to the non-borrowers pre- and post-treatment, I can recover an estimate for this treatment parameter under the assumption that student-borrowers would continue selecting majors at similar rates compared to non-borrowers in a counterfactual world with no expansion. That is, if

$$\begin{aligned} P[J=j|A,y,l>0,NoIDR,T=1] &= P[J=j|A,y,l>0,NoIDR,T=0] \\ &+ P[J=j|A,y,l=0,IDR,T=1] - P[J=j|A,y,l=0,NoIDR,T=0] \end{aligned}$$

I can recover an estimate of the treatment effect. The model suggests that particular *types* of majors should become relatively more attractive after the policy expansion; namely, majors that have lower initial earnings but higher growth trajectories. In addition, the extended model presented in Appendix A.2 highlights that majors with higher probability of low wage draws become relatively more attractive after IDR expansion. This motivates my use of estimated labor market outcomes for particular majors as outcome measure. Appendix A.8 reports results from the more agnostic comparison of probability of selecting specific majors as well.

However, this simple comparison has ignored the *borrowing* decision of the agent. That is, it is not necessarily true that borrowing status is independent of the repayment plan structure. In fact, borrowing here is determined by:

$$u'(w_j^1 A_i^j - R_1) w_j^1 \frac{\partial A_i^j}{\partial z_i} + u'(w_j^2 A_i^j - R_2) \frac{\partial A_i^j}{\partial z_i} = -u'(w_j^1 A_i^j - R_1) \frac{\partial R_1}{\partial l_i} - u'(w_j^2 A_i^j - R_2) \frac{\partial R_2}{\partial l_i}$$

That is, borrowers will equalize the marginal costs of borrowing and their alternative financing (recall, ability is decreasing in z_i). Note that under fixed repayment loans, the right hand side of the equation is $-u'(w_j^1 A_i^j - (1+R)l_i)(1+R)$; however, under IDR loans, the right hand side is zero. This suggests that the IDR itself induces borrowing, as it results in borrowers eschewing alternative financing in favor of borrowing.[31] Notably, the difference-in-differences estimators will conflate the treatment effect–that is, how the IDR expansion shifted major selection–with the change in composition of borrowers, since $P(l>0|A,y,T=1) \neq P(l>0|A,y,T=0)$.[32]

[31] If there is some level of "free" alternative financing, i.e. a value z^* such that $A_i^j(z) = A_i^j(0)$ for any $z \leq z^*$, then agents will select this level of z. The broader point that IDR induces a corner solution for the alternative borrowing level stands.

[32] Here, it is worth noting that if the entirety of the change in selection can be explained by observable characteristics, then this issue can be dealt with using a reweighting approach that explicitly accounts for compositional shifts. I outline this approach in Section 1.5 below.

Note that the difference-in-differences estimator can be written as:

$$
\begin{aligned}
& P[J|A,y,l_1>0,T=1] - P[J|A,y,l_1=0,T=1] \\
& \qquad\qquad - \{P[J|A,y,l_0>0,T=0] - P[J|A,y,l_0=0,T=0]\} \\
= \; & P[J|A,y,l_1>0,T=1] - P[J|A,y,l_1=0,T=1] \\
& \qquad\qquad - \{P[J|A,y,l_1>0,T=0] - P[J|A,y,l_1=0,T=0]\} \\
+ & [P[J|A,y,l_1>0,T=0] - P[J|A,y,l_1=0,T=0] \\
& \qquad\qquad - \{P[J|A,y,l_0>0,T=0] - P[J|A,y,l_0=0,T=0]\}]
\end{aligned}
$$

Here, I suppress the $\{IDR, NoIDR\}$ indicators as all observations in period $T=1$ have are exposed to IDR and vice versa for $T=0$ and $NoIDR$. The first two lines lines (left hand side of the equation) is the difference in difference estimator. I indicate with l_1 and l_0 the idea that endogenous selection means those who are borrowers in period $T=1$ may differ from those in period $T=0$. Hence, the middle two lines (first element on the right hand side of the equation) captures the "treatment effect" of the policy expansion on borrowers while the last two lines (second element on the right hand side of the equation, in square brackets) captures the idea of endogenous selection into treatment induced by the policy itself. This element is not directly observable, meaning I cannot directly account for the endogenous selection.

To address the potential for endogenous selection into borrowing in my empirical section, I use the timing of observation to preclude shifts in borrowing behavior induced by the policy expansion. Specifically, for a subset of borrowers who are in school during the policy expansion, I observe their borrowing status *prior* to the policy expansion. That is, in the model presented here, I am able to observe individuals who make borrowing decisions under the fixed repayment regime, then the policy expansion occurs, and finally they select their major under the IDR policy regime; i.e., treatment occurs between steps (ii) and (iii) of the agents' sequential decision-making process outlined above.

Finally, note that model suggests I examine individuals of similar ability and family background–which motivates my use of conditional parallel trends assumption in Section 1.5. In addition to this, if there are more than two possible majors–as is the case in reality–it may be that certain types of major-specific abilities are correlated. For example, a student who is particularly good at Engineering may also be good at Mathematics, or another STEM related field. This suggests that, after the IDR expansion, student-borrowers may be induced to change majors, but will be more likely to select into "similar" majors that have different labor market prospects. I revisit this hypothesis in Section 1.6.

In sum, the model makes the following predictions: (i) fixed repayment loans discourage selection of majors with higher chance of low remunerative returns as well as majors with steeper earnings profiles; (ii) the introduction of IDR plans specifically protects against low-monetary return labor market realizations, which encourages selection of majors associated with higher probability of low initial wage draws; and (iii) in the absence of a borrowing mechanism to smooth consumption, IDR allows partial consumption smoothing which leads borrowers to trade-off low initial wages with higher earnings trajectories.

Using these model predictions, I turn to my empirical exercises with specific measures of a major's labor market outcomes to assess the degree to which IDRs change borrower behavior. Because the largest benefits of IDR relative to fixed repayment are concentrated on very low wage draws, I specifically examine the share of graduates with a particular major who have early incomes that fall below the poverty line. Martin (2022) notes that college majors with initially low incomes also experience higher wage growth. Furthermore, if credit markets do not allow for complete consumption smoothing, the IDR expansion also offers more flexibility to borrowers in a tradeoff between immediate remunerative returns and longer run gains. In light of this, I also examine income growth as an outcome of interest.[33]

1.3.2 Efficiency

The main purpose of this paper is to document whether borrowers respond to changes in repayment plan structure. However, as noted above, there are two main effects that may lead to shifts in behavior after the introduction of IDR: insurance against low-wage realizations on the labor market and subsidization through eventual forgiveness of remaining debt. In asking whether the behavioral responses induced by IDR expansion are *desirable*, a natural question is if the responses coincide with the efficient behavioral response–that is, do borrowers act as if a benevolent social planner assigned them to a major and career path in order to maximize expected lifetime utility. The particular concern here is that borrowers may take advantage of the IDR by selecting unproductive majors (Looney (2022)).

A full examination of this question would require data on both educational decisions *and* career paths under both pre-IDR expansion regime and post-IDR expansion regime. This

[33] My model suggests that fixed repayment plans discourage risk-taking in wages and wage growth profiles. Higher income growth as an outcome of interest, then, can be motivated either by higher growth being compensation for riskier wage draws or as a risky outcome itself.

would allow for the modeling of educational and career decisions with financial constraints, in the spirit of Carneiro and Heckman (2002) and Johnson (2013). An exercise of this type is beyond the scope of this paper. However, I first note that if the subsidy benefit of IDR expansion is the main driver of behavioral shifts borrowers should select majors with lower *lifetime earnings*; whereas, if the insurance value is more important borrowers should select majors with "riskier" earnings distributions, but not necessarily lower average earnings.[34] Secondly, if major selection differs by borrower status after IDR expansion, after controlling for other plausible drivers of differences, then it is plausible that credit constraints remain an important factor in major selection.

In Section 1.6, I report that, while male borrowers are more likely to select majors that have higher initial share of low-income graduates but also higher wage growth, they do not select majors that have lower initial average wage income. Indeed, I present some evidence that the majors borrowers select into have higher average earnings at mid-career ages, 40-44, though this result is sensitive the sample frame–specifically, when I focus on borrowers in cohorts where I can measure borrowing status prior to the policy change. I argue that this is suggestive evidence that it is not the subsidy value that is the primary driver of shifting major selection among borrowers, but rather the insurance value for major associated with initially risky and variable earnings (Dynarski (2014)). However, using a multinomial choice model of major choice after IDR expansion, I find that there are significant differences in major selection by borrowing status after controlling for parental education, household income, race, high school GPA, and residency status. This suggests that, while IDR expansion does not appear to induce students to select majors associated with lower wages, there still may be credit constraints that impact which major they do select.

1.4 Data

To answer the question of how IDR expansion affected educational investments, I construct a dataset that has rich academic records linked to borrowing data. Without a detailed dataset with academic and borrowing decisions that straddles a policy change, like the one I construct here, answering this question would be significantly more difficult. This data acquisition

[34]This is an imperfect measure of the importance of the subsidy–for example, borrowers may select majors with "riskier" wage realizations and then, via their labor market behavior, extract the value of the subsidy. In effect, these individuals would select *into* the low end of the earnings distribution. However, as a first pass, I believe measuring whether IDR expansion induces selection into less productive majors *on average* can shed some light on the efficiency question.

required separate data use agreements with the provider of academic records as well as the credit bureau as well as a linking process that entailed the creation of study-specific, anonymized identifiers that were separately linked to credit bureau records and academic records, then used to create an anonymized dataset for research purposes. This siloed the research process from linking, and ensured that no academic records nor credit records were shared with identifying information.[35]

Administrative academic records are provided by a large, flagship public university. These data include course selection and performance, major declaration, permanent address ZIP code at time of enrollment, degrees awarded, and certain demographic information, including sex, age, race, residency status, parental income bins, parental education, and number of dependents in the student's family.[36] Notably, I *do not* have financial aid information for specific students. These data are linked to continuing education data using the National Student Clearinghouse (NSC) that includes other enrollment and degree information. These data are also linked to data from Experian Credit Bureau, which is a credit reporting bureau that, among other things, collects information on student loans for particular borrowers. The credit bureau data include aggregate borrowing statistics for various debt categories–including student debt–and loan performance information. In order to be included in the credit bureau data, a borrower must have (i) a credit report and (ii) be at least 18 years old at the time of observation. The credit bureau data *do not* include exact type of repayment plan for student loans, though they do contain information on repayment amounts. Due to funding limitations, I have credit bureau data from 2009, 2013, 2016, and 2019, rather than a full panel. Appendix A.3 contains more information on the linking procedure.

The linking process resulted in 97.6% match rate. Graduate cohorts 2009-2013 match rates exceed 99%. Graduate cohorts 2014-2016 match rate exceeds 98%. Match rate for 2017-2019 was 97.25%, 94.91%, and 90.46%, respectively. This suggests that non-matched observations that meet age qualifications likely did not have a credit report. I treat such individuals as non-borrowers for the purposes of analysis.

[35]Funding for this data acquisition was generously provided by the Weinberg Endowments, the Population Studies Center at the Institute for Social Research at the University of Michigan, the Rackham Graduate School at the University of Michigan, and Economics Departmental Fellowships from the University of Michigan.

[36]The family income data used available are categorical and are not available for all observations. I use average household income in the student's home ZIP code as a proxy for family income, which allows for estimation with a larger sample. The results are robust to instead using the categorical income variable.

Crucially, while the academic records cover the entire academic experience of students, credit bureau data are snapshots from four moments in time: March 2009, March 2013, March 2016, and March 2019.[37] Because of this, borrowing levels for students are not necessarily reflective of their total borrowing: for example, a student who enrolls in Fall 2008 would have borrowing observations (i) at the end of their first year in college (2009); (ii) one year post-graduation (2013); four years post-graduation (2016); and seven years post-graduation (2019).[38] Table 1.4 depicts this debt-observation timing for entry cohorts.

This complicates the measurement of borrowing for students, as I only observe complete borrowing for students graduating in 2009, 2013, 2016, and 2019. Alternatively, I can observe debt balances *after* graduation for students, though this necessitates excluding students who attend graduate programs, as I cannot differentiate between borrowing for undergraduate studies and borrowing for graduate studies. In my main specifications, I use binary borrowing status as my explanatory variable, with debt measured *after* graduation.

I use American Community Survey (ACS) data from the students entry year as a measure of their information set on labor market outcomes (Ruggles et al. (2022)).[39] In particular, I define ages 25-29 as early-career and calculate early-career wage growth from 25-29 to 30-34. To capture longer-run labor market outcomes, I also estimate a Mincer earnings function separately for each major and calculate earnings for prime working years, 25-54, as well as earnings variance in mid-career, ages 40-44. I calculate mean wage income, standard deviation in wage income, share of graduates in particular major with incomes less than the Federal Poverty Line (FPL), income growth over early career, and labor force participation statistics by major.

These measures are used to order discrete major decisions into quantifiable outcomes and are selected because the motivating theory suggests that the expansion of IDR should lead borrowers to make "riskier" investments. In particular, the model suggests borrowers would

[37] These dates were selected for two reasons: (i) loans do not appear on credit reports until after their disbursement. By selecting a date at the end of the academic year, I capture all borrowing up through that particular academic year. And (ii) the first three years immediately pre-date the IDR expansions. The Income Based Repayment (IBR) plan became available after the Title 34 of the Code of Federal Regulations was updated in July 2009. Similarly, Pay As You Earn (PAYE) became available in December 2012 and Revised Pay As You Earn (RePAYE) became available in December 2015. While PAYE and RePAYE enrollment was announced in December, they did not enter the Code of Federal Regulations until July and the National Student Loan Data System (NSLDS) reports that the first loans enrolled in RePAYE after this date.

[38] Assuming the student graduates in four years.

[39] ACS data on majors begins in 2009; hence, for entry cohorts prior to 2009, I use 2009 data. Results are robust to using different ACS years as measures of labor market outcomes by major.

Table 1.4: Year of debt observation relative to year of entry (academic cohort)

Entry Year	Year after entry (debt observation marked with "X")									
(Academic Cohort)	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
2004-2005					X				X	
2005-2006				X				X		
2006-2007			X				X			X
2007-2008		X				X			X	
2008-2009	X				X			X		
2009-2010				X			X			X
2010-2011			X			X			X	
2011-2012		X			X			X		
2012-2013	X			X			X			n/a
2013-2014			X			X			n/a	n/a
2014-2015		X			X			n/a	n/a	n/a
2015-2016	X			X			n/a	n/a	n/a	n/a
2016-2017			X			n/a	n/a	n/a	n/a	n/a
2017-2018		X			n/a	n/a	n/a	n/a	n/a	n/a
2018-2019	X			n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note: academic years 2021-2022 and beyond are not observed and marked "n/a"

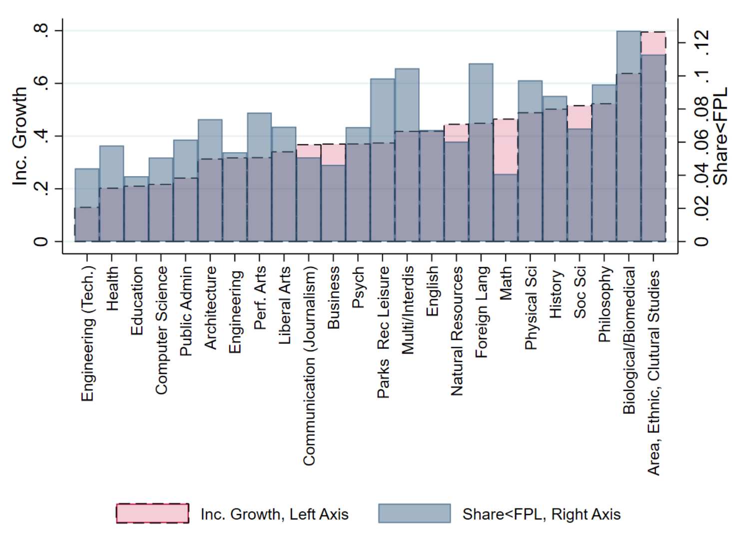
select majors that are associated with lower wage draws (more weight in the bottom of the wage distribution) and higher income growth, since IDRs allow more flexibility in trading off immediate remunerative returns for larger future gains. Recognizing that particular labor market outcomes may differ by sex (see, e.g., Altonji et al. (2016)), I calculate the associated labor market outcomes for each major separately by sex. Figure 1.3 highlights the income growth and share under the FPL for each major. Appendix A.4 contains a detailed table of labor market outcomes by major.

Table 1.5 reports summary statistics of the undergraduate sample by graduating cohort. Each of these characteristics are determined prior to postsecondary enrollment. Because the reported family income is categorical, not available for all students, and not adjusted for inflation, I report here the mean household income from the student's ZIP code as a proxy for family income.[40] There is relatively little movement in these characteristics. This suggests that the characteristics of the student body *as a whole* have not changed dramatically over this time frame.[41]

[40] I use this proxy for family income in my weighting equations. In alternative specifications, I use the categorical family income variable. The results are qualitatively similar.

[41] There was a large, well-publicized effort on the part of this particular University to attract more low-income instate students through targeted aid and advertising. This program began with the 2016 entering cohort, the vast majority of whom graduated after my sample ends: only 1.6% graduated in 2018 or 2019.

Figure 1.3: Selected Labor Market Outcomes by Major



Left axis reports income growth between early- (age 25-29) and mid-career (age 30-34); right axis reports share with that major with incomes under the FPL in early-career. Data are from ACS for student's entry year.

However, Table 1.6 reports the real average loan debt and share of borrowers by family income decile. There is a clear pattern of decreasing borrowing–on both the extensive and intensive margins–for lower income families. This follows a broader pattern of increasing cost progressivity at public research universities, identified by Cook and Turner (2022). Indeed, Table 1.7 highlights the change in the average net price charged to students at the university within particular income ranges. The shift in net price is driven largely by increases in institutional aid: at the university, gross charges increased by 78.5% between 2008-2019 while institutional need-based grants and scholarships increased by 350.5%.[42] This suggests that, while the characteristics of the overall population are relatively stable, the selection into borrowing status may differ over time. Appendix A.5 compares characteristics of borrowers and non-borrowers over time.

Excluding these individuals from analysis does not significantly change results.

[42]This outstrips the national average growth in institutional aid, which grew 94.3% (Ma and Pender (2021)).

1.5 Empirical Methodology

The theoretical model suggests a difference-in-differences analysis: while the expansion of the IDR induces borrowers to shift towards majors with riskier labor market outcomes or steeper early career earnings profiles, it should have *no* effect on non-borrowers. Hence, the 2013 policy expansion acts as a "treatment," where exposure to treatment is determined by borrowing status. In this setting, I use graduation cohorts 2009-2013 as the "pre-treatment" groups and cohorts 2014-2019 as the "post-treatment" groups, with the actual treated and untreated subgroups determined by borrowing. This is a repeated cross sections setting, the cross sections determined by graduation cohort and treatment exposure for the pre-treatment cross sections determined by borrowing status.

The difference-in-differences analysis assumes that the non-borrowers are a suitable substitute for capturing the trends of the borrowers in the absence of the policy expansion–that is, they identify the unobservable counterfactual. While borrowers and non-borrowers are not randomly selected, this assumption states that, on average, the *trend* in the (unobservable) major selections borrowers make in the absence of treatment does not differ from that of the non-borrowers, which is observable. As highlighted in Appendix A.5, these groups differ in a number of dimensions: borrowers are less white, less affluent, come from families with less education, and enter college with fewer academic credits. Differences in these characteristics could plausibly lead to differential trends in the groups, even without the policy expansion, e.g., if students from poorer families react differently to changes in labor market conditions.[43] Furthermore, the theory outlined in Section 1.3 suggests that students *of similar abilities and means* are the comparison of interest. In light of these potential issues, I also invoke the *conditional* parallel trends assumption outlined by Heckman et al. (1997) and Abadie (2005).[44]

In my context, there are three additional concerns I highlight: first, the standard models with conditional parallel trends assume–either explicitly or implicitly–that there is no compositional change in the treatment and control groups. That is, the *selection mechanism* into treatment is constant across time. Here, treatment exposure is determined by borrowing status, which is in part determined by the other resources made available to students.

[43] Testing for pre-trends is frequently used as a substitute for a test of non-parallel trends over treatment.

[44] If the unconditional parallel trends assumption holds, then the conditional parallel trends assumption also holds; hence, this exercise can be thought of as test of the standard difference-in-differences model.

Table 1.5: Summary Statistics by Graduating Cohort Year

Grad. Cohort	Instate	Female	White	Underrep. Minority	Parent Col.+	Mean HH Inc. (ZIP – 2019 $)	Took SAT	SAT Math	Took ACT	ACT Math	HS GPA
2009	0.612	0.509	0.720	0.129	0.837	$117,500	0.739	28.351	0.633	668.482	3.734
2010	0.641	0.510	0.727	0.129	0.854	$117,100	0.764	28.732	0.608	670.651	3.752
2011	0.634	0.518	0.724	0.115	0.846	$117,700	0.778	28.821	0.562	672.245	3.763
2012	0.669	0.513	0.736	0.112	0.847	$115,100	0.827	29.005	0.428	677.934	3.760
2013	0.645	0.509	0.775	0.098	0.847	$116,000	0.823	29.226	0.411	682.705	3.760
2014	0.610	0.497	0.757	0.099	0.864	$118,400	0.807	29.384	0.385	685.069	3.763
2015	0.600	0.509	0.701	0.101	0.876	$118,500	0.816	29.503	0.356	689.042	3.788
2016	0.590	0.504	0.704	0.099	0.885	$116,800	0.805	29.603	0.356	689.676	3.800
2017	0.592	0.520	0.708	0.097	0.888	$116,600	0.809	29.967	0.338	697.921	3.818
2018	0.561	0.501	0.728	0.104	0.890	$118,800	0.813	29.972	0.323	700.720	3.815
2019	0.574	0.511	0.701	0.118	0.908	$116,200	0.845	30.209	0.275	701.977	3.826

Table 1.6: Avg. Loan Debt (2019 $) & Share of Borrowers by Family Income Decile

Grad Cohort	Bottom Decile	2^{nd} Dec.	3^{rd} Dec.	4^{th} Dec.	5^{th} Dec.	6^{th} Dec.	7^{th} Dec.	8^{th} Dec.	9^{th} Dec.	Top Decile	Overall
	A. *Average Loan Debt (2019$)*										
2009	$32,593	$33,535	$36,528	$34,316	$35,362	$32,623	$35,912	$42,797	$29,347	$26,941	$34,360
2013	$28,173	$33,957	$31,197	$33,792	$32,330	$30,233	$32,012	$30,555	$31,277	$34,609	$31,684
2016	$24,529	$26,323	$29,652	$25,357	$25,582	$28,850	$27,967	$30,481	$21,101	$23,554	$26,437
2019	$21,623	$23,832	$26,173	$25,709	$29,478	$26,604	$22,547	$28,570	$27,356	$38,078	$25,904
	B. *Share borrower*										
2009	0.722	0.569	0.546	0.482	0.500	0.429	0.374	0.320	0.288	0.231	0.458
2013	0.733	0.653	0.614	0.515	0.530	0.470	0.392	0.355	0.296	0.229	0.485
2016	0.699	0.620	0.553	0.513	0.360	0.400	0.319	0.307	0.275	0.164	0.434
2019	0.539	0.565	0.440	0.416	0.439	0.337	0.296	0.297	0.261	0.220	0.391

Table 1.7: Average Net Price by Family Income Range

	Family Income Range				
Grad. Cohort	<$30,000	$30,000-$48,000	$48,000-$75,000	$75,000-$110,000	$110,000+
2008	$6,100	$8,400	$13,200	$18,100	$20,100
2009	$5,300	$7,400	$12,400	$18,400	$21,500
2010	$4,800	$7,100	$13,100	$18,300	$22,100
2011	$5,400	$7,900	$12,300	$18,900	$23,200
2012	$5,500	$9,400	$13,600	$18,600	$23,200
2013	$5,500	$7,700	$11,400	$18,300	$23,300
2014	$3,400	$6,600	$9,800	$17,200	$23,700
2015	$2,700	$5,900	$10,100	$16,700	$24,100
2016	$3,300	$5,600	$9,900	$17,700	$25,000
2017	$4,000	$6,300	$11,000	$16,900	$25,000
2018	$3,000	$5,900	$10,000	$18,000	$26,100
2019	$2,700	$5,700	$10,500	$17,200	$26,500

Source National Center for Education Statistics (NCES) (2021). Note: Income ranges are in nominal dollars, net price in constant 2019 dollars

As highlighted in Tables 1.6 and 1.7, the university I use in my empirical analysis expanded institutional aid–especially for students from less-affluent families–over the timeframe I examine. As such, there has been a shift in composition of borrowers. To address this concern, I extend the reweighting approach of Abadie (2005) to explicitly account for the changing selection mechanism over time. This is a novel contribution to the applied econometric literature on difference-in-differences with compositional shifts.[45]

Second, it is possible that the policy expansion *itself* induced students into treatment–that is, since the policy effectively makes borrowing cheaper relative to potential alternative forms of financing, students may decide to borrow *because* of the policy expansion.[46] Such endogenous selection into treatment could impact identification if, for example, it is partially driven by the expected outcomes themselves. To address this issue, I make use of the fact that my data allows me to observe borrowing status at different points in time: using a sub-sample of cohorts for whom I observe borrowing *before* the policy expansion, I can limit the ability of students to change their borrowing strategy in response to the policy.[47]

[45] To the best of my knowledge, Hong (2013) is the only other paper which considers such settings. He extends the Heckman et al. (1997); Heckman et al. (1998) approach that estimates the outcome using matching estimators. I compare these approaches in Appendix A.6.

[46] An example of an alternative form of financing is working while in school: Keane and Wolpin (2001) highlight that this is an important source of financing for credit-constrained students.

[47] For my main specifications, I use a measurement of debt from *after* graduation, in order to more accurately capture total borrowing. However, this means that treatment cohorts 2014-2019 use debt measurements from

Finally, there are other potential confounding factors which may impact identification. Broader policy changes in student loans–e.g. the increase in federal loans limit and the decrease in interest rates–could also drive shifts in behavior. Additionally, the increase in institutional aid could conceivably lead to an income effect for recipients of this aid, which could also lead to shifts in behavior. To address these potentially confounding factors, I also report results from robustness tests that account for (i) the increase in federal loan limits in 2008 by limiting the sample to individuals who enroll after this increase; (ii) estimating the repayment rate under a standard, fixed repayment plan, which accounts for changes in interest rates over time; and (iii) limiting my analysis to a subsample of individuals who are the least likely to be impacted by changes in institutional aid policy.

The remainder of this section describes the difference-in-differences strategy in more detail.

1.5.1 Difference-in-Differences analysis

The IDR expansion here can be thought of as an exogenous change to *borrowers* choice set, which could lead to changes in investment choices. To test this, the parameter of interest is the *Average Treatment Effect on the Treated* (ATT):

$$ATT = E(Y(1,1) - Y(0,1)|D = 1, T = 1)$$

Where $Y(D,T)$ is the potential outcome for treatment exposure group $D = \{0,1\}$ in time period $T = \{0,1\}$. Here, *treatment* is the policy expansion while *treatment exposure* is determined by borrowing status. In period $T = 1$, all individuals in group $D = 1$ are treated, leading to an inference problem: the second term of the ATT, $E(Y(0,1)|D = 1, T = 1)$, is unobserved. Since the policy expansion does not affect non-borrowers, a plausible identification strategy makes use of the parallel trends assumption, Equation A1.

$$\begin{aligned} &E(Y(0,1)|D = 1, T = 1) \\ &\quad = E(Y(1,0)|D = 1, T = 0) + E(Y(0,1)|D = 0, T = 1) - E(Y(0,0)|D = 0, T = 0) \end{aligned} \tag{A1}$$

after the 2013 policy expansion (either 2016 or 2019) to identify borrowing status, potentially leading to endogenous selection into treatment. As highlighted in Table 1.4, for certain cohorts, I can use an earlier measure of debt to identify borrowing status. In particular, I can use 2013 debt measures for treated cohorts enrolled by that time. This precludes using the entire treated cohorts 2014-2019, but does allow for a sub-sample of treated cohorts in this robustness analysis.

Under the parallel trends assumption, the ATT is identified by a simple comparison of outcomes for different groups over time:

$$ATT = E(Y|D=1,T=1) - E(Y|D=0,T=1) - \{E(Y|D=1,T=0) - E(Y|D=0,T=0)\}$$

Which implies a simple comparison of the empirical analogues to the above expectations serves as an estimate of the ATT. In my context, I define the *pre-treatment* period (T=0) as graduating cohorts 2009-2013 and the *post-treatment* period (T=1) as graduating cohorts 2014-2019. Again, the *treatment exposure* is defined by borrowing status (D=1) with non-borrowers being the control-sample (D=0). A regression approach recovers an estimate of the ATT by estimating the following relationship:

$$Y_{i,c} = \alpha_0 + \alpha_d \mathbb{1}(D_{i,c}=1) + \alpha_t \mathbb{1}(T_{i,c}=1) + \gamma \mathbb{1}(D_{i,c}=1, T_{i,c}=1) + \varepsilon_{i,c}$$

Where $Y_{i,c}$ is the outcome of interest for individual i of graduating cohort c, and $\mathbb{1}(\cdot)$ is an indicator function equal to one if the argument is true and zero otherwise. Here $\hat{\gamma}$ is the estimate for the ATT. In Section 1.6, I call this approach the standard difference-in-differences estimate.

The standard difference-in-differences estimate is valid only insofar as the unconditional parallel trends assumption, A1, is valid: that is, if the trends in the non-borrower sample are a good proxy for the (unobservable) trends the borrower sample would have taken in the absence of the policy expansion. However, as noted above, borrowing status is not randomly selected and borrowers and non-borrowers differ on a number of margins. Borrowers are less white, come from less affluent families, have parents with less educational achievement, and enter college with fewer academic credits. If these differences in characteristics *also* lead to differential trends, then the unconditional parallel trends assumption is violated.

Alternative approaches, developed by Heckman et al. (1997), Heckman et al. (1998), and Abadie (2005), make use of the *conditional* parallel trends assumption, Equation A1′.

$$\begin{aligned} &E(Y(0,1)|X, D=1, T=1) \\ &= E(Y(1,0)|X, D=1, T=0) + E(Y(0,1)|X, D=0, T=1) - E(Y(0,0)|X, D=0, T=0) \end{aligned} \tag{A1′}$$

That is, parallel trends holds only *after* conditioning on characteristics X. As shown by Abadie (2005), under A1′, the ATT is identified using re-weighted outcomes of different groups over time.

$$ATT = E(\frac{P(D=1|X)}{E(D=1)}\varphi_0 Y)$$

Where,

$$\varphi_0 = \frac{T-\lambda}{\lambda(1-\lambda)}\frac{D-P(D=1|X)}{P(D=1|X)P(D=0|X)} \quad \text{and} \quad \lambda = P(T=1)$$

This approach requires a two-step procedure: first, a researcher estimates the propensity score; then, they estimate the empirical analogue to the above weighted expectation. In my setting, I parametrically estimate the propensity score using a logit model of borrowing status, using parental education, the number of other dependents in the students' household, a quadratic in household income (proxied by home ZIP-code income), an indicator for whether the student is white, an indicator for residency status, an indicator for if they enter college with enough credits to be considered a second- or above-year student, and their high school GPA. These variables are selected because (i) they all are determined *prior* to attendance and (ii) they include measures of a student's means and ability.[48] I then bootstrap this two-step estimation procedure in order to conduct inference. In Section 1.6, I call the estimates from this procedure the "standard weights" results. A particular attraction of this approach is, if there are elements of X that are unobservable but this unobservable selection is time-invariant, the approach still recovers an unbiased estimate of the ATT.

However, the evidence that institutional aid has shifted borrowing on the part of different *types* of students suggests that the mechanism that determines treatment exposure–that is, borrowing status–has changed. Such settings are usually ruled out–either implicitly or explicitly–in difference-in-differences analyses.[49] Econometrically, this implies that the propensity score changes over time: $P(D=1|X,T=0) \neq P(D=1|X,T=1) \neq P(D=1|X)$. Relatively little work has been done on this problem in difference-in-differences settings. Hong (2013) considers the problem explicitly, and notes that the standard reweighting approach proposed by Abadie (2005) is not sufficient for identification.

Here, I argue that properly redefined weights in the spirit of Abadie (2005) *can* recover the ATT. In particular, by estimating the propensity scores separately for each time period, and using a common integration term that is specific to the treatment period, the ATT is given by:

$$ATT = E\left[\frac{P(D=1|X,T=1)}{P(D=1|T=1)}\omega Y\right]$$

[48]The motivating theory suggests that students of similar means and ability should make similar decisions, with differences attributable to either idiosyncratic preferences or driven by borrowing status.

[49]See, e.g. Sant'Anna and Zhao (2020) assumption A1.b.

Where

$$\omega = \frac{T-\lambda}{\lambda(1-\lambda)} \times \frac{D - TP(D=1|X,T=1) - (1-T)P(D=1|X,T=0)}{TP(D=1|X,T=1)P(D=1|X,T=0) + (1-T)P(D=1|X,T=0)P(D=0|X,T=0)}$$

The details of this approach are given in Appendix A.6. Again, I parametrically estimate the propensity score as in the standard weighting case, but I separately estimate by treatment time-period. Then I calculate the empirical analogue to the weighted expectation above. I bootstrap this estimation procedure in order to conduct inference. In Section 1.6, I call the results from this method the "compositional shift weights."

An advantage of this approach is that it explicitly accounts for the changing selection mechanism over time. However, as noted in Hong (2013), selection on unobservables remains an issue. In particular, because the overall integration weight, $\frac{P(D=1|X,T=1)}{E(D=1|T=1)}$, differs from the weights ω for individuals from pre-treatment period, even time-invariant selection on unobservables is an issue. In recognition of the relative strengths and weaknesses of each approach, I present results from both the standard weighting approach and the approach that accounts for the changing selection mechanism. I argue that, given each approach results in significant shifts in borrower behavior, there is a real response to the policy expansion in terms of optimal human capital investment.

1.6 Results

1.6.1 Main Results

Tables 1.8 and 1.9 reports the main results. Columns (1)-(3) report the results from a standard 2×2 difference-in-differences analysis for the entire sample, with graduating cohorts 2009-2013 considered *pre-treatment* and graduating cohorts 2014-2019 considered *post-treatment*. Column (1) reports the results from a standard difference-in-differences estimation, which is valid under the unconditional parallel trends assumption. Column (2) reports the results from the standard weighting procedure suggested by Abadie (2005), which accounts for conditional parallel trends with no compositional shifts in the treatment and control groups, and column (3) reports the results for my proposed extension to the reweighting technique that explicitly accounts for compositional shifts. Distributions for the estimated propensity scores using both methods are in Appendix A.7. Panel A. reports the results for men and Panel B. reports the results for women.

From Table 1.8, there is no shift in major selection when measured by the majors' average income or standard deviation of income in early-career labor market (ages 25-29). However, there *is* a more subtle change in the distribution of labor market returns associated with majors. In particular, after the policy expansion, borrowers select majors with 1.8-2.6 percentage point higher income growth between early- and mid-career, but also 0.25-0.37 percentage point higher poverty rates. For context, this represents a 4% to 5.7% higher income growth and 3.2%-4.7% higher poverty rate, relative to the average labor market returns as measured by major selection. Furthermore, while there are no significant movements in the early-career first or second moments of the income distribution associated with these major selections, there is some evidence of an increase in the variance of mid-career income (ages 40-44) associated with these major selections. Finally, the point estimates for average career earnings over the entire prime-working age are positive, though noisily measured. What these estimates do imply is that borrowers *do not* select into majors with significantly *lower* lifetime earnings in response to the IDR expansion, a point which I revisit in my discussion of efficiency below. Here it should be noted that the average career earnings are about \$2.5 million for men (\$1.7 million for women), so these estimates rule out shifts into majors that have larger decrease than 1.4% decline in prime-age earnings among men.

This suggests that, after the policy change, borrowers are more willing to sort into majors associated with low initial remunerative labor market returns, but higher income growth. In particular, this effect is concentrated at the bottom of the wage distribution–that is, it is not the *average* returns that shift, but the probability of especially bad labor market draws that increase.

This is consistent with the motivating theory that IDR specifically insures against bad labor market draws and, in the absence of complete credit markets in the labor period, allows borrowers the flexibility to trade off lower initial returns for higher growth. Notably, this evidence suggests that these earnings differences result in similar average lifetime earnings, though higher earnings variance at later ages. That is, it suggests that male borrowers shift into majors with higher risk of low initial earnings *and* riskier long-run earnings.

To examine *which* majors are driving these results for men, I also estimate a multinomial logit model where the latent function for each major choice follows the structure,

$$Y_{i,c}^{j} = \alpha_0^{j} + \alpha_d^{j}\mathbb{1}(D_{i,c} = 1) + \alpha_t^{j}\mathbb{1}(T_{i,c} = 1) + \gamma^{j}\mathbb{1}(D_{i,c} = 1, T_{i,c} = 1) + \varepsilon^{j}$$

Table 1.8: Difference-in-Differences Estimates, Men

	2009-2019			2009-2015		
	(1) Std. DiD	(2) Std. Wts.	(3) Comp. Shift Wts.	(4) Std. DiD	(5) Std. Wts.	(6) Comp. Shift Wts.
Avg. Income (early career)	-\$576*	\$26	-\$582	-\$62	-\$479	-\$396
(s.e.)	(333)	(801)	(470)	(481)	(1148)	(704)
S.D. of Inc. (early career)	\$203	\$682	\$188	-\$30	-\$55	\$188
(s.e.)	(299)	(643)	(412)	(408)	(889)	(607)
Income Growth (early career)	2.01 p.p.***	2.56 p.p.***	1.82 p.p.***	2.13 p.p.***	2.49 p.p.**	2.78 p.p.***
(s.e.)	(0.48)	(0.92)	(0.67)	(0.74)	(1.21)	(0.96)
Poverty (early career)	0.28 p.p.***	0.37 p.p.**	0.25 p.p.**	0.24 p.p.*	0.31 p.p.	0.35 p.p.**
(s.e.)	(0.09)	(0.15)	(0.12)	(0.14)	(0.21)	(0.17)
Avg. Prime Age Income (25-54)	\$4,825	\$32,779	\$1,221	\$29,226**	\$49,785	\$16,584
(s.e.)	(11,383)	(35,070)	(17,787)	(13,891)	(39,686)	(22,118)
S.D. of Inc. (mid-career, 40-44)	\$1,356*	\$3,064*	\$1,678*	\$2,299**	\$2,503	\$3,020**
(s.e.)	(699)	(1,629)	(1,010)	(934)	(2144)	(1413)
		n=17,895			*n=10,017*	

$*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .10$; Standard errors are calculated via bootstrapping estimation procedure 1,000 times. Outcomes are labor market realizations for major selection, measured using ACS from year of student entry. Early career income measures are calculated for ages 25-29; income growth over early career calculated using changes in average incomes from 25-29 and 30-34. Mid-career measures are calculated for ages 40-44. See Section 1.4 for more detail. Borrower status is measured post-graduation here.

Table 1.9: Difference-in-Differences Estimates, Women

	2009-2019			2009-2015		
	(1) Std. DiD	(2) Std. Wts.	(3) Comp. Shift Wts.	(4) Std. DiD	(5) Std. Wts.	(6) Comp. Shift Wts.
Avg. Income						
(early career)	-$358	-$320	$4	$509	-$468	$290
(s.e.)	(238)	(765)	(402)	(334)	(1116)	(525)
S.D. of Inc.						
(early career)	-$445***	-$440	-$224	$45	-$463	$17
(s.e.)	(151)	(505)	(268)	(195)	(721)	(327)
Income Growth						
(early career)	-0.34p.p.	-1.03 p.p.	-0.63 p.p.	0.15 p.p.	-0.50 p.p.	-0.10 p.p.
(s.e.)	(0.38)	(0.75)	(0.52)	(0.51)	(0.96)	(0.67)
Poverty						
(early career)	-0.06 p.p.	-0.21 p.p.	-0.13 p.p.	-0.21 p.p.*	-0.34 p.p.	-0.22 p.p.
(s.e.)	(0.08)	(0.16)	(0.11)	(0.11)	(0.23)	(0.15)
Avg. Prime Age						
Income (25-54)	-$7,225	-$16,430	-$3,376	$13,220	$5,253	$11,739
(s.e.)	(7,594)	(28,366)	(13,915)	(8,863)	(32,320)	(16,007)
S.D. of Inc.						
(mid-career, 40-44)	-$135	-$759	-$185	$1,192*	-$362	$655
(s.e.)	(507)	(1301)	(790)	(644)	(1701)	(951)
		n=18,219			*n=10,318*	

$*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .10$; Standard errors are calculated via bootstrapping estimation procedure 1,000 times. Outcomes are labor market realizations for major selection, measured using ACS from year of student entry. Early career income measures are calculated for ages 25-29; income growth over early career calculated using changes in average incomes from 25-29 and 30-34. Mid-career measures are calculated for ages 40-44. See Section 1.4 for more detail. Borrower status is measured post-graduation here.

Where j indexes each major and the explanatory variables are as defined in the simple difference-in-differences approach above. This unordered analysis of major selection can be thought of as an agnostic examination of how probability of selecting a particular major shifts after the policy expansion among borrowers. The results for this approach are presented in Appendix A.8. The largest relative shifts in probability of selection are increases in History (1.25 p.p.), Biological/Biomedical Sciences (1.17 p.p), and Psychology (0.98 p.p.) and decreases in Computer Sciences (-2.15 p.p.), Business (-1.63 p.p.), and Social Sciences (-1.14 p.p.). It should be noted that there is no clear pattern in relative changes in major selection along Science, Technology, Engineering, and Mathematics (STEM) majors as a group, a point I revisit below.

Interestingly, Table 1.9 shows there is no such effect for women. The entirety of this effect is concentrated among men shifting investment behavior after the policy expansion. Zafar (2013) reports that, in their selection of majors, men are more likely to be motivated by the pecuniary returns to a major on the labor market whereas women are more motivated by non-pecuniary returns during college. This evidence is consistent with those findings in that it is men who are sensitive to the pecuniary returns *accounting for loan repayment plan* in selecting their majors.

There is a particular type of selection on unobservables that could potentially bias these results. In particular, treatment exposure (borrowing status) is possibly *endogenous* to the policy expansion itself. Cadena and Keys (2013) highlight a significant minority of low-income students who are eligible for subsidized loans turn them down. In 2012, 37.3% of students who applied for financial aid reported turning down a federal loan, and 24.4% of borrowers reported accepting less than the maximum amount they were offered (National Center for Education Statistics (NCES) (2012)). Because IDR makes loans more attractive due to the insurance value, this suggests a sizable minority of students *could* adjust their borrowing behavior in response to the policy expansion.

In order to account for this potentially endogenous selection into treatment exposure, as determined by borrowing status, I limit my sample to graduating cohorts 2009-2015 and use a measurement of debt from *before* the policy expansion.[50] This should limit the ability of borrowers to endogenously respond to the policy. Columns (4)-(6) repeat the above exercises on this sub-sample. I do not find evidence that the effect is driven by endogenous selection; in fact, the point estimates for sorting into majors with higher income growth and higher

[50]See Section 1.4 for a description of timing of debt measurement relative to college entry.

share of graduates below the federal poverty line are larger for this subsample, though not significantly so.

Furthermore, here I observe some marginal evidence for increased career earnings over prime working ages (ages 25-54) as well as increased variance in earnings. This suggests that the majors borrowers select have initially lower remunerative returns–reflected in higher risk of very low earnings–but steeper growth trajectories that translate into higher, but also more variable, long-run returns. These long run results are consistent with a standard appetite for risk story in the sense of Rothschild-Stiglitz risk tolerance models. Specifically, the IDR expansion led borrowers to make investment decisions with riskier life-time earnings returns. However, given the sensitivity of this result to specification, as well as the consistency of the shifts in measures of early-career returns, I argue that the evidence is most suggestive of insurance for borrowers over their early career. Again, this is consistent with the model that emphasizes the consumption smoothing benefits of the IDR program over the early career, when borrowers are most likely to experience low and variable incomes.

1.6.2 Leveraging heterogeneity in time and borrowing levels

The main results presented in Tables 1.8 and 1.9 use the standard 2×2 difference-in-differences set-up. However, it is possible that there is variation in the effect of the policy expansion on human capital investments across time *and* borrowing level.

In particular, the motivating theory presented in Section 1.3 and detailed in Appendix A.2 notes that, under fixed repayment, the degree of "risk avoidance" is *increasing* in loan balance. That is, for two otherwise identical borrowers with different loan balances, the borrower with higher loan balances has *lower* expected utility from a risky wage draw than the borrower with lower balances. On the other hand, the IDR eliminates this relationship between loan balance and risk avoidance for low wage draws, due to the subsidy value of loan forgiveness under IDR. In addition, the risk of default increases with loan balance under the fixed repayment, further discouraging selection of a major associated with especially low wage draws. Again, IDR eliminates this risk.

To assess heterogeneity across time and borrowing level, I estimate an event-study style regression with continuous treatment levels.

$$Y_{i,c} = \sum_{c=2009}^{2019} \gamma_c LoanPct_{i,c} + \phi_c + \beta_c X_{i,c} + \varepsilon_{i,c}$$

Where $LoanPct_{i,c}$ is the percentile of loan balance for individual i within cohort c.[51] This measurement constrains $LoanPct \in [0,1]$ and is calculated separately by graduating cohort. This allows a consistent measurement of borrowing across cohorts–given the data limitations regarding debt observation timing, see Section 1.4. Here, I also use a continuous-version of the reweighting technique that accounts for a changing selection mechanism.[52]

In particular, the regression is weighted using $\Psi_{i,c} = \frac{f(d|X,C=2013)}{f(d|C=2013)} \frac{1}{P(C=c)f(d|X,C=c)}$. This is similar to the weights suggested by the semiparametric reweighting that accounts for compositional shifts with two key differences: first, the overall integration term reweights to the base year distributions–that is, 2013–rather than the treatment year. This is so the weights are well-defined for a multi-treatment period regression. Second, the weights are strictly positive. This is because the parametric structure of the estimation equation already estimates changes in the relationship over time, negating the need for negative weights.

The weights are estimated using the quantile-binning approach presented in Naimi et al. (2014). In particular, the continuous treatment exposure variable, $LoanPct_{i,c}$, is split into categories: non-borrowers, and deciles of borrowing status. Then, a multinomial logit estimates the probability of each borrower having loan balances within a particular categorical grouping. This is used to estimate the probability density function, $f(d|X, C = c)$. Estimated propensity score distributions are in Appendix A.7.

Recent work on difference-in-differences with continuous treatments by Callaway et al. (2021) cautions the interpretations of such models. In particular, stronger assumptions are required to identify causal parameters relating different levels of debt to each other (rather than the non-borrower, control group). Furthermore, the estimated parameters are a weighted sum of these causal parameters, but the weights do not correspond to the distribution of the treatment variable, in general.[53] As such, while I use these results as auxiliary results to those presented above, I do not interpret them as strong evidence independent of the main results.

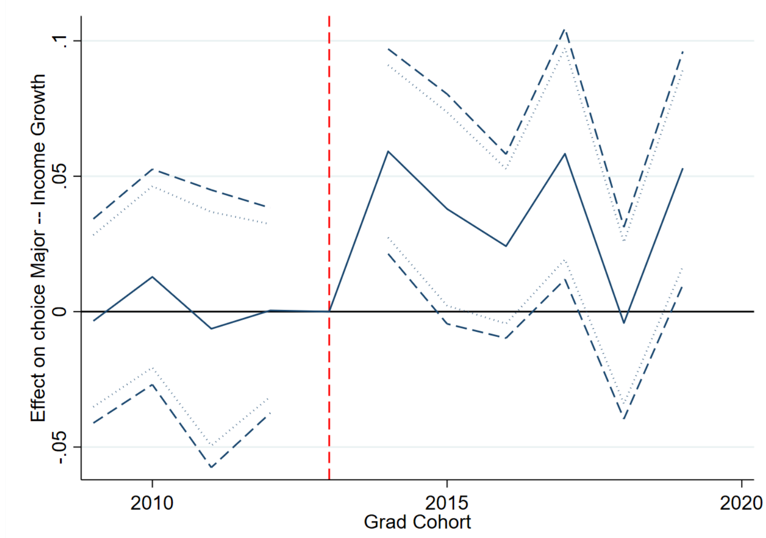
Figures 1.4 and 1.5 display the results of these exercises for men. In both cases, the estimated

[51]Note: because my sample is repeated cross sections of graduating cohorts, and not a true panel, this regression is identifiable with no constant term. In practice, I will include a constant and exclude 2013 from estimation, for ease of interpretation.

[52]See Appendix A.6 for details on the continuous version of reweighting.

[53]In my context, all weights are positive and excess weight is placed near the expectation of the treatment variable and away from the tails, relative to the distribution of the treatment exposure variable, $LoanPct_{i,c}$.

Figure 1.4: Effect on Major Selection–Income Growth for Graduates, Men



Dotted line indicates 90% CI, dashed line indicates 95% CI; Regressions are weighted using regression-version of compositional shift weights with robust standard errors.

Figure 1.5: Effect on Major Selection–Share of Graduates <FPL, Men

Dotted line indicates 90% CI, dashed line indicates 95% CI; Regressions are weighted using regression-version of compositional shift weights with robust standard errors.

effects up to twice as large as the estimated effect from the 2×2 difference-in-differences model, suggesting that there is a significant amount of heterogeneity. In particular, men select majors with between 2.5-6 percentage point higher income growth and 0.2-0.8 percentage point higher poverty rates relative to 2013 pre-policy selections.[54] These results are also noisy; however, taken in conjunction with the main results, I interpret this as supporting evidence that borrowers change their optimal investment behavior in response to the policy expansion.

1.6.3 Potential confounding effects & Robustness

In addition to the possibly endogenous selection in response to the policy change considered above, there are other potentially confounding factors that could affect the ATT estimates of the policy expansion. Here, I briefly address several potential issues.

There were two broader changes to the student loans policy environment during this time. First, federal loan limits were increased in 2008, see Table 1.1. This ended a period when private loan issuance made up a significant portion of the market.[55] Because borrowers who substitute federal loans for private loans may change behavior independently of any effect of IDR, I redo my analysis after excluding students who were in school prior to the loan limit increase. Including only students who entered college after the federal loan limit was raised does not change the qualitative results. Table 1.10 reports difference-in-differences estimates for this sub-sample.

A second change in federal policy was a general decline in interest rates over this period, see Table 1.2. Because interest rates directly impact the repayment *amount* under fixed repayment and repayment *length* under IDR, it is possible that this would impact major selection independently of the policy expansion. To explore this issue, I estimate *repayment amount* under a fixed, 10-year repayment plan and estimates the continuous event study using this as the explanatory variable. Unfortunately, without more detailed records on specific timing of loan disbursement and loan disbursement amount, this estimate is a noisy measure for true repayment amount. Nevertheless, as Figures 1.6 and 1.7 report, the qualitative results remain similar. Note, the measure here is the effect of a $100 increase in estimated monthly repayment. Among borrowers, the average estimated monthly repayment is $162 and the median monthly repayment is $155. This shows that, after controlling for the

[54] Because the explanatory variable is scaled to $[0, 1]$, this would be the effect of moving from a non-borrower to the highest level of borrowing.

[55] Private loan issuance peaked at 25% of all loan issuance in 2006-07 and 2007-08 before falling to around 10%, where it remains (Ma and Pender (2021)).

Table 1.10: Difference-in-Differences Estimates (select cohorts)

	2012-2019		
	A. *Men*		
	(1)	(2)	(3))
	Std. DiD	Std. Wts.	Comp. Shift Wts.
Avg. Income (early career)	-$767*	-$1484	-$450
(s.e.)	(433)	(1269)	(663)
S.D. of Inc. (early career)	$123	-$268	$344
(s.e.)	(384)	(1001)	(563))
Income Growth (early career)	2.14 p.p.***	2.19 p.p.	2.78 p.p.***
(s.e.)	(0.61)	(1.48)	(0.98)
Poverty (early career)	0.37 p.p.***	0.31 p.p.	0.42 p.p.**
(s.e.)	(0.12)	(0.24)	(0.16)
Avg. Prime Age Income (25-54)	$4,545	-$20,490	$22,454
(s.e.)	(14,938)	(56,499)	(25,951)
S.D. of Inc. (mid-career, 40-44)	$1,584*	$1,716	$3,114**
(s.e.)	(946)	(2591)	(1,416)
	n=14,232		

$*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .10$; Standard errors are calculated via bootstrapping estimation procedure 1,000 times. Outcomes are labor market realizations for major selection, measured using ACS from year of student entry. See Section 1.4 for more detail. This sample excludes students who began school prior to the final increase in federal loan limits.

change in interest rates, borrowers change behavior after the IDR expansion.

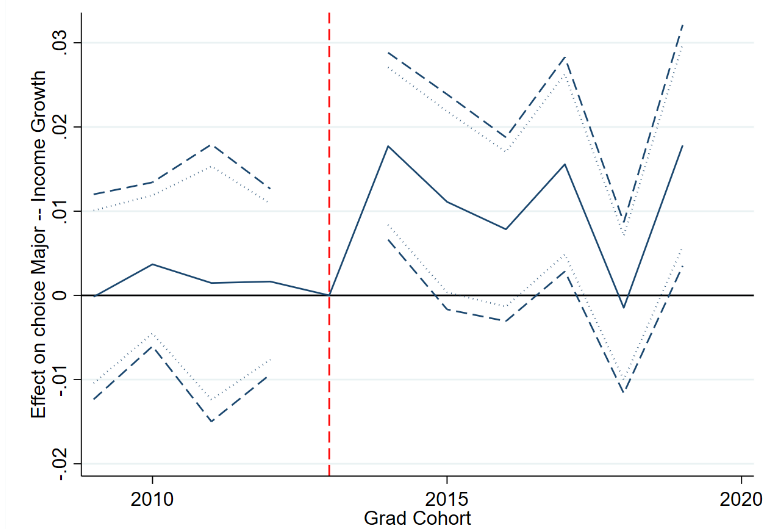
Finally, the increase in institutional aid could conceivably create an income effect for borrowers. In the theoretical model, institutional aid offsets the need for borrowing; however, if there are alternate methods of financing higher education, e.g. working while in school, then the increase in institutional aid could increase the resources of students as they enter the working period either directly, through increased savings, or indirectly, through higher intensive margin investments in human capital.[56] Without more detailed records on the complete financial aid packages of students, I cannot directly test this hypothesis.[57]

However, recall that the increase in institutional aid was largely directed at students from less-well off families. Table 1.7 reports the average net price for students by their family income bin. Note that there is little movement in the net price charged to families with

[56]Keane and Wolpin (2001) document working as an important means to cover pecuniary costs. I do not model this decision; however, it is reasonable to assume that working decreases the effectiveness of human capital investments, resulting in lower ability post-graduation.

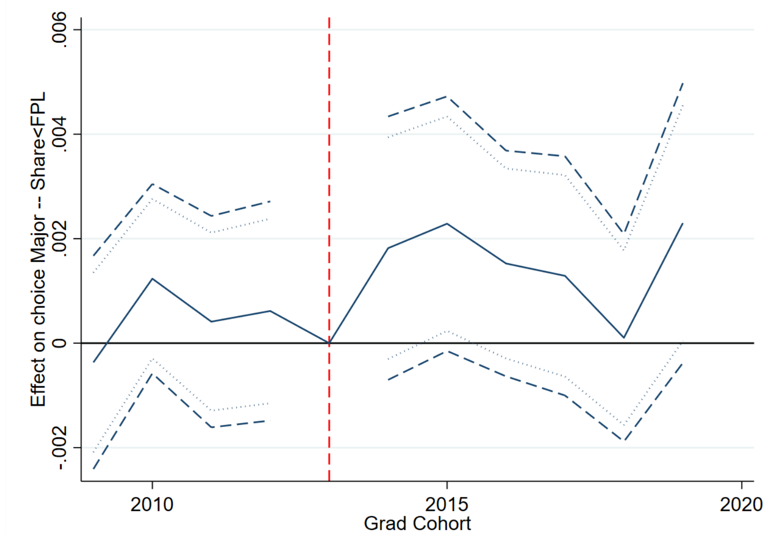
[57]This is an interesting avenue for future research.

Figure 1.6: Effect on Major Selection–Income Growth for Graduates, Men–Continuous Exposure



Dotted line indicates 90% CI, dashed line indicates 95% CI; Regressions are weighted using regression-version of compositional shift weights with robust standard errors. Explanatory variable is estimated monthly repayment on standard 10-year fixed repayment schedule (in $100s). Mean of explanatory variable among borrowers is $162

Figure 1.7: Effect on Major Selection–Share of Graduates <FPL, Men–Continuous Exposure



Dotted line indicates 90% CI, dashed line indicates 95% CI; Regressions are weighted using regression-version of compositional shift weights with robust standard errors. Explanatory variable is estimated monthly repayment on standard 10-year fixed repayment schedule (in $100s). Mean of explanatory variable among borrowers is $162

Table 1.11: Avg. Loan Debt (2019$) & Share of Borrowers by Family Income Bin

Grad. Cohort	Estimated Family Income Range <$25,000	$25,000-$50,000	$50,000-$75,000	$75,000-$100,000	$100,000+
	A. *Avg. Loan Balance among Borrowers (2019$)*				
2009	$28,500	$34,100	$36,300	$36,700	$31,800
2013	$27,900	$31,000	$35,400	$35,700	$32,200
2016	$22,400	$25,400	$28,400	$28,600	$30,000
2019	$16,800	$20,100	$25,600	$28,200	$29,900
	B. *Share Borrowers*				
2009	.81	.71	.67	.57	.26
2013	.75	.74	.70	.66	.32
2016	.70	.70	.73	.62	.30
2019	.53	.58	.65	.65	.28

incomes greater than $75,000 a year. Table 1.6 reports average loan balance by family income decile (proxied by average household income in the students' ZIP-code) and Table 1.11 reports average loan balance by estimated family income reported directly from the university. These data highlight the fact that most institutional aid–and thus movement in net-pricing and borrowing–was directed towards low-income families.

To explore the validity of the ATT estimates, I split my sample into households with estimated incomes below $75,000 and those with estimated incomes greater than $75,000. I re-run my analysis on these groups separately. Because households with incomes above $75,000 had little change in the institutional resources made available to them, this sample should capture the responsiveness to the policy expansion alone, net of any possible income effect. Table 1.12 reports the results of this exercise. Interestingly, there are no discernible effects for lower income families; however, the effects do persist for higher income families.

These exercises address particular sources of potentially confounding effects. In each case, the effects I identified in the main results in Table 1.8 for men remain, for at least a subset of the population. I interpret this as evidence that the theoretical implication–that is, borrowers have higher appetite for risk after the IDR policy expansion–are indeed affecting students' optimal human capital investments.

Table 1.12: Difference-in-Differences Estimates, Heterogeneity by Family Income., Men

	Est. Fam. Income<$75,000			Est. Fam. Income≥$75,000		
	(1) Std. DiD	(2) Std. Wts.	(3) Comp. Shift Wts.	(4) Std. DiD	(5) Std. Wts.	(6) Comp. Shift Wts.
Avg. Income (early career)	-$130	-$568	-$143	-$909**	-$762	-$1197**
(s.e.)	(802)	(1327)	(913)	(428)	(780)	(522)
S.D. of Inc. (early career)	$318	$391	$527	$146	-$128	-$443
(s.e.)	(669)	(1087)	(802)	(385)	(604)	(456)
Income Growth (early career)	1.03 p.p.	1.03 p.p.	0.91 p.p.	1.94 p.p.***	1.88 p.p.**	1.54 p.p**
(s.e.)	(1.20)	(1.58)	(1.33)	(0.63)	(0.93)	(0.77)
Poverty (early career)	-0.14 p.p.	-0.17 p.p.	-0.16 p.p.	0.32 p.p.***	0.41 p.p.**	0.34 p.p.**
(s.e.)	(0.21)	(0.27)	(0.23)	(0.12)	(0.17)	(0.14)
Avg. Prime Age Income (25-54)	-$5,129	-$25,215	-$10,030	-$3,761	$2,519	-$19,230
(s.e.)	(26,278)	(54,757)	(33,600)	(13,959)	(32,187)	(18,440)
S.D. of Inc (mid-career, 40-44)	$933	$1,446	$1,747	$1,190	$957	$137
(s.e.)	(1,635)	(2,723)	(1,932)	(911)	(1,537)	(1,142)
		n=3,528			*n=11,899*	

$***= p-val < .01$; $** = p-val < .05$; $* = p-val < .10$; Standard errors are calculated via bootstrapping estimation procedure 1,000 times. Outcomes are labor market realizations for major selection, measured using ACS from year of student entry. See Section 1.4 for more detail.

1.6.4 Additional Heterogeneity

The heterogeneity analysis presented in Table 1.12 is primarily to specifically address concerns about the confounding effect of changes in institutional aid; however, heterogenous effects are of interest in their own right.[58] However, such heterogenous effects *within* the student population are of interest in their own right. Here, I explore two additional sources of heterogeneity in the student population to assess whether the policy expansion had differential effects on these students' behavior.

First, I examine *first generation college students.* Patnaik et al. (2021) note that there are substantial differences in major selection between students who are first-generation college students–that is, those whose parents are *not* college graduates–and those whose parents are college educated. It is plausible that such differences are driven in part by sensitivity to labor market returns. Because loan repayment system impacts the expected labor market returns, such differences in sensitivity could lead to different treatment effect sizes for the two populations. For example, if first generation students are more likely to emphasize labor market returns when they select their major, then I would expect a larger change in behavior for these students.

Table 1.13 reports the results of this heterogeneity analysis. The point-estimates associated with first-generation students about twice as large in magnitude than the corresponding point estimates for non-first generation students, which is consistent with a hypothesis that first-generation students are more sensitive to labor market expectations than students with college-educated parents. While the differences are not statistically significant, which limits any strong conclusions from this exercise, this may be a fruitful avenue for future research into the effect of loan repayment plan on student behavior, given particular interest in the behavior and success of first generation students. In particular, higher education is frequently cited as a means to promote intergenerational mobility. These results suggest that *repayment plan structure* can impact investment decisions of first-generation students, which implies greater variance in realized measurements of intergenerational mobility, both in early career and over time, as first-generation students invest in majors with higher likelihood of initially low-wages but also steeper income trajectories. Furthermore, while the effects on average income are noisy, they are much larger in magnitude for this sub-sample. It is possible that this population also selects into majors associated with lower wages, not simply

[58]Note that the heterogenous effects in Table 1.12 are difficult to interpret as differential *treatment effects*–that is differential responses to the IDR expansion–or due primarily to *confounding effects* of the change in institutional aid.

higher probability of low-wage draws.

A second source of heterogeneity I analyze is by race. Lochner and Monge-Naranjo (2016) report that there are significant differences in post-graduation loan repayment rates by race, with black borrowers experiencing more difficulty. Indeed, these different outcomes by race were a motivating factor in the Biden Administrations newly announced actions regarding student loans, including their proposed expansion of IDR options (Biden (2022)). Here, I examine whether underrepresented minorities react to the policy expansion differently from other students. Such heterogeneity is especially of interest if, for example, underrepresented minority students are more responsive to shifts in labor market realizations than other students. Table 1.14 reports the results of this heterogeneity analysis. Interestingly, there is little evidence that underrepresented minority students are responsive to the policy change; however, these results are relatively noisy and sensitive to specification, suggesting caution in interpreting these results.

1.6.5 Additional Outcomes

The main results focus on measures of labor market returns for particular majors, since the motivating theory suggests that it is the insurance value on the labor market that drives response. However, it is plausible that other margins could also be affected. For example, Akers (2012) and Millet (2003) both document a *decrease* in graduate school attendance among those who have undergraduate debt. Both of those studies examined students during a period when fixed-repayment loans were the dominant form of student debt. Here, I ask whether there any effects of IDR expansion on graduate school decisions among borrowers. In particular, I focus on graduate school enrollment within one year of completion.[59] Note that I cannot use the same measurement of debt as above for this sample, since the *outcome* is graduate school enrollment.[60] Here, I instead use a measurement of debt from while the borrower is in undergraduate.[61]

[59]This allows me to examine all my treatment cohorts. Furthermore, for students who enroll in post-baccalaureate studies, approximately 60% do so within one year.

[60]Recall, the above specifications use measurement of debt at or after graduation, depending on the cohort, conditional on not being enrolled in graduate school.

[61]It is possible that there is misclassification here for some cohorts, as a non-borrower could become a borrower prior to graduation. However, in settings with binary treatment such as this, this should attenuate results.

Table 1.13: Difference-in-Differences Estimates, Heterogeneity by Parental Education, Men

	First Generation (no parental college degree)			Non-First Generation (parental college degree)		
	(1) Std. DiD	(2) Std. Wts.	(3) Comp. Shift Wts.	(4) Std. DiD	(5) Std. Wts.	(6) Comp. Shift Wts.
Avg. Income (early career)	-\$1,895*	-\$1,209	-\$1,263	-\$448	-\$350	-\$221
(s.e.)	(1,015)	(1691)	(1,203)	(379)	(698)	(446)
S.D. of Inc. (early career)	-\$1,289	-\$805	-\$760	\$536	\$535	\$581
(s.e.)	(870)	(1,317)	(1,009)	(329)	(548)	(403)
Income Growth (early career)	3.88 p.p.***	4.01 p.p.**	3.92 p.p.**	1.77 p.p.***	1.35p.p.*	1.34 p.p.**
(s.e.)	(1.44)	(1.99)	(1.74)	(0.54)	(0.77)	(0.65)
Poverty (early career)	0.53 p.p.**	0.62 p.p.*	0.60 p.p.*	0.24 p.p.**	0.21 p.p.	0.20 p.p.*
(s.e.)	(0.27)	(0.36)	(0.32)	(0.10)	(0.14)	(0.12)
Avg. Prime Age Income (25-54)	-\$31,362	-\$5,895	-\$9,661	\$9,078	\$10,606	\$13,294
(s.e.)	(33,629)	(69,478)	(46,658)	(12,271)	(27,944)	(15,860)
S.D. of Inc. (mid-career, 40-44)	-\$133	\$844	\$737	\$1,813**	\$2,069	\$2,115
(s.e.)	(2,087)	(3,287)	(2,555)	(782)	(1,324)	(971)
		n=2,177			*n=15,718*	

$***=p-val<.01$; $**=p-val<.05$; $*=p-val<.10$; Standard errors are calculated via bootstrapping estimation procedure 1,000 times. Outcomes are labor market realizations for major selection, measured using ACS from year of student entry. See Section 1.4 for more detail.

Table 1.14: Difference-in-Differences Estimates, Heterogeneity by Race, Men

	Underrepresented Minority			Non-Underrepresented Minority		
	(1) Std. DiD	(2) Std. Wts.	(3) Comp. Shift Wts.	(4) Std. DiD	(5) Std. Wts.	(6) Comp. Shift Wts.
Avg. Income (early career)	$268	$2553	$4,040*	-$673*	-$20	-$918*
(s.e.)	(1,102)	(3,024)	(2,337)	(369)	(870)	(506)
S.D. of Inc. (early career)	$380	$1,886	$3,407*	$273	$841	$98
(s.e.)	(1,031)	(2,434)	(2,037)	(308)	(637)	(401)
Income Growth (early career)	-1.25 p.p.	-0.54 p.p.	1.96 p.p.	2.51 p.p.***	3.37 p.p.***	2.36 p.p.***
(s.e.)	(1.65)	(3.55)	(4.20)	(0.52)	(0.89)	(0.67)
Poverty (early career)	-0.20 p.p.	-0.25 p.p.	0.10 p.p.	0.31 p.p.***	0.50 p.p.**	0.33 p.p.***
(s.e.)	(0.28)	(0.51)	(0.50)	(0.10)	(0.16)	(0.12)
Avg. Prime Age Income (25-54)	-$11,298	$75,688	$155,817	$7,917	$41,534	-$3,441
(s.e)	(37,172)	(133,556)	(107,309)	(12,369)	(26,495)	(18,426)
S.D. of Inc. (mid-career, 40-44)	-$910	$3,107	$7,313	$1,861**	$3,676**	$1,703*
(s.e.)	(2,310)	(6,084)	(5,338)	(742)	(1,596)	(1,003)
		n=1,627			*n=16,205*	

$*** = p-val < .01$; $** = p-val < .05$; $* = p-val < .10$; Standard errors are calculated via bootstrapping estimation procedure 1,000 times. Outcomes are labor market realizations for major selection, measured using ACS from year of student entry. See Section 1.4 for more detail.

Additionally, it is possible that, instead of expected labor market returns, it is *difficulty of major* or some other characteristic of the major itself which drives investment decisions. Here, I also examine final cumulative GPA and selection of STEM majors as alternative outcomes.[62] Recall that, with many possible majors and potentially correlated major-specific abilities, the model would suggest that shifts in major selection would occur *within* related major types towards majors with higher probability of initially low-wage realizations but higher growth–that is, the model I present suggests that there should be little movement in larger aggregations of major categorization, such as STEM. However, if an alternative hypothesis is that students select into "easier" majors, for example, there could be shifts in these larger major aggregations.

Table 1.15 displays the estimates for these outcomes. Columns (1)-(3) report results for men while columns (4)-(6) report results for women. There is virtually no change in undergraduate GPA among borrowers after the policy shift. Furthermore, there is not evidence that borrowers are more likely to select into STEM fields. I interpret this as further evidence that the main effect of IDR expansion is the insurance value on the labor market.

There is evidence that men are more likely to enroll in graduate programs after the policy change. While the corresponding effects are also positive for the professional programs in law and medicine–significantly so for law programs–they cannot fully explain the shift in graduate enrollment, suggesting that the increase in graduate school attendance is occurring for other types of programs. Again, there is virtually no response among women to the policy change. These results suggest that the diminished probability of attending graduate school among borrowers is also affected by the structure of the repayment plan, at least among men. Insofar as graduate school is a risky endeavor with uncertain labor market returns–especially for degrees other than professional degrees in law or medicine–these results are consistent with the hypothesis that the IDR expansion offers insurance against such risk.

[62] These outcomes use similar measures of debt as the main results.

Table 1.15: Difference-in-Differences Estimates–Additional Outcomes

	A. *Men*			B. *Women*		
	(1) Std. DiD	(2) Std. Wts.	(3) Comp. Shift Wts.	(4) Std. DiD	(5) Std. Wts.	(6) Comp. Shift Wts.
	Graduate School Enrollment					
Any Grad School	4.78 p.p.***	4.03 p.p.***	3.93 p.p. **	1.48 p.p.	0.07 p.p.	0.34
(s.e.)	(1.21)	(1.50)	(1.53)	(1.26)	(1.58)	(1.16)
Law School	1.67 p.p.***	1.27 p.p.**	1.46 p.p.**	1.04 p.p.**	0.99 p.p.	1.05 p.p.*
(s.e.)	(0.53)	(0.62)	(0.65)	(0.52)	(0.61)	(0.63)
Medical School	0.74	0.85 p.p.	0.75 p.p.	-0.15 p.p.	-1.25 p.p.*	-1.14 p.p.*
(s.e.)	(0.59)	(0.69)	(0.73)	(0.58)	(0.66)	(0.67)
	n=22,169			*n=23,577*		
	Additional Undergraduate Outcomes					
STEM Major	1.81 p.p.	1.59 p.p.	0.40 p.p.	1.01 p.p.	0.12 p.p.	0.30 p.p.
(s.e.)	(1.57)	(2.06)	(1.79)	(1.53)	(2.11)	(1.85)
Final GPA	0.01	0.07 *	0.03	0.01	-0.00	0.00
(s.e.)	(0.01)	(0.04)	(0.02)	(0.01)	(0.05)	(0.02)
	n=17,895			*n=18,219*		

$*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .10$; Standard errors are calculated via bootstrapping estimation procedure 1,000 times. Outcomes are labor market realizations for major selection, measured using ACS from year of student entry. See Section 1.4 for more detail. Top panel uses measurement of debt *during* undergraduate to determine borrowing status. Bottom panel uses measurement of debt *on or after* undergraduate completion to determine borrowing status

1.6.6 Discussion on Efficiency Gains vs. Moral Hazard

The empirical exercises presented here suggest that male borrowers respond to the IDR expansion by shifting their major selections. Specifically, they are more likely to select majors that have more graduates with incomes below the federal poverty line in their early career. However, there is also evidence that this involves a dynamic tradeoff: these majors also have higher income growth over the early career. Notably, there is no significant effect on *average* incomes in the early career and similarly no significant effect on average incomes over prime-age working years, though there is evidence of higher long-run earnings variance. Together, this evidence suggests that the IDR expansion allowed borrowers to select majors that shifted the *timing* of their earnings as well as the longer-run variance of earnings.

As noted in Section 1.3, there are two potential effects of an IDR: (i) it insures against initially low-monetary return majors by automatically providing some consumption smoothing, which, in the absence of perfect credit markets, also allows for accepting lower initial earnings but higher wage trajectories; and (ii) it may subsidize low-return majors via eventual forgiveness, which entails lower *lifetime* earnings. When considering the efficiency of a policy, the first effect noted here (insurance) can improve efficiency while the second effect (subsidy) is not, in general, an improvement in efficiency.

The fact that the empirical results suggest borrowers change their investments to take advantage of the dynamic tradeoff between initial wages and wage growth but *do not* have lower average wages suggests that the former effect is the more important consideration. That is, if the subsidy effect were the dominant consideration for borrowers in major selection, we should see a decrease in average earnings as well, as borrowers take advantage of the forgiveness. It should be cautioned, however, that this is only *suggestive* evidence that the insurance effect is dominant: major selection is only one margin on which borrowers can adjust their behavior. While I do not find evidence that borrowers shift into majors with lower average wages, it is possible that their actual wage realizations *are* lower, as they adjust behavior on the labor market as well. A more comprehensive study of this comparison between insurance and subsidy would examine the entire trajectory, from major selection to eventual labor market behavior.

As noted in Section 1.6.4, the point estimates for first-generation students suggest selection into lower average-wage majors *and* lower variance-wage majors, though these results are not significant. This suggests that it is possible that the subsidy effect of IDR expansions *is* relevant for key populations. Further research into this question is necessary to understand whether the 2013 IDR expansion–and potentially future expansions–effectively subsidize low-monetary return human capital investments.

Furthermore, I examine the role of borrowing status in the post-treatment period in a multinomial choice model of major selection. In particular, I use a multinomial logit model with major selection as the outcome of interest. In addition to borrower status, I control for parental education, a quadratic in household income, high school GPA, an indicator for if the student is white, and an indicator for residency status. I report the marginal effect of borrower status on probability of selecting each major in Appendix A.8. Notably, significant differences in major selection probability remain in the post-treatment period between borrowers and non-borrowers. This suggests that credit constraints may remain an important consideration in major selection. I believe these questions on the role of credit constraints, repayment plan, and efficiency merit further research.

1.7 Conclusion

This paper presents the first evidence of how the structure of student loan repayment plans alters human capital investments. Using the 2013 introduction of Pay As You Earn (PAYE), which significantly increased IDR generosity and preceded a dramatic shift in student loan repayment *away* from standard repayment plans *towards* IDR plans, I document that male borrowers were more likely to select majors that had lower initial remunerative returns–represented by a larger share of graduates with incomes below the poverty line–but higher wage growth. These findings are consistent with theoretical implications that the IDR offers insurance against low monetary returns *and* offers the flexibility to trade off these initially low income realizations with higher income growth.

These results are robust to specifications that account for non-random selection into borrowing status–leading to non-parallel trends–as well as changing compositions of borrowers and non-borrowers over time. In addressing the latter concern, I contribute to the nascent literature in applied econometrics on difference-in-differences with compositional changes by extending the reweighting approach of Abadie (2005) to such contexts. Furthermore, the results are robust to endogenous selection into treatment induced by the policy change itself as well as potential confounding factors driven by broader policy changes in student loans.

I do not find significant responses among women. Zafar (2013) finds that men are more likely to emphasize monetary returns on the labor market in selecting their major, while women are more likely to emphasize non-pecuniary returns while in school. This finding, then, is consistent with a hypothesis that men are more sensitive to changes in expected monetary return *net of repayment plan* in selecting their major.

This work suggests that borrowers are responsive to repayment plan structure in selecting their majors–though the response is driven by increased *probability* of low wage draws (with compensatory increased growth) rather than lower returns on average, suggesting that the insurance value is a more dominant consideration than the subsidy value. Furthermore, the evidence suggests relatively little change in lifetime earnings associated with these major selections–if anything, there is marginal evidence of positive effects–and increased long-run variance in earnings, suggesting that the primary effects are a shift in timing of earnings allowed by the IDR consumption-smoothing mechanism. Given further proposed expansions to IDR generosity, however, it is possible that the subsidy value may become more important for future borrowers. In order to capture the potential efficiency gains of IDR policies, future work should examine the entire set of borrowers' decisions, starting with investment choices and leading into labor market activities.

Additionally, there are historically barriers to IDR enrollment that could limit its use, dampening the effect of the IDR expansion (Cox et al. (2020)). Furthermore, the policy expansion only affects those *who know about it.* Informational frictions such as this would also attenuate results, as borrowers without full information would continue to act as if the old policy regime was in effect, i.e. as if IDR expansion had not occurred. In 2016, 56% of graduating student borrowers from public research universities reported being aware of IDR programs (National Center for Education Statistics (NCES) (1993, 2016, and 2019)). Furthermore, 45% of all undergraduate borrowers reported being aware of IDR programs (National Center for Education Statistics (NCES) (2016)).[63] This suggests that the ATT estimates for the policy effect are *lower* than the true value.[64] Richer data that includes information on about IDR awareness would allow for an extension of my approach with more precise estimates of the ATT. Finally, my results are specific to a single, selective research university. If students at selective research universities respond differentially than other

[63] 44% of first-year students, 40% of second-year students, 39% of third-year students, 51% of fourth-year students, and 53% of fifth-year students reported being aware of IDR programs.

[64] In this sense, I estimate the Intent to Treat (ITT) rather than ATT parameter. Note that it is *awareness* of the program, not necessarily eventual use, that matters in this context. That is, students make human capital investment decisions under uncertainty, and the *availability* of the IDR should alter those decisions via the insurance and subsidy value. However, when ability and labor market realizations occur–i.e., when uncertainty is eliminated ex post–a student who has made decisions with the knowledge of IDR availability need not make use of the IDR.

students to expected monetary returns to major selection, this would limit external validity of these results. Future research is necessary to validate the generalizability of these estimates.

Nevertheless, this work documents for the first time how *repayment plan structure* is an important consideration in major selection of student-borrowers. These shifts in investment behavior have implications for the skills mix and the earnings distribution of the college-educated workforce, as graduates are more likely to leave school with degrees for which there is higher probability of low initial earnings but also higher earnings trajectories. In future work, I hope to examine how Universities and employers respond to this changing loan policy landscape, as well as the interplay between loan policies and broader student aid environment.

CHAPTER II

College Finance, Repayment Regimes, and Job Search

2.1 Introduction

> "[Income driven repayment] will be nothing less than liberating for many students [...] who graduate with big debts and then feel driven into careers with higher pays but lower satisfaction. A student torn between pursuing a career in teaching or corporate law, for example, will be able to make a career choice based on what he or she wants to do, not how much he or she can earn to pay off the college debt."
>
> –President Clinton

In the fourth quarter of 2022, outstanding student loan balances were $1.6 trillion, the vast majority of which was either directly administered or backed by the Federal government. This represents the largest category of non-mortgage consumer debt, overtaking auto loans in 2010. Federally held or backed student loan debt grew continuously through the Great Recession, more than doubling from $624.54 billion[1] in 2007. The increase in these aggregate numbers is in no small part driven by increases in the number of borrowers, which increased by over 50% between 2007 and today. The real average amount borrowed increased by approximately the same amount over the same time frame, implying that about half the increase in the aggregate loan balance is attributable to increased numbers of borrowers and half to increased levels of borrowing.[2]

In addition to these broad trends in borrowing, *how* people have paid their loans

[1] in 2018 USD

[2] See Federal Reserve Bank of New York, Center for Microeconomic Data (2022) and National Student Loan Data System (NSLDS) (2022)

back has shifted in recent years. The Clinton administration introduced income-driven repayment plans (IDRs) into the US federal loan system in the mid 1990s; however, there was very low uptake of this repayment option over the following 20 years. In 2013, just over 10% of all borrowers of Federal Direct Loans were enrolled in IDR plans, constituting just over a fifth of total balances. By the end of 2020, about a third of all borrowers of Federal Direct or Department of Education held Federal Family Education Loans were enrolled in IDR plans, representing over half of all outstanding balances (National Student Loan Data System (NSLDS) (2022)). Most of this shift in repayment plans is due to the introduction of the Pay As You Earn (PAYE) and Revised Pay As You Earn (RePAYE) plans by the Obama administration in 2013 and 2016, respectively, which expanded eligibility and generosity of these IDR in the United States.

As noted in Chapman (2006), a key consideration in the introduction of the US IDR plan was concern that high debt burdens may push graduates away from certain occupations, whereas justification for other notable cases of IDR plans–e.g. the Australian and United Kingdom plans–were explicitly about perceived regressivity of free-college-for-all and insurance against future job market outcomes. Chapman cites the quote at the beginning of this article as well as other contemporary evidence that occupational sorting was a primary concern for the Clinton administration. In remarks on the expansion of IDR plans, President Obama (2014) similarly cited occupational choice as a concern, saying "[w]e want more young people becoming teachers and nurses and social workers." In fact, empirical and theoretical evidence suggests that the presence of student loans and the structure of repayment matters for occupational decisions made after graduation, though whether this pushes people to take on higher wage or lower wage jobs is not always clear.[3]

In this paper, I examine a theoretical model that combines an education decision, a financing decision, and labor market decisions into a unified framework. The student loans system in the United States is a large and complex system, and it is national in scope with few–though not zero–natural experiments available for study.[4] Hence, any analysis of the system is inherently difficult. The structural model allows me to examine general equilibrium in labor market behavior effects of specific institutional changes. It also helps explain seemingly contradictory effects documented in the existing literature relatively straightforwardly. Understanding how behavior changes in different institutional settings

[3]e.g., see Chapman (2006), Rothstein and Rouse (2011), Ji (2021), Kaas and Zink (2011), Abraham et al. (2018)

[4]McPherson and Schapiro (2006), describing the entire higher-education finance system, quip: "if indeed such a tangled and decentralized set of arrangements warrants the label 'system"'

is key to understanding the total impact such changes have–for example, I will argue that the payment structure affects job search decisions post education; hence, changing the repayment system will change the composition of jobs in the labor market, the earnings distribution, the schooling investment decision, and the total revenue of the repayments. Finally, the structural model allows for aggregate welfare calculations for policy changes. This paper also contributes to the literature on how search frictions affect optimal investment.[5]

In the following section, I review the related literature. In Section 2.3, I introduce the model and define equilibrium under different repayment regimes. Section 2.4 considers comparative statics of the optimal decision rules. Section 2.5 considers (brief) thoughts on efficiency and Section 2.6 concludes.

2.2 Related Literature

Stiglitz (2014) notes that a primary risk of non-dischargeable student loans is the additional risk of being unemployed–he notes that "the consequences of not having a job are more severe with a conventional [fixed repayment] loan" leading individuals to increase search intensity and potentially lower reservation wages. Furthermore, he notes that individuals with uncertain prospects of their future labor market outcomes may avoid investing in human capital altogether, even in the presence of fixed repayment loans, due to the high costs associated with being unemployed. These two effects–the insurance problem and the holdup problem[6]–are those that are mainly addressed in the existing theoretical literature. Moen (1998) also considers the hold-up problem in human capital investment. In the context of a random search model, he argues that uncertainty over bargained wages in the future cause underinvestment in human capital today, and the introduction of a type of IDR can solve the holdup problem.

In empirically assessing how debt impacts labor market behavior, Rothstein and Rouse (2011) use a natural experiment wherein a highly selective university replaced loans with grants in their aid packages. Using this exogenous variation in the presence of debt on graduates' labor market decisions, they document that loan-free graduates were more likely to take on lower paying jobs, evidence consistent with credit constrained lifecycle agents. Similarly, Field (2009) uses an experiment that offered two mathematically identical aid packages at NYU Law–one which consisted of loans that converted to grants if individuals

[5]see, e.g., Acemoglu and Shimer (1999), Kaas and Zink (2011), Ottonello (2018)

[6]These are closely related, really the difference being the time when the problem is considered

took low-paying public sector jobs and one which consisted of grants which converted to loans if individuals *did not* take public interest jobs. She found that those with the latter aid packages were significantly more likely to work in the public-interest law field following graduation. Field's findings point more towards behavioral biases when considering debt burdens, but the main effect consistent throughout Rothstein and Rouse and Field is it appears as though the presence of debt pushes individuals towards higher paying jobs. Of course, the treatment here is binary–either individuals have debt or not–and the sample is likely to be more high ability than the general population or even the college-going population. Nevertheless, the evidence suggests that debt is an important consideration in individuals' labor market behavior. In particular, these papers highlight the "repayment burden," the borrower behavior of taking on higher-wage jobs than they would in the absence of debt to ensure they can repay their debts.

Abraham et al. (2018) focus on the "default penalty" posed by student loans: if borrowers are unable to make payments in periods of unemployment–or underemployment more generally–leading to default on debt, then they may avoid riskier, higher paying jobs that they would otherwise be willing to take.[7] This is because defaulting on debt carries significant penalties: in the United States, failure to pay federal student loans for 9 consecutive payments can potentially lead to (i) credit score penalties due to adverse credit reports which begin after 3 missed payments; (ii) loan acceleration, where the entire balance is immediately due; (iii) loss of benefits, such as loan deferral, forbearance, or access to alternative repayment plans (including IDR plans); (iv) wage garnishment and seizure of tax returns and other federal benefits to cover debts; and (v) additional fees added onto the loan (CFR (2021)). Furthermore, bankruptcy laws in the United States preclude automatic discharge of student debt, meaning these increased financial penalties cannot be easily removed (see, e.g. Taylor and Sheffner (2016)). However, Abraham et al. (2018) note that IDR plans can erase this result through the insurance properties of IDR plans, empirical evidence for which is provided by Herbst (2023), who documents a 22 percentage point decline in delinquency rates among distressed borrowers who switch to IDR options from standard repayments.

Building on this work that documents distortionary behavior in labor market activity among borrowers, specifically due to the default risk associated with fixed repayment loans, Ji (2021)

[7] In addition, Abraham et al. highlight the importance of behavioral biases in the ex post welfare among borrowers when they are presented with choices over repayment plans, i.e. choosing between fixed repayment and income-driven repayment compared to no choice in assignment to one or the other.

documents that fixed repayment systems can suboptimally reduce the reservation wage of workers in a dynamic random search model. He also argues that IDR systems can ease this result through insurance protection.

As noted above, there have been empirical studies that have documented workers who search for *increased* wages in the presence of debt. However, this phenomenon has been less addressed in the theoretical literature. An exception is Kaas and Zink (2011), who take the Rothstein and Rouse (2011) results to motivate the construction of their theoretical model. In the context of a directed search model, they show that fixed repayment systems can induce individuals to search in suboptimally high wage markets. They argue that this is a form of moral hazard, since lenders cannot contract on which markets borrowers search in and hence bear some of the risk associated with searching in these high wage markets, namely the higher risk of unemployment. However, their model does not explicitly include default penalties, allowing borrowers to avoid any debt-related risk of being unemployed.[8]

This paper is most closely related to Ji (2021) and Kaas and Zink (2011). In contrast to the latter, I allow for a dynamic working period during which borrowers repay debt and also explicitly model default penalty that increases with default amount. This causes agents to consider the "default penalty" in making their optimal search decisions. In contrast to the former, I allow for directed search, which allows agents to explicitly trade off wages with unemployment risk, and introduce heterogenous agents in terms of productivity.[9] Whereas Ji (2021) argues that fixed repayment systems lead to suboptimally low reservation wages–due to default risk–and Kaas and Zink (2011) argue that fixed repayment systems lead to borrowers to search in suboptimally high wage-markets–due to repayment burden and the inability to contract on seach behavior leading borrowers to "pass off" the costs of searching in these markets to the lender–I show here that both effects can be true for different *types* of borrowers, where type is determined by heterogeneity in labor productivity. Notably, it is the high skill workers who are more likely to search in higher wage markets due to fixed repayment contracts, which is consistent with the Rothstein and Rouse (2011)

[8]That is, the only risk of being unemployed from the perspective of the borrower is that they will be unemployed and hence have lower consumption, but the actual level of consumption is common across all borrowers.

[9]Ji (2021) allows for heterogeneity among agents only in terms of initial wealth and pecuniary costs of college–which determine total debt conditional on going to school–as well as non-pecuniary costs of college. He argues that allowing for heterogeneity in labor productivity through on-the-job training or learning by doing does not quantitatively impact his results. However, my argument here is that heterogeneity in baseline abilities–which translates to heterogeneity in labor productivity–*can* impact the optimal search behavior in a directed search model.

finding, insofar as the students at their selective university were more likely to be drawn from high ability types. Furthermore, as both Ji (2021) and Kaas and Zink (2011) argue that the repayment schemes that account for income–in the former case an IDR and in the latter a graduate tax–can ease the distortions they describe–i.e. either raise reservation rates or lead to lower wages in directed search, respectively. Here, I also show that a graduate tax can ease distortionary behavior, as in Kaas and Zink (2011), but that the direction of this effect depends on labor productivity and critically is only true under rather strong assumptions. I also highlight that IDR plans continue to exhibit distortionary search behavior relative to the no-loans case, but that certain features of an IDR system can help to alleviate these distortions.

2.3 Dynamic Model

I set out to create a model that captures the investment decision under a known financing rule, and subsequent labor market behavior in a directed search model. The model is discrete time with an infinite horizon, indexed by t. There are two "stages," investment and working. In $t = 0$, individuals make an investment decision with access to a known financing plan. For time $t = 1, 2, \dots$, they work a labor market with search frictions, where they can search for specific wages and firms.

2.3.1 Investment

Individuals, indexed by i, enter time 0 with initial ability, a_i^0, and wealth, b_i^0. Individuals can make a dichotomous choice to invest in education or not to augment their abilities according to a production function $a_i = f(a_i^0)$. There is some non-pecuniary cost, $v_a(a_i^0)$, where $v_a > 0$ and $v_a' < 0$. After augmentation, individuals' abilities are denoted by a_i and remain so for the rest of their lives. If they do not invest, they maintain their endowed ability, a_i^0.

Additionally, there is an exogenous, pecuniary cost to education investment, given by c_i, which they initially pay out of their endowed wealth. If, however, they cannot completely finance their education out of endowed wealth, they can borrow the difference and repay during their working years under a known repayment plan. That is, debt is $x_{i,0} = \max\{c_i - b_i^0, 0\}$.

2.3.2 Repayment Plans

In this section, ω will be per-period income, either wage income or some unemployment income, and s will be savings. I give more detail on these variables in the following section.

The repayment plan under the FR regime is a known function depending on the total amount of debt outstanding, an exogenous interest rate, r, and time in working stage:

$$\rho(x_t, t) = x_t \frac{r}{1 - \left(\frac{1}{1+r}\right)^{\hat{t_{FR}} - t}}$$

This formula gives the fixed repayment amount per period given interest rate r. After $\hat{t_{FR}}$ periods, they exit repayment.[10] During repayment, there is some risk of default that is a function of the difference between the fixed repayment amount and total liquid assets–earned income plus any savings, $\omega + s$–net of a minimum consumption threshold, $\underline{c_{FR}}$.

$$\psi(\omega, s, x_t) \text{ such that } \begin{cases} \omega + s - \underline{c_{FR}} - \rho \to -\infty & \Rightarrow \psi \to 0 \\ \omega + s - \underline{c_{FR}} - \rho \to \infty & \Rightarrow \psi \to 1 \\ \omega + s - \underline{c_{FR}} - \rho = 0 & \Rightarrow \psi = .5 \end{cases}$$

This parameter stochastically captures default probability as a function of ability to repay.[11] In the event of default, I assume that the borrower's liquid assets net of the minimum consumption threshold are used to pay down debt. This means that debt evolves during the repayment period according to the transition rule:

$$x_{t+1} = \begin{cases} (1+r)x_t - \rho(x_t, t) & \text{if no default (probability } \psi(\omega, s, x_t)) \\ (1+r)x_t - (\omega + s - \underline{c_{FR}}) & \text{if default (probability } 1 - \psi(\omega, s, x_t)) \end{cases}$$

Note that, if the borrower remains out of default the entire repayment period, they pay off their debt and there is never any accrued interest–debt evolves according to a simple interest rule. However, if they enter default and their liquid assets do not cover accruing interest, this interest is capitalized into the principal, x_{t+1}.[12]

Under the IDR regime, the repayment plan is a known function depending on income in each period and a set amount of time in the labor force. Furthermore, there is some minimum

[10]Under the standard repayment plan in the United States, this is 10 years, or 120 payments.

[11]This allows for standard procedures to define decision rules in labor search, see below. Note that this allows for some borrowers with liquid earnings above the minimum consumption threshold to default and vice versa. However, with a parametric form of, for example, the logistic function such that $\psi(\omega + s - \underline{c_{FR}} - \rho) = \psi(x) = \frac{1}{1+\exp(-kx)}$, a large enough scaling parameter, k, will ensure that almost all borrowers with earnings below this threshold default and almost all borrowers with earnings above this threshold do not.

[12]This is slightly different than the actual process in the United States. In reality, debt accrues according to a simple interest rule and is only capitalized after exiting default. For tractability, I have instead considered this accrual rule here.

income disregard, below which there are no required payments (though interest still accrues). After a set amount of time, $\hat{t_{IDR}}$, the borrower exits repayment.[13] However, if the borrower ever pays off all of their principal and accrued interest, they exit repayment. Interest accrues according to a simple interest rule and is not capitalized into the principal.[14]

$$\rho(\omega, A_t, x_t) = \begin{cases} 0 & \text{if } \omega < \underline{c_{IDR}} \\ \iota(\omega - \underline{c_{IDR}}) & \text{if } \omega \geq \underline{c_{IDR}} \text{ and } \iota(\omega - \underline{c_{IDR}}) < A_t + x_t \\ A_t + x_t & \text{if } \iota(\omega - \underline{c_{IDR}}) \geq A_t + x_t \end{cases}$$

That is, the individual makes payments that are a fraction $\iota \in (0, 1)$ of income over the income disregard in period t for all periods before $\hat{t_{IDR}}$ unless their payments will completely pay off their total debt. Note that the IDR plan eliminates default risk by assumption.[15] Here, the debt and accrued interest evolve according to the following transition rules:

$$\begin{aligned} x_{t+1} &= x_t - \max\{\rho(\omega, A_t, x_t) - A_t, 0\} \\ A_{t+1} &= \max\{A_t + (1 + r)x_t - \rho(\omega, A_t, x_t), 0\} \end{aligned}$$

After making this investment decision, individuals advance to period $t = 1$ and enter the labor market, which operates according to a directed search model and has an infinite horizon.

2.3.3 Labor Market

The labor market operates as a standard directed search model.[16] The key difference here will be the presence of debt that individuals consider in their job search. I am assuming a mass of firms and workers of similar endowed ability and wealth in order to avoid the case where either firms or workers can exercise market power.

A mass of identical firms pay a cost, k, to post a vacancy, which is the couple, (a, w), and workers who meet the posted criteria (i.e. $a_i = a$) may search for jobs in that submarket. Intuitively, firms post necessary criteria for a job as well as the wage for that jobs. Workers observe firm postings and decide which market (wage) to search in, provided they meet the posting criteria.

[13] In the United States, IDR plans last for 20 or 25 years, depending on the plan. Additionally, any remaining balance is forgiven, though the IRS considers forgiven debt taxable income, in general.

[14] This corresponds to the RePAYE plan in the United States, under which interest only capitalizes if the borrower leaves the IDR plan.

[15] Recall, IDR plans dramatically reduce default risk, see e.g. Herbst (2023).

[16] Moen (1997) and Shimer (1996) pioneered the development of directed/competitive search models. Wright et al. (2021) offer a useful guide on the structure of these models and key features.

The mass of workers searching in a particular submarket is given by $u(a, w)$ and the number of vacancies posted in the submarket is given by $v(a, w)$. Vacancies and unemployed searchers are matched according to a matching function $m(u(\cdot), v(\cdot))$, which has constant returns to scale, is increasing and concave in both arguments. Additionally, I will assume that the elasticity of substitution, $\sigma \leq 1$.[17]

Define $\theta(a, w) = \frac{v(a,w)}{u(a,w)}$ to be the market tightness in a particular submarket, (a, w). Then $\frac{m(u(\cdot),v(\cdot))}{u(\cdot)} = m(1, \theta(\cdot)) = p(\theta(\cdot))$ is the probability a worker finds a job in a particular submarket. Similarly, $\frac{m(u(\cdot),v(\cdot))}{v(\cdot)} = m(\frac{1}{\theta(\cdot)}, 1) = q(\theta(\cdot))$ is the probability a firm posting a vacancy is matched to a worker. Note that $p(\theta(\cdot))$ is increasing and $q(\theta(\cdot))$ is decreasing in $\theta(\cdot)$ and $p(\theta(\cdot)) = q(\theta(\cdot))\theta(\cdot)$.

In each period, $t \geq 1$, workers who enter the period unemployed may search for jobs. Firms post vacancies at fixed cost k as long as posting is expected to yield a non-negative profit, i.e. free-entry dictates the firm demand for workers.

Workers who enter the period employed continue their relationship with firm with probability $\phi \in (0, 1)$ and with probability $(1 - \phi) \in (0, 1)$ their match dissolves and they become unemployed. Importantly, workers who enter the period employed and subsequently lose their jobs *cannot* search that period–they *must* remain unemployed for one period before they may search. Firms that do not have workers, either because they did not match or because their match dissolved, cease to exist.

Finally, production, consumption and savings occur. Employed workers earn their wage w_t and unemployed workers earn some unemployment value b. If workers are not in default, they repay student loans according to the repayment rule outlined in Section 2.3.2 and they may choose to save for next period, s_{t+1}, at savings rate $1 + R$. Returns on savings are not contingent on employment state and the interest rate is fixed and known. I restrict savings to be positive, $s_{t+1} \geq 0$. If workers are in default, their liquid assets net of a minimum consumption threshold are used to pay loan debt and they may not save. Workers begin their working life with any leftover endowment that was not used on education expenses, $s_1 = \max\{b_i^0 - c_i(z_i^*), 0\}$. Workers have flow utility of $u(\omega)$ where $u' > 0$ and $u'' \leq 0$. I

[17]This will have useful properties for the following section analyzing comparative statics, and, in the case of matching markets, is relatively intuitive–the number of job searchers and vacancies posted are productive complements in this case.

abstract away from ownership of the firms and lenders–the profits that firms keep (which will be non-zero for operating firms) and the loan payments leave the system and are not consumed by any workers in my model. Firms and workers share a common discount factor, β.

With these features, I can identify six value functions that describe the labor market: four for workers dictating continuation values of (i) being unemployed and not in default; (ii) being unemployed and being in default; (iii) being employed and not being in default; (iv) and being employed and in default. There are two continuation values for firms dictating the continuation values of (i) having a vacancy and (ii) a filled position entering period t.

2.3.3.1 Workers

Unemployed, not in default:

$$\begin{aligned} U_t(a_i, s_t, x_t, A_t) = \max_{w} \Bigg\{ & p(\theta(a_i, w))W_t(w, a_i, s_t, x_t, A_t) + \\ & (1 - p(\theta(a_i, w))) \max_{s^u_{t+1}} \Bigg[u\Bigg(b + s_t - \rho(b, x_t, A_t) - s^u_{t+1}\frac{1}{1+R} \Bigg) \\ & + \beta\psi(b, s^u_{t+1}, x_{t+1})U_{t+1}(a_i, s^u_{t+1}, x_{t+1}, A_{t+1}) \\ & + \beta(1 - \psi(b, s^u_{t+1}, x_{t+1}))U^D_{t+1}(a_i, s^u_{t+1}, x_{t+1}, A_{t+1}) \Bigg] \Bigg\} \end{aligned} \tag{2.3.1}$$

Recall: if the worker does not match to a job, they *then* select their savings s.t. $s^u_{t+1} \geq 0$. Furthermore, note that debt and accrued interest here evolve s.t.

$$x_{t+1} = \begin{cases} (1+r)x_t - \rho(x_t, t) & \text{if Fixed Repayment} \\ x_t - \max\{\rho(\omega, A_t, x_t) - A_t, 0\} & \text{if IDR} \end{cases}$$

$$A_{t+1} = \begin{cases} 0 & \text{if Fixed Repayment} \\ \max\{A_t + (1+r)x_t - \rho(\omega, A_t, x_t), 0\} & \text{if IDR} \end{cases}$$

Finally, note that workers who enter a time period unemployed and not in default will remain not in default during that period. They can only transition to default in the *next* period, depending on their savings decision today.

Unemployed, in default

$$
\begin{aligned}
U_t^D(a_i, s_t, x_t, A_t) = \max_w \Bigg\{ & p(\theta(a_i, w)) \Big(\psi(w, s_t, x_t) W_t(w, a_i, s_t, x_t, A_t) \\
& + (1 - \psi(w, s_t, x_t)) W_t^D(w, a_i, s_t, x_t, A_t) \Big) \\
& + (1 - p(\theta(a_i, w))) \Big[u(\underline{c_{FR}}) - \delta + \beta \psi(b, 0, x_{t+1}) U_{t+1}(a_i, 0, x_{t+1}, A_{t+1}) \\
& + \beta (1 - \psi(b, 0, x_{t+1})) U_{t+1}^D(a_i, 0, x_{t+1}, A_{t+1}) \Big] \Bigg\}
\end{aligned}
\tag{2.3.2}
$$

Here, note that (i) workers pay a default penalty, δ, corresponding to various negative aspects of entering default; (ii) workers who enter period t unemployed and in default can leave default *within period t* if they become employed and their earnings are sufficiently high; (iii) there is no savings decision here, as the worker's liquid assets (above consumption threshold $\underline{c_{FR}}$) are used to pay down debt; and (iv) if they do not leave default today, there is some probability they leave next period. As I assume default probabilities under IDR are 0, the only relevant transition function is for debt:

$$x_{t+1} = x_t(1 + r) - (b + s_t - \underline{c_{FR}})$$

Employed, not in default:

$$\begin{aligned}
W_t(w, a_i, s_t, x_t, A_t) = \max_{s^e_{t+1}} \Bigg\{ & u\left(w + s_t - \rho(x_t, w, A_t) - s^e_{t+1}\frac{1}{1+R}\right) \\
& + \beta\phi\Big[\psi(w, s^e_{t+1}, x_{t+1})W_{t+1}(w, a_i, s^e_{t+1}, x_{t+1}, A_{t+1}) \\
& + (1-\psi(w, s^e_{t+1}, x_{t+1}))W^D_{t+1}(w, a_i, s^e_{t+1}, x_{t+1}, A_{t+1})\Big] \\
& + \beta(1-\phi)\Big[\psi(b, s^e_{t+1}, x_{t+1}) \max_{s^u_{t+2}} \Big(u\left(b + s^e_{t+1} - \rho(x_{t+1}, b, A_{t+1}) - s^u_{t+2}\frac{1}{1+R}\right) \\
& + \beta\psi(b, s^u_{t+2}, x^{ND}_{t+2})U_{t+2}(a_i, s^u_{t+2}, x^{ND}_{t+2}, A^{ND}_{t+2}) \\
& + \beta(1-\psi(b, s^u_{t+2}, x^{ND}_{t+2}))U^D_{t+2}(a_i, s^u_{t+2}, x^{ND}_{t+2}, A^{ND}_{t+2})\Big) \\
& + (1-\psi(b, s^e_{t+1}, x_{t+1}))\Big(u(\underline{c_{FR}}) - \delta + \beta\psi(b, 0, x^D_{t+2})U_{t+2}(a, 0, x^D_{t+2}, A^D_{t+2}) \\
& + \beta(1-\psi(b, 0, x^D_{t+2}))U^D_{t+2}(a, 0, x^D_{t+2}, A^D_{t+2})\Big)\Big]\Bigg\}
\end{aligned} \tag{2.3.3}$$

For employed workers, the continuation value of being employed is the flow benefit to being employed plus the discounted benefit to being employed tomorrow, accounting for the exogenous probability that the match dissolves and default risk. The worker selects their savings, s^e_{t+1} s.t. $s^e_{t+1} \geq 0$, and moves to the next period with these savings.

In the next period, if they become unemployed, they *must* live one period in unemployment–i.e. there is no opportunity to search. This is why the second half of this continuation value function is identical to the unemployment value function above *with no search opportunity.* Since savings decisions are made after search is complete, the savings decision in unemployment (s^u_{t+2}) is identical to the savings rule in the unemployment value function above. Note that here we must keep track of the evolution of debt and accrued interest over two periods, though the second evolution only occurs if the worker is unemployed next period. The debt and accrued interest evolve according to:

$$x_{t+1} = \begin{cases} (1+r)x_t - \rho(x_t, t) & \text{if Fixed Repayment} \\ x_t - \max\{\rho(\omega, A_t, x_t) - A_t, 0\} & \text{if IDR} \end{cases}$$

$$A_{t+1} = \begin{cases} 0 & \text{if Fixed Repayment} \\ \max\{A_t + (1+r)x_t - \rho(\omega, A_t, x_t), 0\} & \text{if IDR} \end{cases}$$

The evolution of debt and accrued interest in the second period–which, again, only occurs in the event of a match being disolved and the worker must spend one period in unemployment–is:

$$x^{ND}_{t+2} = \begin{cases} (1+r)x_{t+1} - \rho(x_{t+1}, t+1) & \text{if Fixed Repayment} \\ x_{t+1} - \max\{\rho(\omega, A_{t+1}, x_{t+1}) - A_{t+1}, 0\} & \text{if IDR} \end{cases}$$

$$A_{t+2} = \begin{cases} 0 & \text{if Fixed Repayment} \\ \max\{A_{t+1} + (1+r)x_{t+1} - \rho(\omega, A_{t+1}, x_{t+1}), 0\} & \text{if IDR} \end{cases}$$

and,

$$x^D_{t+2} = x_{t+1}(1+r) - (b + s^u_{t+1} - \underline{c_{FR}})$$

Again, there is no accrued interest under fixed repayment and there is no risk of default probability under IDR, so A^D_{t+2} is always 0.[18]

Employed, in default

$$\begin{aligned}
W^D_t(w, a_i, s_t, x_t, A_t) =& u(\underline{c_{FR}}) - \delta + \beta\phi\Big[\psi(w, 0, x_{t+1})W_{t+1}(w, a_i, 0, x_{t+1}, A_{t+1}) \\
&+ (1-\psi(w, 0, x_{t+1}))W^D_{t+1}(w, a_i, 0, x_{t+1}, A_{t+1})\Big] \\
&+ \beta(1-\phi)\Big[\psi(b, 0, x_{t+1}) \max_{s^u_{t+2}} \Big(u\Big(b - \rho(x_{t+1}, b, A_{t+1}) - s^u_{t+2}\frac{1}{1+R}\Big) \\
&+ \beta\psi(b, s^u_{t+2}, x^{ND}_{t+2})U_{t+2}(a_i, s^u_{t+2}, x^{ND}_{t+2}, A^{ND}_{t+2}) \\
&+ \beta(1-\psi(b, s^u_{t+2}, x^{ND}_{t+2}))U^D_{t+2}(a_i, s^u_{t+2}, x^{ND}_{t+2}, A^{ND}_{t+2})\Big) \\
&+ (1-\psi(b, 0, x_{t+1}))\Big(u(\underline{c_{FR}}) - \delta + \beta\psi(b, 0, x^D_{t+2})U_{t+2}(a, 0, x^D_{t+2}, A^D_{t+2}) \\
&+ \beta(1-\psi(b, 0, x^D_{t+2}))U^D_{t+2}(a, 0, x^D_{t+2}, A^D_{t+2})\Big)\Big]
\end{aligned} \tag{2.3.4}$$

Here, there is no employed savings decision as liquid assets net of the consumption threshold are used to pay debt. The unemployed savings decision in the next period–which only occurs if the worker loses their job and must spend one period in unemployment before they can

[18] This means I could drop it from the value function above, but I include it for completeness and flexibility of applying these value functions to either a fixed repayment system or an IDR system. That is, if we consider a fixed repayment system, we could drop all A_t state variables from the value functions. Similarly, if we consider a IDR system, we could drop all ψ functions and the default continuation value functions, U^D and W^D.

search again–has the same rule as above. Note that debt evolves according to:

$$
\begin{aligned}
x_{t+1} &= (1+r)x_t - (w + s_t - \underline{c_{FR}}) \\
x_{t+2}^{ND} &= (1+r)x_{t+1} - \rho(x_{t+1}, t+1) \\
x_{t+2}^{D} &= (1+r)x_{t+1} - (b - \underline{c_{FR}})
\end{aligned}
$$

Again, default does not occur under the IDR system, so there are no accrued interest equations here.

2.3.3.2 Firms

Firms with vacancy

$$
J_t^V = \max_{a_i, w} q(\theta(a_i, w))[a_i - w + \beta\phi_j J_{t+1}^E(a_i, w)] - k \tag{2.3.5}
$$

Firms with filled job

$$
J_t^E(a_i, w) = a_i - w + \beta\phi_j J_{t+1}^E(a_i, w) \tag{2.3.6}
$$

The firm side of the model is standard. Note that firms do not care about the borrowers' debt nor their repayment plan, except insofar as it affects market tightness.

2.3.3.3 Equilibrium

Equilibrium will consist of:

- A market tightness function, $\theta(a, w)$
- Value functions for the worker, $W(w, a_i, s_t, x_t, A_t)$; $W^D(w, a_i, s_t, x_t, A_t)$; $U(a, s_t, x_t, A_t)$; and $U^D(a, s_t, x_t, A_t)$
- Value functions for the firm, J^V; $J^E(a, w)$
- Law of motion for debt, $x_t(x_{t-1}, s_{i,t-1}, A_{t-1}, \omega)$ and accrued interest $A_t(A_{t-1}, x_{t-1}, s_{i,t-1}, \omega)$
- Optimal search behavior, $w_t^*(a_i, x_t, A_t)$
- Savings rules $s_t^{u*}(b, x_t, A_t)$ and $s_t^{e*}(w_t, x_t, A_t)$
- Optimal investment behavior, $z^*(a_i^0, b_i^0)$

- Laws of motion dictating the transition from unemployment to employment for workers of type a_i with debt x_t and accrued debt A_t and the transition from employment to unemployment, which depends on exogenous separation parameterϕ; and
- Laws of motion dictating transition between default and non-default states, dictated by the exogenous function $\psi(\cdot)$

Following the same logic as Menzio and Shi (2010), an equilibrium exists for these heterogeneous agents.[19] Further, such an equilibrium features block recursivity, where the individual decision rules depend only on individual features and aggregate state of the economy, not on the proportion of workers employed or unemployed at any given point in time. This makes the model more tractable and is especially useful in the next section where I consider how individual decision making changes under either (i) different loan balances or (ii) different loan repayment regimes.

2.4 Decision Rules & Comparative Statics

2.4.1 Labor Market

From the firm continuation value for a vacancy and the free entry condition, I can define the market tightness as:

$$\theta(a,w) = q^{-1}\left\{\frac{k}{\frac{1}{1-\beta\phi}[a-w]}\right\} \tag{2.4.1}$$

Firms do not care about individuals' loans or repayment plans–only those individuals' productivity and the wage they will pay. Market tightness is increasing in individual-specific labor productivity, a, as well as the probability a match continues next period, ϕ. It is decreasing in wage, w, and cost of posting a vacancy, k.

Using the continuation value of being unemployed, not in default and solving for the optimal wage an individual will search for at period t, I can define optimal search

[19]While Menzio and Shi deals specifically with ex ante heterogenous agents and on-the-job search, here agents can no longer alter their characteristics after entering the labor market, making them ex ante heterogenous from the perspective of the labor market–i.e. the initial choice of education is based on initial expectations and cannot be changed after. Further, while I don't have on-the-job search here, this is a sub-class of the model in Menzio and Shi, where the probability an employed worker can search is set to zero.

behavior as:

$$w_t = a_i - \frac{\varepsilon_{p,\theta}}{1-\varepsilon_{p,\theta}}$$
$$\times \frac{W_t(\cdot) - \Big[u\Big(b + s_t - \rho(b, x_t, A_t) - s^{u*}_{t+1}\frac{1}{1+R}\Big) + \beta\psi(\cdot)U_{t+1}(\cdot) + \beta(1-\psi(\cdot))U^D_{t+1}(\cdot)\Big]}{u'(\cdot)(1-\rho_w - \frac{\partial s^{e*}}{\partial w}\frac{1}{1+R}) + \beta\phi\Big[\psi(\cdot)\frac{\partial W_{t+1}}{\partial w} + (1-\psi(\cdot))\frac{\partial W^D_{t+1}}{\partial w} + \frac{\partial \psi(\cdot)}{\partial w}\Big(W_{t+1}(\cdot) - W^D_{t+1}(\cdot)\Big)\Big]} \tag{2.4.2}$$

I suppress the arguments to the value functions and default probabilities for clarity here. Note, ρ_w is the derivative of the repayment plan with respect to the wage, which is 0 under the fixed repayment plan. Under the IDR, it is ι in the repayment region, or 0 if earnings fall below the income disregard, $\underline{c_{IDR}}$, or if the borrower completely repays their debt and accrued interest. The elasticity of the probability of matching with respect to the market tightness is $\varepsilon_{p,\theta} = \frac{\partial p}{\partial \theta}\frac{\theta}{p}$.

Unemployed in default workers have the following optimal search behavior:

$$w_t = a_i -$$
$$\frac{\varepsilon_{p,\theta}}{1-\varepsilon_{p,\theta}} \frac{\psi(\cdot)W_t(\cdot) + (1-\psi(\cdot))W^D_t - \Big[u(\underline{c_{FR}}) + \beta\psi(\cdot)U_{t+1}(\cdot) + \beta(1-\psi(\cdot))U^D_{t+1}(\cdot)\Big]}{\psi(\cdot)\frac{\partial W_t(\cdot)}{\partial w} + (1-\psi(\cdot))\frac{\partial W^D_t(\cdot)}{\partial w} + \frac{\partial\psi(\cdot)}{\partial w}\Big(W_t(\cdot) - W^D_t(\cdot)\Big)} \tag{2.4.3}$$

Where:

$$\frac{\partial W_t}{\partial w} = u'(\cdot)(1-\rho_w - \frac{\partial s^{e*}}{\partial w}\frac{1}{1+R})$$
$$+ \beta\phi\Big[\psi(\cdot)\frac{\partial W_{t+1}}{\partial w} + (1-\psi(\cdot))\frac{\partial W^D_{t+1}}{\partial w} + \frac{\partial\psi(\cdot)}{\partial w}\Big(W_{t+1}(\cdot) - W^D_{t+1}(\cdot)\Big)\Big]$$

$$\frac{\partial W^D_t}{\partial w} = \beta\phi\Big[\psi(\cdot)\frac{\partial W_{t+1}}{\partial w} + (1-\psi(\cdot))\frac{\partial W^D_{t+1}}{\partial w} + \frac{\partial\psi(\cdot)}{\partial w}\Big(W_{t+1}(\cdot) - W^D_{t+1}(\cdot)\Big)\Big]$$

The Euler equations dictating optimal savings are:

$$\frac{1}{1+R}u'\Big(b + s_t - \rho(b, x_t, A_t) - s^u_{t+1}\frac{1}{1+R}\Big)$$
$$= \beta\Big[\frac{\partial\psi(\cdot)}{\partial s^u_{t+1}}\Big(U_{t+1}(\cdot) - U^D_{t+1}(\cdot)\Big) + \psi(\cdot)\frac{\partial U_{t+1}(\cdot)}{\partial s^u_{t+1}} + (1-\psi(\cdot))\frac{\partial U^D_{t+1}(\cdot)}{\partial s^u_{t+1}}\Big]$$

and

$$
\begin{aligned}
&\frac{1}{1+R}u'\Big(w+s_t-\rho(w,x_t,A_t)-s^e_{t+1}\frac{1}{1+R}\Big)\\
&=\beta\phi\Big[\frac{\partial\psi(\cdot)}{\partial s^e_{t+1}}\Big(W_{t+1}(\cdot)-W^D_{t+1}(\cdot)\Big)+\psi(\cdot)\frac{\partial W_{t+1}(\cdot)}{\partial s^e_{t+1}}+(1-\psi(\cdot))\frac{\partial W^D_{t+1}(\cdot)}{\partial s^e_{t+1}}\Big]\\
&\beta(1-\phi)\Big\{\frac{\partial\psi(\cdot)}{\partial s^e_{t+1}}\Big[u\Big(b+s^e_{t+1}-\rho(x_{t+1},b,A_{t+1})-s^u_{t+2}\frac{1}{1+R}\Big)+\beta\psi(\cdot)U_{t+2}(s^u_{t+2})\\
&+\beta(1-\psi(\cdot))U^D_{t+2}(s^u_{t+2})-u(\underline{c_{FR}})+\delta-\beta\psi(\cdot)U_{t+2}(0)-\beta(1-\psi(\cdot))U^D_{t+2}(0)\Big]\\
&+\psi(\cdot)u'\Big(b+s^e_{t+1}-\rho(x_{t+1},b,A_{t+1})-s^u_{t+2}\frac{1}{1+R}\Big)\\
&+(1-\psi(\cdot))\Big(\beta\frac{\partial\psi}{\partial x^D_{t+2}}(U_{t+2}(\cdot)-U^D_{t+2})+\beta\psi(\cdot)\frac{\partial U_{t+2}(\cdot)}{\partial x^D_{t+2}}+\beta(1-\psi(\cdot))\frac{U^D_{t+2}(\cdot)}{\partial x^D_{t+2}}\Big)\Big\}
\end{aligned}
$$

These optimal choice functions are recursive in nature, highlighting a particular difficulty in analyzing them analytically. However, we can still note several features of optimal behavior implied by this model. Note that worker's optimal search decision depends on whether they are in default or not, but whether an otherwise identical worker searches in a higher (lower) wage submarket when they are not in default relative to when they are in default is ambiguous. I explore this in more detail below, in addition to changing the basis of the repayment rule. Second, note that there are two effects of default risk on optimal savings, both of which induce higher savings than in the absence of default risk. First, higher savings help to avoid default altogether, by reducing the ψ term. Second, higher savings reduce the disparity between $\frac{\partial U_{t+1}}{\partial s^u_{t+1}} < \frac{\partial U^D_{t+1}}{\partial s^u_{t+1}}$ by ensuring a larger amount of outstanding debt is paid off in the event of default. This immediately highlights one way the structure of repayment plans impacts optimal decision-making: IDR plans eliminate default risk, eliminating this effect on optimal savings.

Additionally, consider how the savings decision is impacted by the required loan repayments even in the absence of the default risk. Savings here operate as self-insurance against unemployment risk, as in Chaumont and Shi (2022), who show that higher savings lead workers to search in higher wage submarkets and vice versa.[20] However, under a fixed repayment system, any savings level, s_{t+1}, results in lower wealth next period, due to the required repayment. Hence, fixed repayment systems reduce the insurance potential of

[20] An important difference here is I do not allow on-the-job search. This means firms do not care about a worker's savings in their job posting here, and workers cannot insure against lower initial earnings by searching later.

savings at a given level, and workers will tend to search in lower wage submarkets than they otherwise would. On the other hand, IDR systems act as a proportional reduction in *income*; hence, workers with a given savings level will tend to search in a higher wage submarket than they otherwise would. These results are conceptually similar to those of Ji (2021). However, these only flow from the role of savings as a self-insurance mechanism against unemployment risk and do not consider other impacts on optimal search behavior.

Next, I focus on the decision rule for optimal search in order to characterize how debt and repayment systems alter optimal behavior in this model.

2.4.1.1 Fixed Repayment Loans

Here, for expositional purposes and clarity, I focus on the search decision for an unemployed individual not in default. Furthermore, I assume that if they successfully match to a job, that job does not transition to default, i.e. $\psi(w, s_{t+1}) = 1$. First, note that under the fixed repayment system, I can sign the change in optimal wage searched for using the implicit function theorem. Namely:

$$
\begin{aligned}
\frac{\partial w}{\partial x} \approx\ & \overbrace{[-\frac{\partial W_t}{\partial x} + u'(c_u)(\rho' + \frac{\partial s^{u*}_{t+1}}{\partial x}) + \beta\psi\frac{\partial U_{t+1}}{\partial x} + \beta(1-\psi)\frac{\partial U^D_{t+1}}{\partial x} + \beta\frac{\partial \psi}{\partial x}(U_{t+1} - U^D_{t+1})]}^{?} \\
& \times \overbrace{\left[\frac{1-(\beta\phi)^{t_{\hat{FR}}-(t+1)}}{1-\beta\phi}u'(c^x_e)(1 - \frac{\partial s^{e*}}{\partial w}\frac{1}{1+R}) + \frac{(\beta\phi)^{t_{\hat{FR}}-(t+1)}}{1-\beta\phi}u'(c^0_e)(1 - \frac{\partial s^{e*}}{\partial w}\frac{1}{1+R})\right]}^{+} \\
& + \underbrace{[W_t - u(c_u) - \beta\psi U_{t+1} - \beta(1-\psi)U^D_{t+1}]}_{+} \\
& \times \underbrace{\frac{1-(\beta\phi_j)^{t_{\hat{FR}}-(t+1)}}{1-\beta\phi_j}\Big[u'(c^x_e)(-\frac{\partial^2 s^{e*}}{\partial w \partial x} - \frac{u''}{u'}(c^x_e)(1 - \frac{\partial s^{e*}}{\partial w})(\rho' + \frac{\partial s^{e*}}{\partial x})}_{+}
\end{aligned}
$$

Where $\approx$ means "is the same sign as" and $c_u = b + s_t - \rho(x) - s^{u*}_{t+1}\frac{1}{1+R}$ is consumption in the unemployed state; $c^x_e = w + s_t - \rho(x) - s^{e*}_{t+1}\frac{1}{1+R}$ is consumption in the employed state before the debt is repaid and $c^0_e = w + s_t - s^{e*}_{t+1}\frac{1}{1+R}$ is consumption in the employed state after debt is repaid. Furthermore, note that $\rho' = \frac{r}{1-\left(\frac{1}{1+r}\right)^{t_{\hat{FR}}-t}}$ under the fixed repayment plan.

Intuitively, the first line above captures the reduction in the continuation values of being employed at a given wage, which itself is separable into the "repayment burden" and the "default penalty." The former of these captures the idea that, at any given wage, higher repayments mean lower consumption, driving the individual to search for higher wage jobs. The latter captures the idea that, because of the risk of entering default in the event of unemployment, there is greater risk to not finding a job, driving the individual to search for lower wage jobs. These ideas have been captured separately in existing literature by Kaas and Zink (2011) Kaas and Zink (2011) in the context of repayment burden and Ji (2021) Ji (2021) and Abraham et al. (2018) Abraham et al. (2018) in context of default risk. The net effect here is ambiguous.

The second term reflects the curvature of the flow utility function. This second effect is unambiguously positive. To sign the second term,[21] note that an increase in debt is equivalent to a decrease in earnings such that $\partial x = -\partial w \frac{1}{\rho'}$ so the final line simplifies to[22]

$$\underbrace{\frac{1-(\beta\phi_j)^{\hat{t_{FR}}-(t+1)}}{1-\beta\phi_j}\Big[u'(c_e^x)(\rho')^2(\frac{\partial^2 s^{e*}}{\partial w^2}-\frac{u''}{u'}(c_e^x)(1-\frac{\partial s^{e*}}{\partial w})^2)\Big]}_{+}$$

Note that the second term includes the coefficient of absolute risk aversion, $-\frac{u''}{u'}$ evaluated at the employed consumption level. Combined, the net effect of an increase of loan balances on optimal search behavior is ambiguous.

Here, note that for an individual to search for lower wages in the presence of higher debt, the effects of the default penalty must outweigh the effects of the repayment burden:

$$|\beta\psi\frac{\partial U_{t+1}}{\partial x}+\beta(1-\psi)\frac{\partial U^D_{t+1}}{\partial x}+\beta\frac{\partial\psi}{\partial x}(U_{t+1}-U^D_{t+1})|>|\frac{\partial W_t}{\partial x}-u'(c_u)(\rho'+\frac{\partial s^{u*}_{t+1}}{\partial x})|$$

Here, I write the repayment burden as a measure relative to the flow utility of being unemployed. The opposite is also true: for higher debt to induce workers to search for higher wages, the repayment burden must outweigh the default penalty. Further, note that, all else equal, higher ability individuals have higher continuation values than lower ability individuals: they can command higher wages due to the fact that labor market tightness increases with labor productivity. Finally, note that the relevant term to understand repayment burden here can be rewritten as

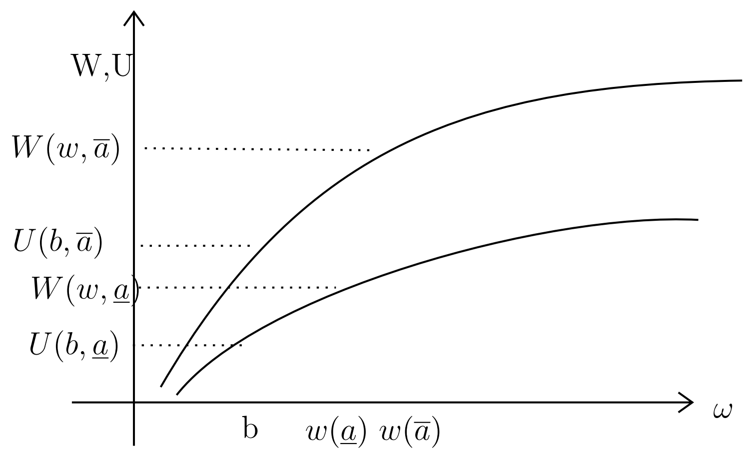
$$-\frac{\partial W}{\partial x}=\frac{\partial W}{\partial w}\frac{1}{\rho'}=-\frac{p'\theta'}{p}\{W-U\}$$

[21] Allowing a small abuse of notation

[22] Here, I use the result from Chaumont and Shi (2022) that optimal savings rules are concave in wealth to sign $\frac{\partial^2 s}{\partial w^2}>0$.

Where the first equality again comes from the fact that debt impacts wages proportionally, $\partial x = -\partial w \frac{1}{\rho'}$, and the second equality comes from the first order condition for the search problem. Now, note that it is the higher (lower) ability individuals for whom value of employment, W, likely to be sufficiently large (small) relative to the continuation value for unemployment, U, such that the repayment burden will be large relative to the default risk. The following sketches out this relationship:

Figure 2.1: Effect of higher loans on value functions on high ability, a_i, and low ability, a_j, individuals



In Figure 2.1, I sketch a contrived example to highlight that it is the higher ability people for whom the repayment burden dominates the default risk. I show the effect of higher loans on the continuation values given the optimal selection of wages for a high ability and low ability person, denoted by $\overline{a}$ and $\underline{a}$, respectively. The x-axis here measures per-period income, ω, and the y-axis the continuation values. Recall that, all else equal, a higher ability person will search in higher wage submarkets. Furthermore, because higher ability persons necessarily have higher continuation values at all incomes–including the unemployment income, b–the high ability value curve must lie above the low ability value curve. Finally, because the continuation values are simple additive combinations of concave functions, they must retain concavity.

An empirical implication of this model is that, under a FR loan system, *if* higher loan balances induce anyone to search for a lower wage, this must be concentrated among lower ability individuals. In other words, the default penalty is more likely to dominate for lower ability persons. This implication is testable in the data. To see if this implication is at least consistent with what we observe in the real world, I use the Baccarlaureate and

Beyond 1993 sample.[23] I run a simple regression–which cannot be interpreted as causal but only correlational here–of wages in 2003 on original student loan balances in 1994.[24] I also include indicators for gender and ethnicity. Then, I run the same specification separately for individuals with SAT scores in the top and bottom quartile, which I use as a proxy for ability here. The results are showin in Table 2.1.

Table 2.1: Association between loan balance in 1994 and wage in 2003

Group	Coefficient on undergraduate debt in 1994
Overall	0.285*** (0.071)
Highest Quart. SAT	0.432*** (0.137)
Lowest Quart. SAT	0.075 (0.154)

Standard errors reported in parentheses, *** = .01 sig. level, regressions include an indicator for if the individual is female an indicator for if the individual is non-Hispanic white

While this is a relatively weak test, the results are consistent with the idea that high ability persons (as measured by SAT scores) with higher levels of debt are *more* likely to search for higher wages, whereas lower ability persons exhibit no such relationship. Additionally, recall that the second term in the equation that signs the change in optimal wage given a change in debt is mitigated by the coefficient of absolute risk aversion. With decreasing absolute risk aversion, it is less likely that we will observe negative relationships between optimal search behavior and debt, since the second (unambiguously positive) term will be relatively larger for lower ability workers.

Here, it is worth noting how this result fits into the existing empirical and theoretical literature. Recall that Rothstein and Rouse (2011) found that student loans caused graduates to take on higher wage jobs post-graduation. Kaas and Zink (2011) argued, using a similar methodology to mine, that this was due to the reduced consumption at a given wage after

[23]The Baccalaureate and Beyond Longitudinal Study (B&B) 93 is a nationally representative survey of recent 1992-1993 academic year college graduates drawn from the National Postsecondary Student Aid Study (NPSAS). See National Center for Education Statistics (NCES) (1993, 2016, and 2019)

[24]A similar, albeit less significant, pattern emerges if I replace the dependent variable with April 1997 job annual salary. However, no strong relationship emerges when I use April 1994 job salary. It is possible that initial wages are compressed, and certain time variation in wages are what individuals search for rather than the static contracts I examine here. Alternatively, it is possible that individuals make additional, post-graduate investments in human capital that are not captured in this model.

accounting for loan payments. It should be noted here that Rothstein and Rouse (2011) used a natural experiment at a highly selective university, meaning their sample was likely to be mostly made up of high ability individuals. Abraham et al. (2018) and Ji (2021), on the other hand, explicitly model the default penalty, and argue that individuals are the most likely to be driven *away* from riskier but higher wage jobs they otherwise might have taken. In Abraham et al.'s model, lower ability individuals are those who would never have taken those risky jobs or wouldn't have invested in college education, and middle-ability individuals are discouraged from high wage jobs. In Ji's model, there is no directed search and individuals have homogenous abilities. Hence, this model and these results are consistent with both of these strands of literature, which emphasize different elements that are components of the model presented here; namely, the repayment burden and the default penalty. Which of these effects dominants likely differs for different segments of the population, which explains the different findings presented in previous work.[25]

2.4.1.2 Income-Driven Repayment Loans

Returning to equation 2.4.2 and rewriting it under the IDR plan, optimal search behavior is characterized by:

$$w_t = a_i - \frac{\varepsilon_{p,\theta}}{1-\varepsilon_{p,\theta}} \frac{W_t(\cdot) - u\left(b + s_t - \rho(b, x_t, A_t) - s^{u*}_{t+1}\frac{1}{1+R}\right) - \beta U_{t+1}(\cdot)}{u'(\cdot)(1 - \rho_w - \frac{\partial s^{e*}}{\partial w}\frac{1}{1+R}) + \beta\phi\frac{\partial W_{t+1}}{\partial w}} \tag{2.4.4}$$

Here, note that if we (i) shut down the endogenous savings decision, $s_t = 0$; (ii) set the income disregard such that individuals pay the same fraction of income in all states, $\underline{c_{IDR}} \leq b$; (iii) ensure there is no cap on total repayments by, for example, setting interest on debt high enough that no payments cover accruing interest, $R \to \infty$; (iv) set the repayment window s.t. borrowers never exit repayment, $\hat{t_{IDR}} \to \infty$; and if flow utility exhibits Constant Relative Risk Aversion (CRRA), then the parameter related to loan repayment, namely $(1-\iota)$, will cancel out of the numerator and the denominator. That is, if IDR loans are never repaid, then the optimal search behavior coincides with optimal search behavior in the absence of any repayment. This is the same result documented in Kaas and Zink (2011), who note

[25] Abraham et al. (2018) model the default penalty as being excluded entirely from the labor market in the subsequent period. Indeed, defaulting on loans has been associated with worse labor market outcomes, see, e.g. Bos et al. (2018), and has obvious implications for ability to borrow in other credit markets. These effects are not explicitly modeled here, but insofar as they compound the effect on the continuation value to unemployment presented here–e.g. being excluded from the labor market for an additional period would reduce the continuation value to unemployment even further–then these patterns in labor market responses to higher loan levels would be consistent with what I've presented here.

that a graduate tax–which is simply an IDR plan that never exits repayment–eliminates the distortion to labor market behavior found under fixed repayment plans.

However, if any of these conditions does *not* hold, then this will not be true, and the presence of loans will alter optimal search behavior relative to the no-loans case. To characterize how loans affect search behavior during repayment, I introduce each of these features and analyze how they affect decisions.

Repayment window

First, let $\hat{t_{IDR}} < \infty$, such that borrowers exit repayment at some point. I again sign $\frac{\partial w}{\partial \iota}$. Moving from $\iota = 0$ to $\iota > 0$ will locally be captured by this derivative, allowing me to characterize how loan repayments affect behavior. Additionally, I will continue to assume preferences are CRRA, all borrowers repay regardless of repayment history or employment state, and savings are 0.

$$\begin{aligned}\frac{\partial w}{\partial \iota} &\approx \\ &[-\frac{\partial W_t}{\partial \iota} + u'(b(1-\iota))b + \beta\frac{\partial U_{t+1}}{\partial \iota}]\left[\frac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w(1-\iota))(1-\iota) + \frac{(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w)\right] \\ &+ [W_t(w,a_i) - u(b(1-\iota)) - \beta U_{t+1}(a_i)]\frac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}[\gamma u'(w(1-\iota))]\end{aligned} \tag{2.4.5}$$

Where γ is the coefficient of relative risk aversion. This can be further simplified, using the fact that if preferences are CRRA of the form $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$, then $u'(c)c = (1-\gamma)u(c)$.

$$\begin{aligned}\frac{\partial w}{\partial \iota} \approx\ & \tfrac{(1-\gamma)}{(1-\iota)}[\hat{W}_t - u(b(1-\iota)) - \beta\hat{U}_{t+1}]\left[\tfrac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w(1-\iota))(1-\iota) + \tfrac{(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w)\right] \\ &+[W_t(w,a_i) - u(b(1-\iota)) - \beta U_{t+1}(a_i)]\tfrac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}[\tfrac{\gamma(1-\iota)}{1-\gamma}u'(w(1-\iota))]\end{aligned}$$

Where $\hat{W}_t$ and $\hat{U}_t$ are only defined for $t \leq \hat{t}$ and:

$$\begin{aligned}\hat{W}_t =\ & u(w(1-\iota)) + \beta[\phi\mathbb{1}(t \leq \hat{t}-1)\hat{W}_{t+1} + (1-\phi)[\mathbb{1}(t \leq \hat{t}-1)u(b(1-\iota)) \\ &+\beta\mathbb{1}(t \leq \hat{t}-2)\hat{U}_{t+2})] \\ \hat{U}_t =\ & p(\theta(w^*,a_i))\hat{W}_t + (1-p(\theta(w^*,a_i)))[u(b(1-\iota)) + \beta\mathbb{1}(t \leq \hat{t}-1)\hat{U}_{t+1}]\end{aligned}$$

That is, the part of the continuation value that is associated with the repayment period. Importantly, w^* corresponds to the solution of the *full* problem, not simply this sub-problem.

Thus, $\frac{\partial w}{\partial \iota}$ is the same sign as an additive term where the first term is three positive elements multiplied together and the second term is three positive elements multiplied together–that is, $\frac{\partial w}{\partial \iota} > 0$. This implies that, if w_∞ is the optimal wage and individual who has no loans to repay searches for, then $w_t > w_\infty \; \forall t \leq \hat{t}$ under this simplified IDR system.

Recall, an individual of a given level of ability faces the same market tightness function. Hence, these individuals induced into searching for higher wage jobs by the IDR system are taking on more risk, meaning they also face greater levels of unemployment. This is a form of moral hazard identified by Kaas and Zink (2011) and frequently an item of worry for economists considering these systems.[26] Alternative sources of moral hazard, e.g. where people shirk during their repayment period to keep payments relatively low, are not present in this model.[27] The reason for this is entirely due to the simplicity of the system under consideration here, notably that the repayments are due during unemployment as well as employment. Under a more complicated scheme with a floor, where workers with wages lower than this threshold are not required to repay, there would be adverse labor market effects, which I return to below. Additionally, since I have abstracted from the labor market supply decision here, there is no disutility from work. Adding this element would also alter this conclusion.

Returning to the simple IDR system, there are two features I have identified: (i) if $\hat{t} = \infty$ then $w_t = w_\infty \; \forall t$ and (ii) if $\hat{t} < \infty$ then $w_t > w_\infty \; \forall t \leq \hat{t}$. Additionally, for any individual of ability a_i, the particular value of w_t they will search for will depend on t only in determining the length of time remaining in their repayment. That is, if we extend repayment so $\hat{t}' = \hat{t} + \Delta_t$, then the optimal wage they will search for at any time t will become $w'_t = w_{t+\Delta_t}$, that is, the path of the optimal wage will have a level shift. The final feature needed to characterize the time path of the optimal wages under IDR is the difference between any two wages during repayment, $w_t - w_{t+1}$. One way to approximate this difference is to use the linear approximation at any particular time, t–that is, take the derivative of w_t with respect to t. While time is discrete, this linear approximation will capture the general shape of the time path of w_t, and in fact will capture precisely the local behavior under

[26] See, e.g., Stiglitz (2014)

[27] Although as Palacios (2014) argues, this source of moral hazard may be overstated as a potential risk, given the dynamics of human capital formation on the job. This is not present in this model.

continuous time. Thus, I can sign this as:

$$\frac{\partial w_t}{\partial t} = \dot{w}_t \approx \quad [\dot{W}_t - \beta\dot{U}_t]\left[\tfrac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w_t(1-\iota))(1-\iota) + \tfrac{(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w_t)\right]$$
$$+[W_t - u(b(1-\iota)) - \beta U_{t+1}]\tfrac{(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}ln(\beta\phi)[u'(w_t(1-\iota))(1-\iota) - u'(w_t)]$$

$$\dot{w}_t \approx \quad [\dot{W}_t - \beta\dot{U}_t]\left[\tfrac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w_t(1-\iota))(1-\iota) + \tfrac{(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w_t)\right]$$
$$+[W_t - u(b(1-\iota)) - \beta U_{t+1}]\tfrac{(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}ln(\beta\phi)\tfrac{(1-\gamma)}{w_t}[u(w_t(1-\iota)) - u(w_t)]$$

Where the second line again comes from the particular feature of CRRA preferences of the form $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ that $u'(c)c = (1-\gamma)u(c)$ and γ is the coefficient of relative risk aversion. Similar to Section 2.4.1.1 above, it's possible that $[\dot{W}_t - \beta\dot{U_{t+1}}]$ is negative. However, it is true that $\dot{W}_t, \dot{U}_t \geq 0$ when $t \leq \hat{t}$.[28] Additionally, when $t = \hat{t}-1$, it must be that $\dot{U_{t+1}} = \dot{U} = 0$, i.e. the continuation value to being unemployed will reach the stationary value, since the borrower no longer makes payments from that period forward. Thus, in the last period, at least, we know that this term is positive.

Noting that $\beta\phi \in (0,1)$, this implies that if $\dot{W}_t = \beta\dot{U_{t+1}} > 0$ then $\dot{w}_t > 0$. Additionally, this is a sufficient, but not necessary, condition, since the second term above is unambiguously positive.

Using these facts

(i) $\hat{t} = \infty \Rightarrow w_t = w_\infty$[29]

(ii) $\hat{t} < \infty \Rightarrow w_t > w_\infty \ \forall t \leq \hat{t}$ and $w_t = w_\infty \ \forall t > \hat{t}$

(iii) w_t depends on t only through the amount of time left in repayment while $t \leq \hat{t}$

(iv) If $\dot{W}_t - \beta\dot{U_{t+1}} > 0$, $\dot{w}_t > 0$ for $t \leq \hat{t}$[30]

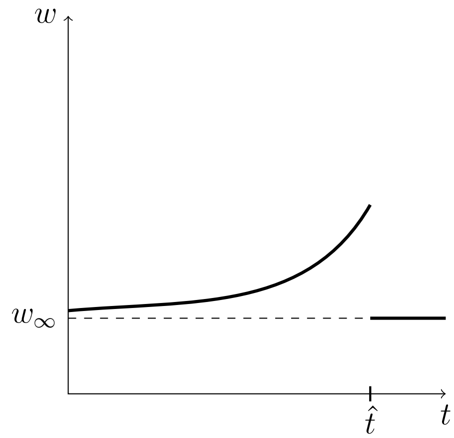
I can characterize the time path of the optimal search behavior during repayment for borrowers in the IDR plan, shown in Figure 2.2.

[28] Since there are fewer periods of payment, the individual gets to keep more of their income moving forward, improving the value.

[29] And, relatedly, that $w_t \to w_\infty$ as $\hat{t} \to \infty$

[30] For the time being, I will ignore the possibility that $\dot{W}_t - \beta\dot{U_{t+1}} \leq 0$, which, again, cannot be true in period $\hat{t}-1$ and is a necessary but not sufficient condition for $\dot{w}_t < 0$

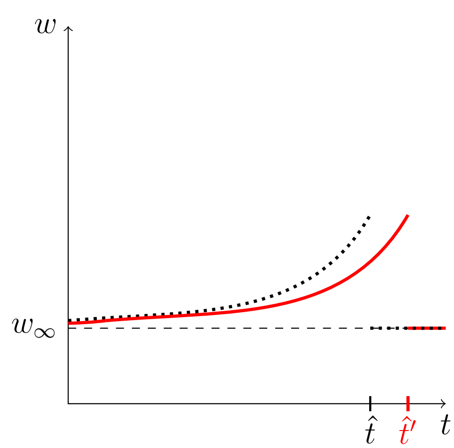
Figure 2.2: Optimal wage with IDR



The intuitive explanation for this is that, initially, from the perspective of the borrower they will be paying off their loan for so long that they ignore (or heavily discount) the time they will eventually be rid of the loan. Thus, their behavior closely mirrors the optimal search behavior of a non-borrower. However, as the time they will eventually pay off her loan approaches, the moral hazard associated with the IDR system becomes more and more prevalent, driving up their optimal search behavior until $t = \hat{t}$, the last period they must repay. From then on, they have no more repayments and behave like a non-borrower.

The third fact above also lets me characterize how a change in the repayment plan length affects search behavior, i.e. changing from $\hat{t}$ to $\hat{t}'$, shown in Figure 2.3.

Figure 2.3: Optimal wage with IDR; increase $\hat{t}$



Lengthening the repayment period under the IDR simply shifts the path out horizontally, with the optimal wages in the earlier periods falling below the optimal wages under the shorter repayment plan.

The intuition here is the same as above–with a longer repayment plan, the borrower discounts the time they will be rid of their loan even more; hence, their behavior looks even more like someone with no repayments. However, as the time they complete repayments approaches, their behavior exhibits the moral hazard described above. Since the time they will be fully repaid has extended, the time this moral hazard begins to worsen is also extended.

Income disregard

Now, suppose $b \leq \underline{c_{IDR}}$, meaning that there is some income range for which there are no required repayments under the IDR. For simplicity of exposition, allow $b = \underline{c_{IDR}}$, s.t. only unemployed individuals do not pay.[31] There are two effects here: (i) the income disregard increases flow utility while in unemployment, offering additional insurance to this state, and (ii) the income disregard *also* increases flow utility to employed states, by the lump sum ιb. This implies that the income disregard increase all incomes *regardless of employment state* by the same lump sum, ιb, while the borrower is in repayment. Hence, we can evaluate the impact of the income disregard on optimal search by examining how this income increase in all states affects marginal utilities. With a slight abuse of notation, we can consider an increase in all income states, ω:

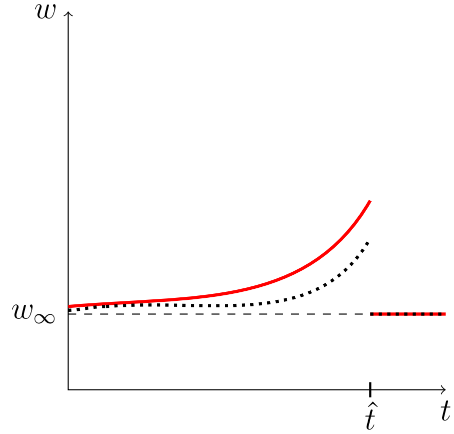
$$\begin{aligned}\frac{\partial w}{\partial \omega} \approx [-\frac{\partial W_t}{\partial \omega} + u'(b) + \beta\frac{\partial U_{t+1}}{\partial \omega}]\left[\frac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w(1-\iota)+b\iota) + \frac{(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w)\right] \\ + [W_t(w,a_i) - u(b(1-\iota)) - \beta U_{t+1}(a_i)]\frac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}[-u''(w(1-\iota)+b\iota)]\end{aligned} \tag{2.4.6}$$

As above, this suggests that all $w_t > w_\infty$ $\forall t \leq \hat{t}$ and that the optimal search path has a similar shape. Note, too, that $\frac{\partial W_t}{\partial \omega} < \frac{\partial U_t}{\partial \omega}$ due to the curvature in flow utility function; hence, the insurance provided by the income disregard will result in a steeper growth path relative to the case considered above, shown in Figure 2.4.

This result is relatively intuitive: the introduction of the income disregard means insures the unemployed state, which induces riskier search behavior. While it also decreases the

[31] Higher income disregards would generate additional insurance over low paying jobs as well.

Figure 2.4: Optimal wage with IDR; no income disregard (black) vs. income disregard (red)



repayment burden–note that $w(1-\iota) < w(1-\iota) + \iota b$–which would result in less-risky search behavior, this effect is dominated by the insurance effect.

Savings

Reintroducing savings into the model complicates the problem, as noted above, limiting our ability to analytically solve the model due to its recursive nature. However, as already noted, the IDR plan operates like a proportional reduction in income for any given wage level. Following Chaumont and Shi (2022), we can infer that this implies that borrowers under an IDR plan will search in relatively higher submarkets compared with non-borrowers who hold the same level of savings. Hence, the IDR plan counteracts the role of savings as a self-insurance mechanism.

Repayment Limit

Finally, consider a repayment limit wherein if borrowers ever completely repay their debt and any accrued interest, they immediately exit repayment. Note that, similar to the repayment window above, this shifts the amount of time a borrower is in repayment. However, unlike the repayment window–where shorter repayment windows result in borrowers searching in higher wage markets at all time periods, here we see the opposite effect–that is, a repayment limit tends to mitigate the distortionary impact of the IDR plan somewhat.

To see this, note first that if a borrower is matched to a job with wage $\tilde{w}$ under which they expect to repay their debt prior to the end of the repayment window, which I will call $\tilde{t} < \hat{t_{IDR}}$, then they exit repayment and the impact of their repayments on the marginal flow utilities drops out of the relevant equation. That is, if we again consider the relationship:

$$\frac{\partial w}{\partial \iota} \approx \\ [-\frac{\partial W_t}{\partial \iota} + u'(b(1-\iota))b + \beta\frac{\partial U_{t+1}}{\partial \iota}]\left[\frac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w(1-\iota))(1-\iota) + \frac{(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}u'(w)\right] \\ + [W_t(w,a_i) - u(b(1-\iota)) - \beta U_{t+1}(a_i)]\frac{1-(\beta\phi)^{\hat{t}-(t+1)}}{1-\beta\phi}[\gamma u'(w(1-\iota))] \tag{2.4.7}$$

But now for a worker who accepts a job that will pay off debt after $\tilde{t} < t_{\hat{IDR}}$ periods, then all the elements $u'(\tilde{w}(1-\iota))(-w)$ for time periods $t > \tilde{t}$ and $t \leq t_{\hat{IDR}}$ will drop out of the $-\frac{\partial W_t}{\partial \iota}$ element. Hence, this effect will shrink in magnitude. How long the borrower expects to remain in repayment affects how much the optimal wage decreases–if they expect to immediately repay their debt with the wage, then there is no distortionary effect. However, if they expect to remain in debt for a longer period of time, recall from above that the length to the end of the repayment window *also* tends to decrease distortionary search. Note: here it be that the optimal wage decreases so much here that the borrower no longer expects to repay their debt–that is, I am exclusively focusing here on how a *binding* repayment cap affects behavior.

Intuitively, the repayment cap limits the repayment burden–the effect that drives borrowers to search in higher wage submarkets. That is, since the borrower gets to keep a relatively larger fraction of their income under the repayment cap compared to the no-cap case, the risk of unemployment has more "bite." This result is conceptually similar to that of Quiggin (2014), who argues that borrowers who repay their debt almost certainly do not exhibit moral hazard in work intensity. Here, it is not true that the cap completely eliminates the distortionary search behavior, but it is also not true that they repay with probability 1, since they are always at risk of unemployment.

Income Driven Repayment–parameters

Here, I have described how IDR plans alter search behavior relative to a no-loans case. While I can replicate the finding of Kaas and Zink (2011) that a graduate tax can result in no distortionary search behavior relative to non-borrowers, this is only true if repayment lasts forever, there is no income disregard, there is no repayment limit, and there are no savings. I then consider each of these features of an IDR and how its policy parameters impact optimal search. There are several lessons.

First, length of repayment impacts optimal search behavior –the longer the time remaining in repayment, the less optimal search behavior deviates from the no-loan case. Second, the income disregard tends to exacerbate the difference between optimal search under the IDR and the no-loan case. The income disregard increases insurance for unemployment (or, more generally, low-wage states), which allows workers to search in higher wage submarkets with lower cost. I note here that this does *not* imply that income disregards are inefficient. The model here does not consider optimal unemployment insurance, e.g., which the income disregard can be construed as with respect to loan debt. These two features of IDR have been noted by existing critiques of particular policy parameter choices by, e.g., Looney (2022). Here, the model highlights a form of the moral hazard that Looney argues can be induced with high income disregards and shorter repayment windows.

Third, the introduction of repayment limits can reduce the difference between optimal search behavior of borrowers under the IDR and non-borrowers. This is because the repayment limit also limits the repayment burden associated with debt, which is the key feature driving workers to search in higher wage submarkets. It should be noted here that this is not a costless feature of IDR: obviously, it reduces revenues. However, in a "fairness' sense, it also limits borrowers' repayments to what they actually borrowed (plus interest). This feature of the model, too, corresponds to existing work in IDR analysis by Quiggin (2014), who argues that the repayment limit is a key way of mitigating moral hazard.

Finally, savings here may interact with optimal search in interesting ways, notably by driving a wedge between what otherwise identical agents–one of whom is a borrower and one of whom is not–do. However, the recursive nature of this problem and the optimal choice functions make it difficult to analyze analytically, limiting insight to what we can infer from existing work on similar problems of savings and directed search by Chaumont and Shi (2022).

2.4.2 Investment

While the crux of this paper examines search behavior of borrowers, I briefly return to the investment decision here. Recall, investment is a static decision in this model–in period 0, the individual with ability a_i^0 and wealth b_i^0 will maximize their expected lifetime welfare by making a dichotomous decision to attend school or not. If they do attend, they pay pecuniary and non-pecuniary costs–c_i and $v_a(a_i^0)$, respectively–and receive augmented ability parameter, $a_i = f(a_i^0)$. If they do not, they move onto the labor market with their original ability. Additionally, any costs of education that are larger than initial wealth are financed with student loans with known repayment plans, $x_{i,1} = \max\{c_i - b_i^0, 0\}$. Any unspent wealth becomes initial savings on the labor market., $s_{i,1} = \max\{(b_i^0 - c_i)(1+R), 0\}$

Hence, an agent will attend school if and only if:

$$\beta U_1(f(a_i^0), s_{i,1}, x_{i,1}, 0) - v_c(c_i) - v_a(a_i^0) \geq \beta U_1(a_i^0, \frac{1}{1+R} b_i^0, 0, 0)$$

Note that here I add a function $v_c(\cdot)$ that converts pecuniary costs to utility. Additionally, I have omitted any flow utility from the investment period–though kept the flow *dis*utility from attending college. This is implies that college-goers and non-goers have the same flow utility. Changing these assumptions will impact choices, but does not change the general points we can draw from this.

With reasonable assumptions on the augmenting function $f(\cdot)$, this implies that, all else equal, higher ability individuals will be more likely to attend college, as they have both greater expected returns on the labor market and lower non-pecuniary costs. This also implies that higher *wealth* individuals will be more likley to attend college, as they will have lower debt, $x_i, 1$, which translates to higher expected returns and potentially no debt and some savings, $s_i, 1 > 0$, which offers precautionary savings value to the worker.

2.5 Efficiency and Discussion

A natural question here is which system is *efficient*. Here, I offer some brief thoughts on efficiency. Here, I consider efficiency as a case where (i) workers have identical optimal search behavior regardless of their debt status and (ii) agents make college attendance decisions independently of their initial wealth endowment. Note that, if (i) holds, then the search behavior will be efficient in the commonly used sense in search models that agents search where the social planner would also assign them to search if the planner could not overcome search frictions (see, e.g. Menzio and Shi (2010)).

As noted in the discussion on IDR plans above, it is possible to achieve no distortions between borrowers and non-borrowers under a graduate tax, as in Kaas and Zink (2011), but this greatly restricts the model. That is, if agents have CRRA preferences, cannot save, remain in repayment forever, have no income disregard, and have no cap on how much they repay in aggregate. If we change any of these–perhaps most notably, if we allow endogenous savings–then this will not be true, in general.

However, in the discussion on IDR plans, I highlighted some lessons that can *mitigate* the distortions between borrowers under an IDR plan and non-borrowers. That is, if the repayment plan is longer, the income disregard is low (or nonexistent), and there is a repayment limit, then borrower behavior will more closely resemble non-borrower behavior–at least early in their repayment window.[32] Here, it's worth noting again that these conclusions are within the context of this model. If I were to introduce, for example, an optimal unemployment insurance problem, then it may be that the "optimal" income disregard is not zero. Nevertheless, the model here suggests, at the very least, that longer repayment windows, lower income disregards, and a repayment limit all serve to mitigate distortionary behavior between borrowers and non-borrowers, features of IDRs that have been highlighted by other researchers (e.g. Looney (2022); Quiggin (2014)).

Turning to education, note that higher ability agents are more likely to attend college, but so are higher wealth agents, all else equal. The efficient policy I discuss here would encourage the former, but discourage the latter. If the only policy tool available is the loan repayment plan–i.e. if the non-pecuniary costs and pecuniary costs are exogenous, as would be the case if, e.g., these were choices made by the college–then the optimal repayment contract would institute *higher* rates for wealthier students and *lower* rates for more able students. Note that this type of optimal financing scheme–one that accounts for (i) realized income, (ii) ability, and (iii) wealth endowment in repayment rates is conceptually similar to that described by Lochner and Monge-Naranjo (2016) and Lochner et al. (2021).

2.6 Conclusion

In this paper, I have presented a structural model that combines an investment decision, a known financing rule for this investment, and a competitive labor market with search frictions. I have shown that, within the context of this model, the repayment plan for the investment financing plays a crucial role in optimal labor market behavior–in particular the appetite to take on risk in exchange for higher wages.

With a fixed repayment system, where individuals make a fixed payment that is a function of the total amount borrowed each period they are employed that also has an

[32]Furthermore, if borrowers also tend to experience wage growth, then as they approach the end of their repayment window, they will also be more likely to hit the repayment cap, which, again, acts to mitigate distortions between borrower and non-borrower behavior.

explicit penalty for failure to pay in the event of unemployment, higher loan balances have two, off-setting effects. The first is the repayment burden–each period they are employed, they can consume less since their payments are higher. This pushes individuals to take on more risk in exchange for higher wages, an effect documented by Rothstein and Rouse (2011) and studied theoretically by Kaas and Zink (2011). The second is the default penalty, the cost associated with failure to pay in a given period. This causes individuals to avoid risk and take lower wages, and effect studied by Abraham et al. (2018) and Ji (2021). Which of these effects dominates differs for different subsets of the population, relating specifically to their ability and thus their earning power.

Under an income-driven repayment system, where individuals make payments that are a function of their per-period income rather than the amount borrowed, the crucial parameters are the income disregard–the amount of income below which no payments are due–the repayment window–the length during which the borrower is in repayment–and the repayment cap–the maximum amount the borrower can repay, which can be a function of time. Under a simple case where repayments are always a fixed fraction of income, even in unemployment; there no end to the repayment period; and there are no savings, borrower behavior in the labor market coincides with that of non-borrowers. This is a graduate tax, and this result coincides with that of Kaas and Zink (2011). However, if any of these is changed, this is no longer true in general, though the deviation between borrower and non-borrower behavior under an IDR can be mitigated through these parameter choices.

While I think this model has useful conclusions regarding how individual behavior changes under different financing schemes, there are a number of simplifying assumptions that limit this realism. I do not allow on the job search; dynamic contracts; other forms of human capital investment, like schools of different qualities or on-the-job learning; endogenous labor supply decisions; or model the education supply decision. Each of these would potentially add new insights into the model, though it would also complicate it further.

In addition to these, I only consider downside risk in this model, i.e. $(1-\phi)$ is the exogenous probability a match dissolves next period. However, I could also consider upside risk, in which case there would be a transition matrix to both a shut-down level of productivity as well as higher levels of productivity, or heterogenous risk. In this case, I would also need to consider state- and firm-type-dependent wage contracts, a simple one of which would be a sharing rule where workers get some share of total output from the match. The usefulness of this consideration would be to highlight how income-contingent plans could have ex-post higher returns to the lender than fixed repayment loans. Note, however, that one result of this would be further segmentation of the market in terms of ability and hence wages. While I do not focus on it here, a result of this model is that higher ability people are able to command higher wages. If we consider a mean-preserving spread of risk in this case, the risk aversion in utility will cause supply of workers for those wages to drop. Hence, the equilibrium wages will have to increase to induce workers to accept this additional risk. Since these will already be those who can command higher wages, the result is further segmentation between high- and low-ability workers in equilibrium. While not the focus of this paper, this result itself is interesting.

Finally, I have also ignored the roles of specific skills in occupations. A richer version of this model could include investments that improve specific skills, jobs that require use of certain skills to produce output, and individual aversion to using certain skills. While I believe this would capture the concept of occupation in a better way than simply wage and job risk, it also complicates the solution to the model significantly.

The theory presented here offers some useful insight into the way that repayment structure can affect individual labor market behavior. However, it is also important to be cognizant of the significant limitations imposed in the model. While certain extensions may richen and strengthen the model, I would still be wary of strong conclusions, especially in the realm of efficiency. Additionally, this model would benefit from calibration to real world data, which would allow for empirical conclusions to the theory outlined above–especially in the cases where the theoretical result is ambiguous.

CHAPTER III

Estimating the Impact of the Opioid Epidemic on Population Mortality: Competing Risks with Economic, Demographic, and Spatiotemporal Heterogeneity

(with John Bound, Timothy A. Waidmann, and Arline T. Geronimus)

3.1 Introduction

Prior to the COVID-19 pandemic, long term upward trends in life expectancy in the U.S. had begun to stagnate, and even reverse. A recent report by the National Academy of Sciences identified rising mortality since the year 2000 among adults between 25 and 64 years of age and the primary driver of this stagnation, and it identified increases in drug and alcohol-related deaths as an important cause of increases in these rates (Committee on Rising Midlife Mortality Rates and Socioeconomic Disparities (2021)).

As measured by the Centers for Disease Control and Prevention, age-adjusted death rates from drug overdose increased more than four-fold between in the first two decades of the 21st century, with deaths due to opioids accounting for 85 percent of that growth (Hedegaard et al. (2021)). When estimating drug-related death rates and their effect on life expectancy, a common approach is to simply use deaths coded as overdose on death certificates(Gomes et al. (2018); Hall et al. (2022); Spencer et al. (2022)). However, understanding the impact of the opioid epidemic on overall death rates or life expectancy, or a change in any other cause of death, requires assumptions about the counterfactual. In this case, what, if anything would have happened to the rates of other causes had the dramatic rise in opioid use not occurred? Under the assumption that the risks of death from different causes are uncorrelated it straightforward to calculate the impact of any single cause (see, e.g., Preston et al. (2001)). While researchers often make the assumption of independence, they do so for lack or tractable alternatives and not out of any belief in the validity of this assumption. Indeed, it is easy think of reasons why causes might not be independent of each other.

In the case of opioid use or any health behavior that may increase the risk of multiple causes of death, and that may itself be correlated with other health risks, standard demographic methods are likely to lead to inaccurate counterfactuals. Health behaviors like opioid use may be direct drivers of increases in mortality from multiple causes, implying that standard demographic methods would understate the impact of the opioid mortality of death rates and life expectancy. At an even more basic level, relying on death certificates properly coding opioid overdose deaths may undercount the impact of opioids due to variation in coding behavior among medical examiners and coroners (Slavova et al. (2015); Ray et al. (2016)). However, opioid use could also act as a mediating factor for other social forces that affect a broad array of causes of death. What is more, those who die from opioid overdoses could, plausibly be drawn from an unhealthy population. In either of these two cases, standard methods will over-estimate the impact of opioids on mortality. Statistically distinguishing between such counterfactuals with observational data would appear to be difficult if not impossible (Cox (1962); Heckman and Honoré (1989); Tsiatis (1975)).

Epidemiologists have often used variation across space and time to identify the effect of an epidemic on overall mortality (e.g. Goldstein et al. (2012)). Thus, for example, in the context of the COVID 19 epidemic, authors have compared overall mortality rates in local areas hit particularly hard by the epidemic to other areas hit less hard to gauge overall effects (e.g. Lee et al. (2022)).[1] In the context of the opioid epidemic, Glei and Preston (2020) have formalized this approach to estimate the overall effect of the opioid epidemic on mortality and life expectancy.[2]They estimate that the overall impact of drug use on mortality for US men and women in 2016 was over twice as large and estimates that simply count overdose deaths. As the authors acknowledge, the validity of this method relies on several assumptions, including that geographic variations in drug overdose death rates are driven by variation in drug use and that there are no other factors driving both drug overdose deaths and the risks of death due to other causes. There are good reasons to doubt these assumptions, however.

Indeed, common narratives regarding the epidemic emphasize the role of economic

[1]Similar approaches have been used in the context of natural disasters.

[2]Glei and Preston build on earlier work by Preston et al. (2010) and Fenelon and Preston (2012) which used spatiotemporal patterns between cause-specific death rates and deaths attributable to lung cancer to estimate the overall impact of smoking on mortality and life expectancy.

and psychosocial factors in explaining the recent increase in mortality due to substance use. Chronically depressed areas of the country such as the Appalachian region were early epicenters of the opioid epidemic. Monnat (2019) and Monnat (2022) describes the evolution of the epidemic, starting in Appalachia and spreading to parts of the industrial Midwest. Other researcher have found evidence that opioid overdose deaths have risen more rapidly in areas of the country that have seen particularly steep declines in manufacturing employment (Charles et al. (2018); Hollingsworth et al. (2017); Seltzer (2020)) or that have been particular hard hit by imports from China (Dean and Kimmel (2019); Pierce and Schott (2020)).

Quite plausibly economic stagnation could be associated with increased mortality even were it not for the opioid epidemic. There is ample evidence of an SES gradient in health and mortality (Williams and Collins (1995)), with researchers identifying economic and psychosocial distress as an important factor in a variety of causes of death through the body's natural hormonal reaction to stress (Geronimus et al. (2006); McEwen (1998); McEwen and Stellar (1993); Seeman et al. (2001)). Importantly, the stress pathway has been shown to affect both cardiovascular disease and cancer, the two causes that account for the largest number of deaths nationwide, by far. Other researchers have documented the association between labor market structures, socioeconomic deprivation and access to regular medical care, which may particularly affect those with chronic disease (Committee on Community-Based Solutions to Promote Health Equity in the United States (2017); Madrian (2006); Thomson et al. (2022)).

Further, the nature of opioid-related mortality has changed over time, and the burden of these deaths is heterogeneous within the population. In the early part of the period examined by Glei and Preston, the largest share of opioid overdose deaths was attributed to prescription opioids containing oxycontin. Since 2016, however, fentanyl has been the predominant cause of opioid overdose (Hedegaard et al. (2021)). If the long term use of prescription opioids increases mortality rates for other causes (Ray et al. (2016)), but fentanyl overdose is sufficiently fatal that the effect of fentanyl use on other risks is inconsequential, the Glei and Preston estimates may overstate the current amount of spillover to other mortality.

This paper examines both the implicit assumptions of the Glei and Preston analysis: that opioid use is exogenous to other risk factors for mortality and that the spillover from drug mortality to other causes is homogenous over time, and type of drug involved. To do so, we use a statistical framework similar to the one used by Glei and Preston, but test for the potential confounding effect of socioeconomic factors using techniques that have been developed in econometrics and epidemiology (Robins (1999); Robins et al. (2000); Hirano and Imbens (2004)). To best account for varying socioeconomic conditions that may drive mortality outcomes, we use commuting zones, rather than states, as our unit of analysis(USDA ERS–Documentation (n.d.)). Controlling for the potential confounding effect of observable socioeconomic factors that vary across the geographic areas we use in our analysis substantially lowers the magnitude of estimated impact of opioids on non-overdose deaths.

In addition, we compare the period from 1990-2009 to the period from 2009-2017. Here we find significant differences in the magnitude of opioid use's impact on total mortality. While in the earlier period we find estimates of the same magnitude as Glei and Preston, we found little evidence of opioid impact on non-overdose deaths after 2009. These findings suggest that the type of drug involved matter to estimates of mortality spillovers. Indeed, specifications that split opioid overdose mortality by opioid type suggest that the estimated effects in the early period, 1990-2009, are driven by prescription opioids rather than illicit opioids.

3.2 Simple Model

Consider the relationship between opioid use and the risk for opioid overdose. As the underlying risk of opioid overdose in the absence of opioid use is zero, only the use of opioids can expose an individual to overdose. Following Glei and Preston (2020), this implies that opioid-coded mortality can be used as an indicator for opioid use within a geographic region at a particular point in time. Because it is possible that opioid use may increase other-cause mortality as well, only using opioid-coded mortality to estimate the impact of opioid use on total mortality understates the full impact.

As noted by Glei and Preston, such an approach also avoids selection bias on an individual level: aggregating up to geographic measures of mortality means they do not need to model individual selection into opioid use that may be correlated with other-cause mortality as well. However, this assumption is only true if the selection mechanism is not also correlated with geography and time. This may be a reasonable assumption for certain risk exposures—for example, smoking, as in Preston et al. (2010) and Fenelon and Preston (2012)—but as we note above, there are well documented hypotheses that the spatiotemporal selection into higher opioid use is not exogenous to other-cause mortality: that is, additional mechanisms may drive *both* higher opioid use rates (and thus opioid overdoses and other-cause deaths caused by opioid use) and other-cause deaths directly.

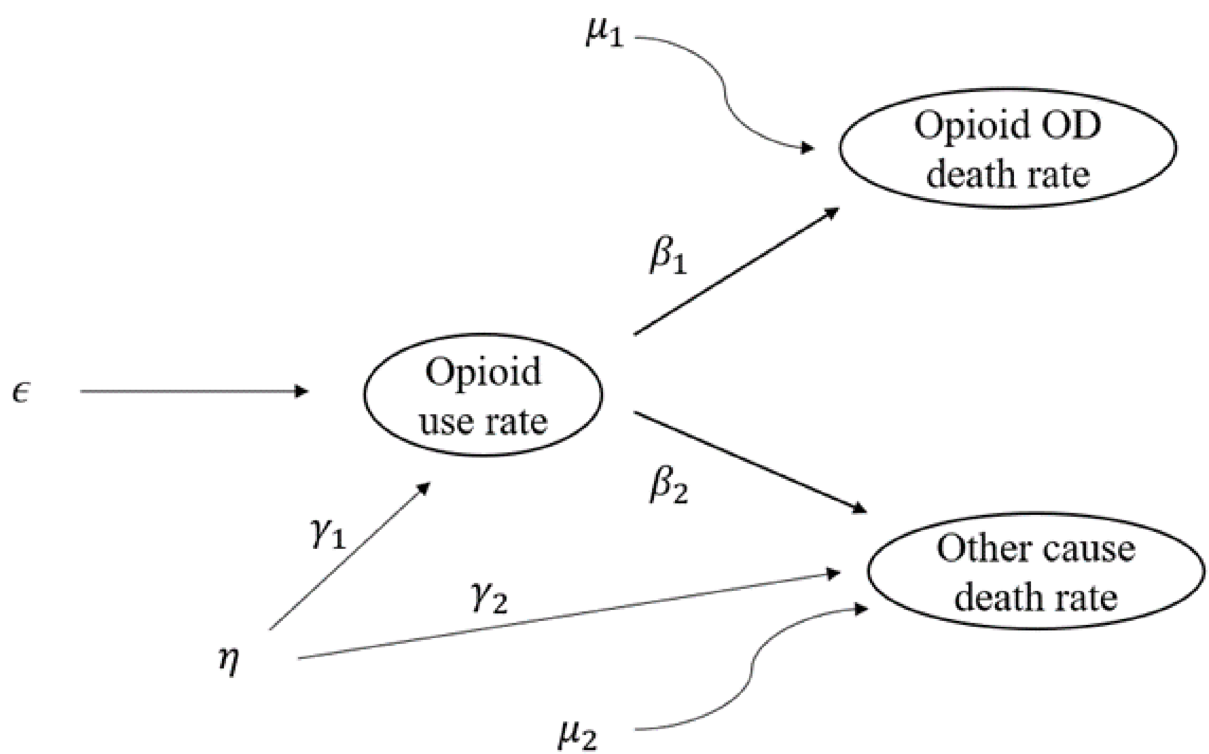
Figure 3.1 illustrates this point with a simple diagram. Here, opioid use rates impact both opioid overdoses, captured by β_1, and other-cause death rates, captured by β_2. In the absence of endogenous selection into opioid use rates, the Glei and Preston method uses opioid mortality as a proxy for opioid use with geographic variation in opioid mortality capturing the geographic variation in opioid use rates This allows for the estimation of the effect of opioid use on other-cause mortality.[3] However, if there is a third factor that affects both opioid use and other-cause mortality—here captured by η through the channels γ_1 and γ_2, respectively—then this method will not consistently measure the relationship between opioid use and other-cause mortality. In our estimation approach, we control for baseline differences between geographies using geographic fixed effects as well as common time-trends using year fixed effects. Hence, any endogeneity impacting identification must occur at the same level of observation as our identifying variation: the variation across geographies over time. However, it is plausible that there is precisely this type of endogenous selection that could impact both opioid use rates as well as non-opioid mortality rates, notably through economic deterioration that occurred unevenly across the United States over this time.

If there is a suitable measurement to capture the channel γ_1, however, we can condition our estimates on this mechanism, which can again recover the relationship between opioid-use and other-cause mortality. As described in our Methodology in Section 3.4 below, econometric techniques pioneered byRobins (1999); Robins et al. (2000); and Hirano and Imbens (2004) address this endogenous selection. This approach only controls for endogenous selection on observables. If there is also selection on unobservables, then this approach will still lead to biased measures of the relationship between opioid use and other-cause mortality. Hence, our results should be interpreted with this caveat in mind. In that sense, we emphasize here that such selection on observables is an important factor that should be considered when estimating these types of empirical models, not that it is the only factor that should be considered.

Additionally, note that the presence of μ_1 implies there is some variance in the measurement of opioid overdose deaths exogenous to the direct channel captured by β_1—note that this does not mean that there is underlying risk of opioid overdose in the absence of opioid use, simply that there could conceivably be variation in this risk relationship. This

[3]While we cannot separately estimate the two beta terms, we can estimate a scaled version of β_2. Details of this approach are available in Appendix B.

Figure 3.1: Simple model illustrating the relationship between opioid use, opioid overdose, and other mortality



Note: if there is no endogeneity, i.e. $corr(\varepsilon, \mu_2) = 0$ and $\gamma_1 = 0$, then the standard method recovers the (scaled) parameter of interest, β_2; however, if there is endogenous selection on observables, $\gamma_1 \neq 0$ and some elements of η are observable covariates, then the standard method will not, in general, recover the parameter of interest. In cases where these observables are correlated across both time and geographies, then standard methods to control for endogeneity without directly controlling for observables, e.g. inclusion of fixed effects, will not recover the parameter of interest. However, the reweighting approach formalized by Robins (1999) and Robins et al. (2000) can recover this parameter. Finally, if there is also selection on unobservables, $corr(\varepsilon, \mu_2) \neq 0$ or elements of η that are unobservable vary across both time and geography, then stronger assumptions are necessary. We note that the method we employ here cannot address this type of selection on observables; hence, our results here are limited to the point that selection on observables are important to consider, not that our method recovers unbiased estimates (in the presence of possible selection on unobservables).

captures an errors-in-variables issue with using opioid mortality as a proxy for opioid use. Intuitively, the smaller the variance of this error, μ_1, the better opioid mortality is as a proxy for opioid use, since most of the variance in opioid mortality must be driven by variation in opioid use and not this error term. On the other hand, the larger this variance, the less useful opioid mortality is as a proxy, which would lead our method to attenuated estimates of the relationship between opioid use and other-cause mortality.

Finally, note that it is possible that both β_1 and β_2 could vary over time. For our purposes, we suggest that it is plausible that, as the dominant form of opioids shifted from prescription opioid to more potent synthetic and other illicit opioids, the strength of the relationship between opioid use and opioid overdose, β_1, could increase. Our model suggests that this would attenuate the estimated relationship between opioid overdoses and other-cause mortality, even in the absence of any change in β_2. This, too, suggests the need for careful interpretation of the model in the context of potential heterogeneity over time.

3.3 Data

Data for these analyses came from several government and private sources. To calculate cause-specific death rates at the commuting zone (CZ) level, we obtained a restricted version of the Multiple Cause of Death (MCD) file including county of occurrence for each death record (National Center for Health Statistics (2020)) and the public use microdata samples from the 1990 decennial census (5%) and the five-year American Community Survey (ACS) for 2019, which includes respondents from the 2015-2019 surveys (Ruggles et al. (2022)).

We limit our analysis to non-Hispanic white men and women. We make this choice, in part, because of the concentration of opioid deaths among this group in the early years of the epidemic, but also because the identification strategy relies on the use of relatively small geographic areas, and the number of areas with sufficient non-white or Hispanic populations are small. Further, the strategy relies on the relative geographic stability of populations, and the share of Hispanic and Asian populations who are recent immigrants reduces the plausibility of that assumption.

Following methods used by the CDC (Centers for Disease Control and Prevention (2022b); Hedegaard et al. (2021)) and other authors (Hollingsworth et al. (2017)) we used the MCD files' underlying cause of death variable and record axis codes to define opioid-related deaths. Drug overdose deaths involving opioids were defined using ICD-9 for 1989-1991 and ICD-10 for 2016-2018. First, we identified all deaths with an underlying cause and manner of death listed as drug poisoning or drug addiction, excluding those where the manner of death was homicide (ICD9:E850.0-E858.9, E950.0-E950.5, E962.0,E980.0-E980.5; ICD10: X40-X44, X60-X64, Y10-Y14). Among these drug-related deaths those classified as opioid-related were those where an opioid was among the substances listed in at least one "record axis" code (ICD9:E850.0-E850.2,965.00-965.02,965.09; ICD10:T40.0-T40.4,T40.6).

We also used Census and ACS data to construct the measures we use to standardize on socioeconomic conditions. Drawing from Monnat (2019)'s examination of determinants of opioid mortality in the United States, we create commuting zone level measures of the shares of the population living in poverty and the share receiving public assistance income and the shares of adults who have less than college education, are not employed, have a work disability, and are divorced or separated. To measure geographic variation in long-term conditions we create levels for each measures in 1990 and the change (in natural log) between 1990 and 2007, just before the great recession. We also measure long term change (in natural log) in population between 1980 and 2007.

3.4 Methodology

Competing risks theory is the workhorse model for demographic analysis of multiple-decrement processes, offering a powerful tool to examine the impact of a particular cause of death as well as differences over time, geography, and population groups. In the context of the opioid epidemic, calculations based on competing risks models suggest the impact of opioid overdoses on mortality has increased dramatically between 2000-2020 (Hedegaard et al. (2021)). However, such analyses may miss the total impact of the opioid epidemic in two ways. First, opioid overdoses account for only part of the opioid-associated mortality (Degenhardt et al. (2011); Ray et al. (2016)). Second, and relatedly, standard counterfactual analysis based on competing risks models imposes strong assumptions to create measures of mortality in the absence of opioid overdoses. Concisely, relying on measures of opioid overdoses alone ignores the interrelationship between opioid use and other causes of death.[4]

Preston and coauthors present an alternative approach to measure the total impact of a particular cause of death (Fenelon and Preston (2012); Preston et al. (2010); Glei and Preston (2020)). They use geographic variation in a coded cause of death to the effect of underlying behavioral causes: for example, lung cancer deaths act as a proxy for smoking

[4]Appendix B presents a more detailed discussion of competing risks models and related assumptions.

and drug overdose as a proxy for drug use. The geographic variation in these deaths is used to attribute additional, other-cause deaths to these underlying causes using regression analysis. In the context of the opioid epidemic, under the assumption that opioid usage is the sole explanation for risk of opioid overdose and that opioid usage impacts other-cause mortality risk as well, a proportional hazard model implies that there will be a log-linear relationship between other-cause mortality and opioid mortality, with the latter acting as a proxy for opioid usage. The details of this modelling background are available in Appendix B.

This model suggests an empirical approach that employs a generalized linear model with exponential link function. The empirical approach suggested by Preston and co-authors (Fenelon and Preston (2012); Preston et al. (2010); Glei and Preston (2020)) estimate a negative binomial regression of other-cause mortality on the mortality of the cause of interest—lung cancer in the former two cases or drug overdoses in the latter. However, as noted by Wooldridge (1999), the use of a negative binomial model in cases of over-dispersion is unnecessary when interest is focused on inference at the mean. In particular, the Poisson quasi-maximum likelihood estimator (QMLE) fixed-effects model requires only that the conditional mean of the model is specified correctly and is robust to distributional misspecification and arbitrary relationships between mean and variance, while the negative binomial imposes additional restrictions on the relationship between the conditional mean and conditional variance.[5] In recognition of these relative advantages of the Poisson model—given our context that does not make specific distributional assumptions—we employ a Poisson QMLE fixed effects model in our empirical analysis.[6]

The assumption that the conditional mean is correctly specified rules out any omitted variables.[7] In particular, there can be no alternative cause that drives both opioid usage and other cause mortality. Recognizing that recent work has highlighted plausible role of deteriorating economic conditions driving both opioid use and other-cause mortality, either through behavioral responses that increase likelihood of death by so-called "deaths

[5]Inference at the mean *would* be impacted if the standard errors are calculated using the Poisson assumption of equality of mean and variance. However, we calculate clustered standard errors that do not make use of this assumption.

[6]We have also examined select results employing an unconditional negative binomial fixed effect approach. The results do not change. Given the computational power required to estimate the unconditional negative binomial fixed effect model and the theoretical advantages of the Poisson QMLE approach, we report results from the Poisson QMLE model.

[7]With an exponential link function, an omitted variable that is independent of the explanatory variable of interest does not impact the estimation of the coefficient on the included explanatory variable. Hence, for our purposes, we only consider omitted variables that are related to both the outcome and the explanatory variable of interest.

of despair" (Case and Deaton (2015); Case and Deaton (2017); Case and Deaton (2020)) or by prolonged stress deteriorating other physiological functions (Geronimus et al. (2006); McEwen (1998); McEwen and Stellar (1993); Seeman et al. (2001)), we test this particular avenue of explanation by controlling for location-specific economic deterioration.

We control for this potential endogeneity via reweighting our model. Intuitively, we create a pseudo-population that accounts for the non-random geographic variation in economic deterioration that could be driving both opioid and other-cause mortality. Because in the observed population, those areas that have both high degrees of economic deterioration and opioid mortality are "overrepresented"—in the sense that the endogenous selection makes them more likely to be observed—we down weight these observations in our pseudo-population. Conversely, those observations with high degrees of economic deterioration but low opioid mortality are underrepresented—again, in the sense that the endogenous selection makes them less likely to be observed—we up-weight these observations. This reweighting approach is akin to a model where opioid mortality is a (proxy) measure of "treatment" status which is endogenously determined. The reweighting accounts for non-random selection into "treatment" status determined by observable factors relating to the economic conditions affecting each geographic unit.

Additionally, the model implies that the "spillover" effect of opioid use on other-cause mortality is constant across time.[8] Yet, the initial period of the opioid epidemic was largely driven by prescriptions opioids (Paulozzi et al. (2011)) while the latter period was marked by increases in heroin (Rudd et al. (2014)) and synthetic opioids, particularly illicitly produced fentanyl (Hedegaard et al. (2021)). It is plausible that the type of opioid has differential effects on other-cause mortality, if, for example, the higher potency of heroin or fentanyl results in significantly higher risk of overdose, precluding the physiological mechanisms that would drive opioid-related other-cause mortality.

To explore this potential source of heterogeneity, we split the sample into two distinct time periods: the first covering the "early" portion of the opioid epidemic which was largely characterized by an increase in prescription opioid use, from 1990-2009, and the second

[8]In linear models, not accounting for heterogeneity in the slope coefficients across two groups results in an estimated coefficient that is a weighted average of the group-specific slope coefficients, where the weights related to the variance of the covariates. Here, from the first order conditions, we can show that the resulting coefficient solves $\sum_i X_i(Y_i - \exp\{X_i\hat{\beta}\}) = 0$ if there are different true parameters for different groups, this is equivalent to solving $\sum_A X_A(\exp\{X_A\beta_A\} + \varepsilon_A - \exp\{X_A\hat{\beta}\}) + \sum_B X_B(\exp\{X_B\beta_B\} + \varepsilon_B - \exp\{X_B\hat{\beta}\}) = 0$. If the model is otherwise correctly specified, it is clear that the estimated $\hat{\beta}$ will fall between these two true beta parameters, though therere is no closed-form solution for the corresponding weights.

covering the “later” portion of the opioid epidemic characterized by the increase in illicitly produced opioid use, from 2009-2017.

We use the methodology of Preston et al. (2010) and Glei and Preston (2020) as the basis for our empirical analysis. We examine the effect of opioid mortality on excess other-cause mortality using a fixed-effect quasi-MLE Poisson regression model:

$$M^{non-opioid}_{CZ,age,year} = \exp\{\sum_{age \in A} \beta_{age} M^{OD}_{CZ,age,year} + \phi_{year,age} + \phi_{CZ,age} + \varepsilon_{CZ,age,year}\} \tag{3.4.1}$$

Where $M^{non-opioid}_{CZ,age,year}$ is the non-opioid mortality rate in commuting zone CZ for age bin $age \in \{[25-29], \ldots, [60-64]\}$ in year $year \in \{1990, 2009\}$ for our first long-difference and $year \in \{2009, 2017\}$ for our second long-difference. Similarly, the dependent variable $M^{OD}_{CZ,age,year}$ is opioid overdose mortality for the same CZ X age bin X year observation. We allow the coefficients on each opioid mortality explanatory variable to differ by age-bin. All analyses are done separately by sex. We include fixed effects for commuting zones and year fully saturated by age-bins. The inclusion of geographic and time fixed effects means our identifying variation comes from differences across geographies over time—that is, we use the uneven geographic development of the opioid crisis over time to identify other-cause mortality attributable to opioid use. It also means that baseline differences across geographies or common trends in mortality across time will not impact our identification. However, any potential endogeneity that may impact both opioid use and other-cause mortality that varies over geographies across time could impact identification—e.g. the uneven economic deterioration over the same time period.

Our first set of specifications includes no additional controls. In our Results Section 3.5, we call the results from this estimation approach Model 1—it uses coefficients from estimation of Equation 3.4.1 with no re-weighting except for population weights. In the alternative specifications that control for endogenous selection by reweighting, we calculate the generalized propensity score following Hirano and Imbens (2004). Rather than using these propensity scores in matching estimators, we use the estimated propensity scores to calculate Inverse Probability Weights (IPW) and estimate a weighted version of our nonlinear specification above, akin to the method suggested by Robins (1999) and Robins et al. (2000). We call the results of this estimation approach Model 2.

In particular, we estimate the conditional distribution of the change in opioid overdose mortality between 1990 and 2009 for our first set of specifications and between 2009 and

2017 in our second set of specifications. The estimated distribution is conditional on the changes in the poverty rate, government assistance rate, non-college rate, non-employed rate, work disability rate, and divorce/separation rate. The estimation assumes that the change in opioid mortality conditional on these explanatory variables:

$$\Delta M_{CZ}^{OD}|X \sim N(h(\gamma, X_{CZ}), \sigma^2)$$

We use the *gpscore* estimation package in Stata (Bia and Mattei (2008)), which estimates the parameter vector (γ, σ^2) by MLE and then recovers the estimated propensity scores assuming a conditional normal distribution:

$$\hat{R} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{\frac{-1}{2\sigma^2}(\Delta M_{CZ}^{OD} - h(\hat{\gamma}, X_{CZ}))^2\} \tag{3.4.2}$$

We use the inverse of these estimated propensity scores to weight our nonlinear model, 3.4.1. A concern with weights based on a continuous exposure variable, such as is the case here, is that if there is strong predictive power between (some) of the X_CZ variables and the continuous exposure, M_{CZ}^{opioid}, then this estimation procedure can produce extremely large weights for a few observations with larger values of $(\Delta M_{CZ}^{OD} - h(\hat{\gamma}, X_{CZ}))^2$. This can result in estimates from a weighted model that fail to achieve asymptotic normality. Following Robins (1999) and Robins et al. (2000), we stabilize these inverse propensity weights (IPW) using the unconditional probability density. That is, we also estimate:

$$\hat{s} = \frac{1}{\sqrt{2\pi\sigma^2_{\Delta M_{CZ}^{OD}}}} \exp\{\frac{-1}{2\sigma^2_{\Delta M_{CZ}^{OD}}}(\Delta M_{CZ}^{OD} - \mu_{\Delta M_{CZ}^{OD}})^2\}$$

Where $\sigma^2_{\Delta M_{CZ}^{OD}}$ and $\mu_{\Delta M_{CZ}^{OD}}$ are the sample standard deviation and mean of the exposure variable, ΔM_{CZ}^{OD}, respectively. Our stabilized weights are defined as:

$$w_{CZ} = \frac{\hat{s_{CZ}}}{\hat{R_{CZ}}}$$

Robins (1999) shows that weighted estimates using these stabilized weights recover consistent estimates. Our functional form for estimating the conditional distribution for ΔM_{CZ}^{OD}, $N(h(\hat{\gamma}, X_{CZ}), \hat{\sigma^2})$, is linear in the explanatory variables capturing economic conditions by commuting zone drawn from the Census and ACS.[9] Balance tests for our weighing approach

[9]Alternative functional forms include (i) quadratics in each of these explanatory variables and (ii) quadratics and interactions between the five strongest explanatory variables selected by a machine learning algorithm. Furthermore, alternative calculations of the weights do not include the stabilization term, $s\hat{_{CZ}}$, and instead trim the smallest and largest 1% of the un-stabilized weights from the sample. Finally, alternative

are reported in the Appendix B. All specifications are weighted by the average share of population in the commuting zone over the years of our analysis: 1990, 2009, and 2017.[10] Standard errors are clustered at the commuting zone X age level.

The long-difference quasi-MLE Poisson model can be re-written as:

$$\%\Delta M_{CZ,age}^{non-opioid} = \exp\{\sum_{age \in A} \Delta M_{CZ,age}^{OD} + \Delta\phi_{age} + \Delta\varepsilon_{CZ,age}\}$$

With $\%\Delta M_{CZ,age}^{non-opioid} = \frac{M_{CZ,age,\text{late year}}^{non-opioid}}{M_{CZ,age,\text{base year}}^{non-opioid}}$ and $\Delta M_{CZ,age}^{OD} = M_{CZ,age,\text{late year}}^{OD} - M_{CZ,age,\text{base year}}^{OD}$.

Thus, the exponent of the coefficients of interest, β_{age}, are interpreted as the percent increase in non-opioid coded mortality associated with a unit increase in opioid overdose mortality. To aid in interpretation, we standardize mortality rates to be per 1,000 population.

We then calculate the share of other-cause mortality that is attributable to opioid use using the relationship:

$$ShM_{non-opioid}^{opioid} = (1 - \exp\{-\beta_{age}^{OD} M_{CZ,age,year}^{OD}\})$$

Which implies the total mortality attributable to opioid use is:

$$M_{CZ,age,year}^{\text{total opioid}} = M_{CZ,age,year}^{OD} + M_{CZ,age,year}^{non-opioid}(1 - \exp\{-\beta_{age}^{OD} M_{CZ,age,year}^{OD}\}) \tag{3.4.3}$$

We then calculate summary measures of the overall effect of opioid use on mortality. Following Andersen et al. (2013), we calculate the opioid-coded Years of Life Lost (YLL) as well as the YLL to opioids, both directly (i.e. deaths coded as opioid overdoses) and indirectly (other-cause deaths that are caused by opioids, as implied by our statistical analysis). This measure accounts for how many potential years between ages 25 and 85 are lost due to an opioid death. That is, a death at age 25 due to opioids contributes 60 lost years to the YLL

calculations of the weights impose a Box-Cox transformation on the exposure term, ΔM_{CZ}^{OD}, to better approximate normality. These alternative specifications yield qualitatively similar results.

[10] Standard Poisson models with counts as the dependent variable and including population as the exposure variable are equivalent to models quasi-MLE Poisson models with rates as the dependent variable and populations as weights. Here, we use the average share of population in a commuting zone over our sample timeframe. This is equivalent to a Poisson model with standardized counts as the dependent variable, where standardization accounts for population shifts, and a constant population as the exposure variable. This prevents commuting zones with large population changes from having differential weights, allowing for the long-difference interpretation of the model. Specifications that use actual population rather than average population share give similar results.

measure whereas a death at age 45 due to opioids contributes 40 lost years to the aggregate YLL measure. When we calculate the YLL to opioids both directly and indirectly, we assume that the opioid-coded mortality for ages 65+ is an accurate measure of opioid-caused mortality.[11] That is, we only adjust opioid mortality rates for ages 25-64 and use the opioid-coded mortality for ages 65+. Notably, this measure places more weight on an early death than a later death. Because the opioid-caused YLL plus any other-cause YLL will add up to the all-cause YLL—the total years of potential life lost to any death—this summary measure allows us to examine the total impact of opioids on observed aggregate mortality. Hence, this summary measure shows the total impact of opioid-coded mortality as well as the all opioid-caused mortality after accounting for the impact of opioids on other-cause mortality using our statistical approach. Comparisons of the difference between opioid-coded YLL and all opioid-caused YLL highlight the additional impact of opioids on mortality not captured by opioid-coded mortality alone. Comparisons between the all opioid-caused YLL as measured by our models that do not account for non-random geographic variation in opioid mortality versus those that do highlight the importance of accounting for such endogeneity in this statistical approach. More details for this accounting approach to mortality are in Appendix B.

Additionally, we report the cause-elimination life-expectancy at age 25 after eliminating all opioid-related deaths, both those caused directly by opioid use (opioid-coded mortality) as well as those indirectly caused by opioid use (implied by our statistical models). Standard cause-elimination approaches generate a counterfactual life-table that substitutes a measure of all-cause mortality that removes opioid-caused mortality risk, $\tilde{M} = M_{observed} - M_{opioid}$.[12] We then calculate the life expectancy for an individual living with this counterfactual mortality risk.

We calculate three measures: one which removes only mortality caused directly by opioid overdoses (opioid coded mortality) and two which remove both mortality directly caused by opioid overdose and mortality indirectly caused by opioids as implied by our statistical models. When we estimate all opioid-caused mortality risk, we assume that the opioid-coded mortality rates for ages 65+ are accurate measures of all opioid-caused mortality for these ages. In contrast to our YLL measures, which essentially assign some additional deaths as caused by opioid use but do not change total observed mortality, the life-expectancy approach uses a counterfactual mortality rate, assuming that some individuals would not

[11]The opioid-coded mortality risk as well as measured additional impact of opioid use on other-cause mortality risk are very low for these age bins.

[12]Alternative methods to remove opioid-caused mortality risk developed by Chiang (1968) yield similar measures. See Appendix B for details.

have died at the ages they did in the absence of opioid use. More details are available in Appendix B. After accounting for the impact of opioid use on other-cause mortality using our statistical approach, under the assumption that the remaining other-cause mortality is independent of the total opioid-related mortality, standard approaches capture the counterfactual life expectancy in the absence of opioid mortality.

Analyses are done in Stata/SE version 16.1. The quasi-MLE Poisson regressions use the *ppmlhdfe* code by Correia et al. (2019).

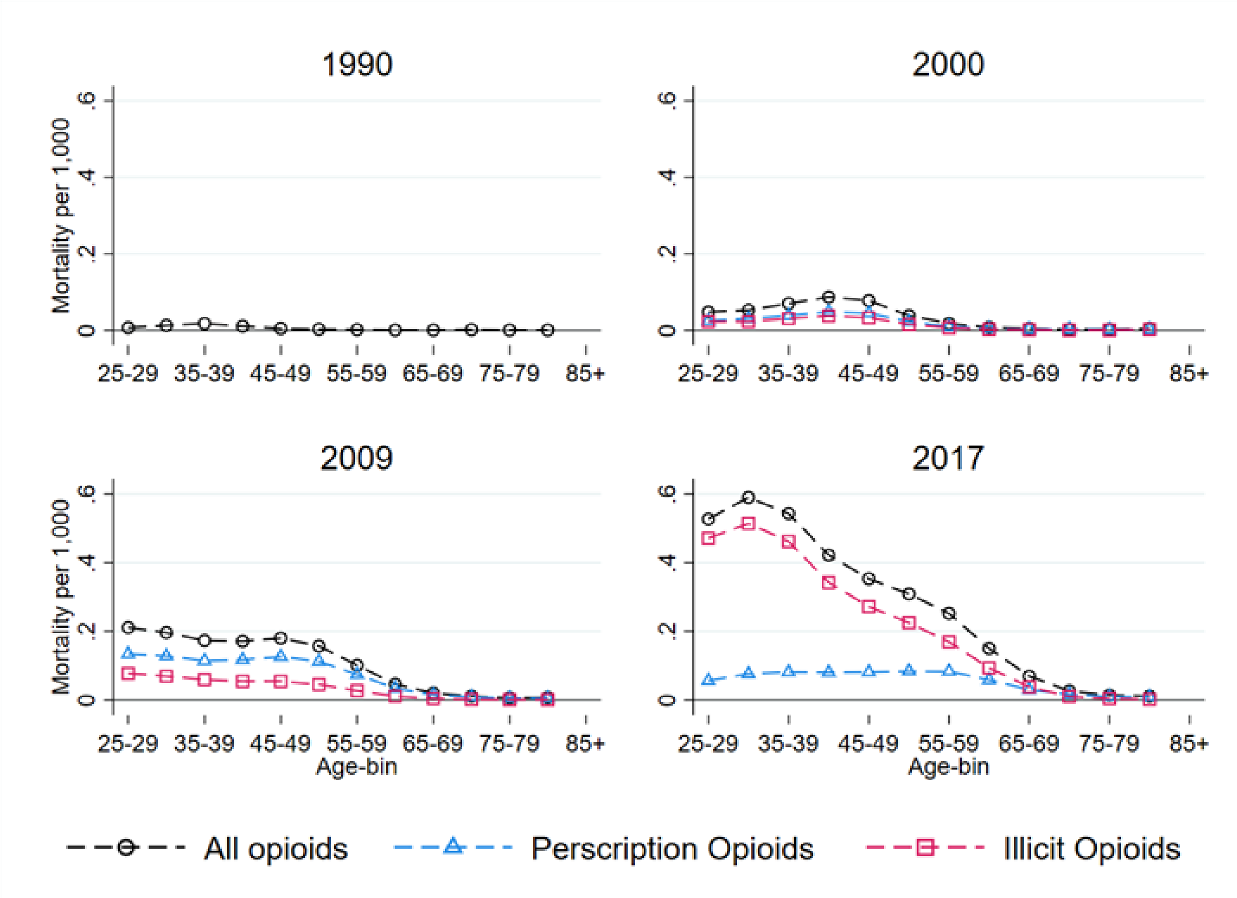
3.5 Results

Figure 3.2 reports the age-specific opioid overdose mortality among non-Hispanic white men for select years by opioid type. Opioid overdoses were relatively rare in 1990 but rose steadily through the late 1990s and 2000s. As previously noted, the early rise in opioid overdoses was largely driven by prescription opioids, with illicit opioid overdoses driving the increase over the 2010s. Figure 3.3 reports the same statistics for non-Hispanic white women. The general patterns of growth are similar for women compared to men, but the opioid mortality among women is significantly less than that of men. However, this growth in opioid mortality was not geographically uniform. Figure 3.4 display the growth in opioid mortality by commuting zone over time. The initial growth in opioid mortality from 1990-2009 was concentrated in particular geographic areas, especially Appalachia, Oklahoma, and Nevada and Northern California. By 2017, opioid mortality had increased in urban areas in the mid-West, New England, and the Mid-Atlantic. Figure 3.5 shows the geographic variation in non-opioid mortality—i.e. mortality by any cause other than opioid-poisoning—over time. Note that the areas that experienced growth in opioid mortality also had elevated non-opioid mortality rates, suggesting a relationship between the two. However, those areas also had relatively elevated mortality rates *prior* to the increase in opioid mortality, suggesting the potential for other mechanisms driving these correlations. Furthermore, the geographies that experienced growth in opioid mortality from 2009-2017 did not uniformly experience growth in other-cause mortality as well.

Figure 3.6 reports our main results for non-Hispanic white men. After estimating the standard and IPW versions of Equation 3.4.1, we calculate the opioid-related mortality rates implied by the model estimates using Equation 3.4.3.[13] The top panel reports the results from 1990-2009 and the bottom panel reports results from 2009-2017. Focusing first on

[13] To be clear: all specifications are weighted by the average share of population in a particular commuting zone. The IPW specifications are additionally weighted by the inverse of the propensity score given in Equation 3.4.2 stabilized as described in Section 3.4.

Figure 3.2: Opioid-overdose mortality, All Non-Hispanic White Men

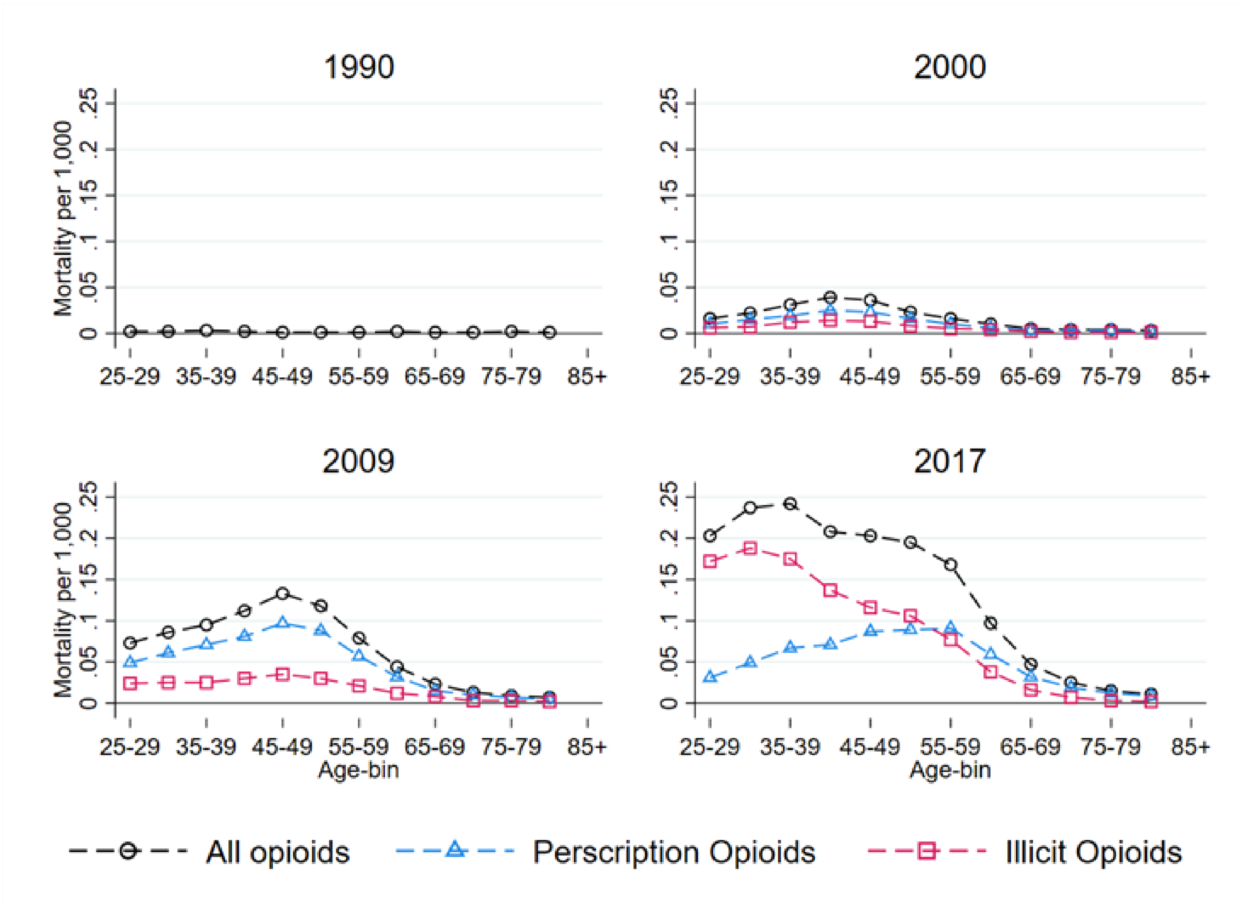


Note: Data on opioid-type not available in 1990

the standard results from Model 1—that is, estimates that assume the geographic variation in opioid mortality is exogenous—we find evidence of significant "spillover" effects for men in the early period, 1990-2009. However, turning to Model 2, these estimated spillovers diminish significantly after accounting for the non-random geographic variation in opioid mortality via our reweighting approach. These results suggest that accounting for the nonrandom variation in opioid mortality—specifically that which is driven by economic deterioration in the commuting zones—mitigates the spillover effects of opioid usage on other-cause mortality.

The bottom panel reports the results for the 2009-2017 long difference. Here, Model 1 suggests that there is very little spillover effect of opioid use. Furthermore, our reweighting approach results in an estimated impact of opioid use on mortality that is not significantly different than the opioid-coded mortality, in general. Full model results are available in

Figure 3.3: Opioid-overdose mortality, All Non-Hispanic White Women



Note: Data on opioid-type not available in 1990

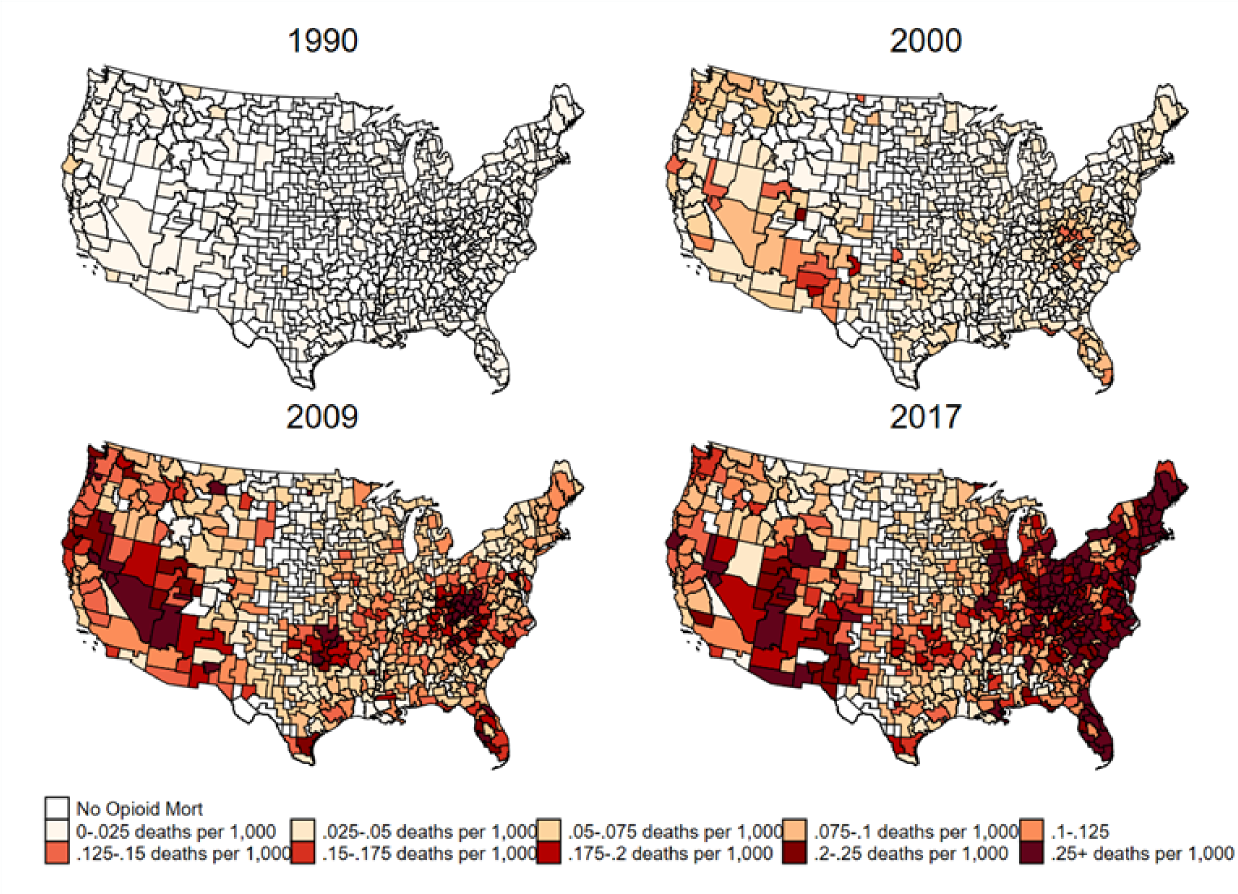
Appendix B.

Several of these findings suggest the need for nuanced interpretation. First, Figure 3.6 shows that spillover effects between 1990 and 2009 estimated using Model 1 are statistically significant and of similar magnitude to prior estimates (Glei and Preston (2020)),[14] but there is virtually no evidence for spillover effects between 2009 and 2017, suggesting temporal heterogeneity in the relationship between opioid usage and other-cause mortality.

Examining the result from Model 2, however, we find that accounting for the non-random

[14]Modelling differences between Glei and Preston (2020) and our approach account for some difference in the magnitude of coefficients, but the qualitative results remain. These differences are outlined in Appendix B as well as a comparison of approaches.

Figure 3.4: Geographic variation in opioid-coded mortality over time

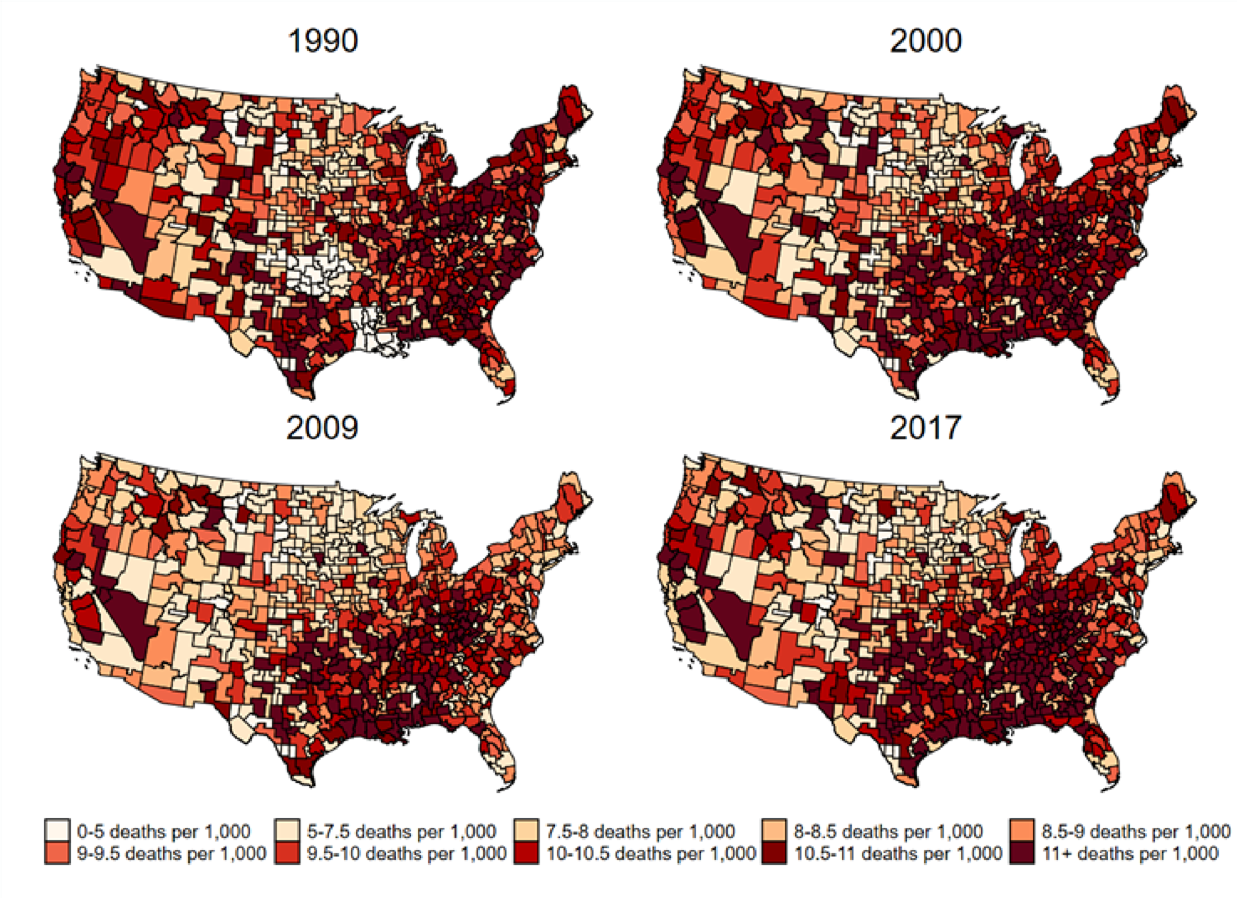


variation in opioid mortality in the early period, 1990-2009, mitigates the estimated spillover effects. This suggests that the use of these nonlinear methods to estimate total impact of opioid use on mortality should carefully consider alternative explanations for the geographic variation in opioid mortality; specifically, that there may be an endogenous relationship between opioid use and other-cause mortality driven by broader economic malaise (Case and Deaton (2020)) or prolonged exposure to particular physiological stressors (Geronimus et al. (2006); McEwen (1998); McEwen and Stellar (1993); Seeman et al. (2001)).

Figure 3.7 reports the corresponding results for white women. Again, using the unadjusted Model 1, we find evidence that there are significant impacts of opioid use on other-cause mortality for the period 1990-2009. For the period 2009-2017, we find some evidence for these "spillover" effects. Model 2 again tends to diminish the estimated effects.

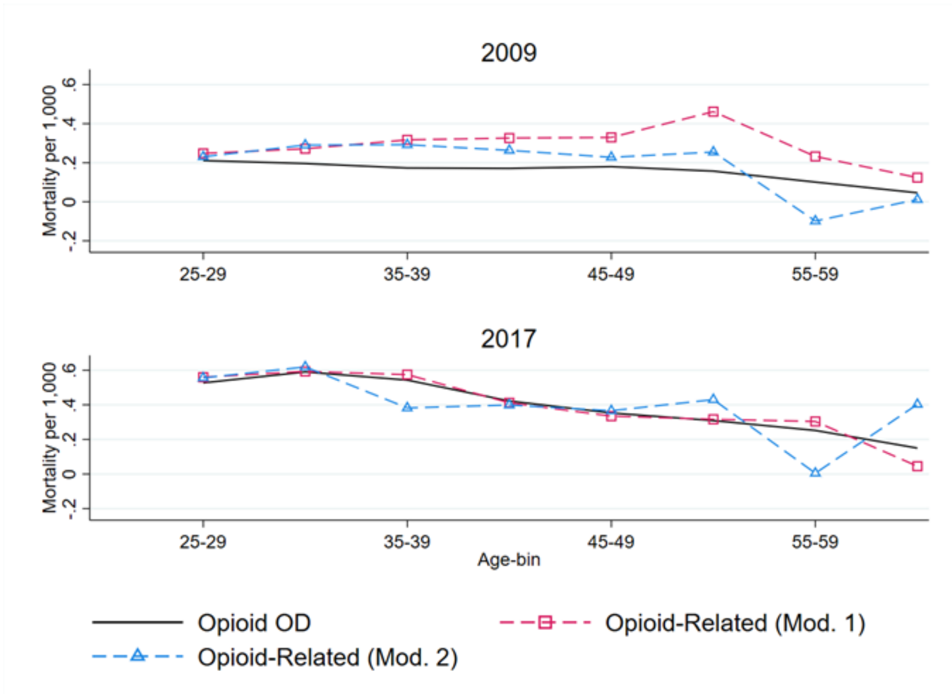
In order to estimate the total impact of opioids on mortality, Table 3.1 compiles these spillover effects into summary measures on total mortality for men and women. Panel A

Figure 3.5: Geographic variation in all other-cause mortality over time



reports the accounting approach by calculating the YLL due to opioid overdose in 2009 and 2017. In particular, the first column reports the actual, overall YLL due to any cause. The second column reports the YLL due to opioid-coded overdose. The third and fourth columns report the implied YLL due to over-doses after assigning a share of other-cause mortality to opioid mortality using the approach out-lined in Equation 3.4.3 using the estimates from Model 1 and 2, respectively. As the YLL to opioid-related mortality is a linear construction of the estimates from Models 1 and 2, we report the associated standard errors using the Delta method. Column 5 reports the difference between column 3 and 2, testing the hypothesis that the YLL to opioid related mortality according to Model 1 is the same as the actual YLL due to opioid-coded overdoses. Similarly, column 6 reports the difference between column 4 and 2, testing the hypothesis that the YLL to opioid related mortality from Model 2 is the same as opioid-coded YLL. Finally, column 7 reports the difference between column 4 and 3, testing the hypothesis that the weighted and unweighted estimates result in the same YLL for opioid-related mortality. Column 7 is essentially the results of a Hausman test, where the null hypothesis is that the unweighted estimates are efficient; however, as our test is

Figure 3.6: Observed Opioid Overdose Mortality and Model-implied Opioid-related Mortality, Non-Hispanic White Men



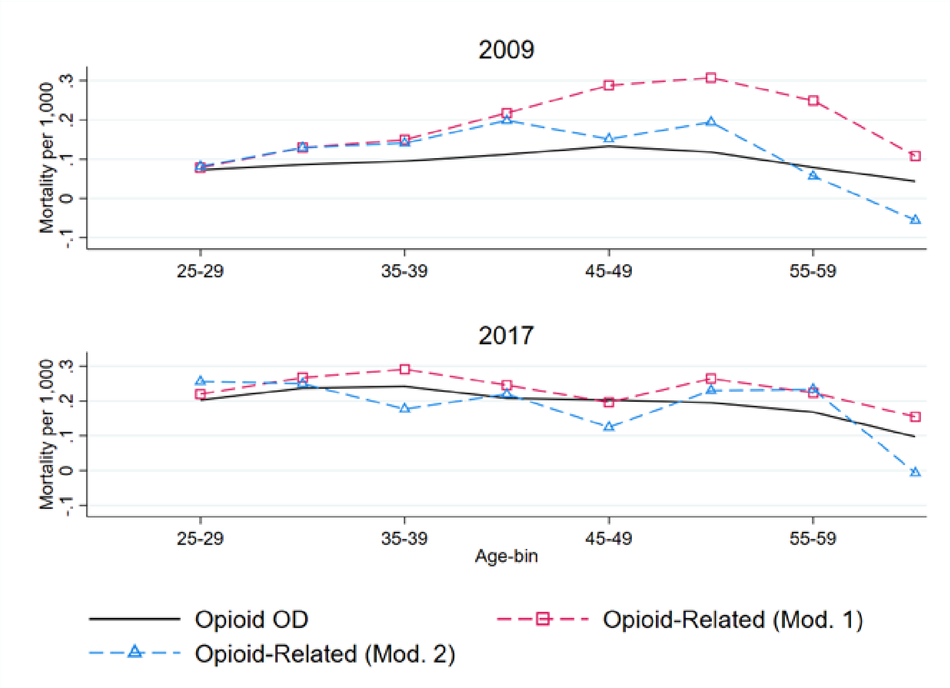
Note: The opioid related estimates use the method presented in Equation 3.4.3. Model 1 uses coefficients from Equation 3.4.1 with no controls for economic conditions. Model 2 uses coefficients from Equation 3.4.1 and re-weights the specification by the stabilized inverse propensity score, where propensity scores are calculated using change in opioid mortality as the outcome and changes in economic characteristics as explanatory variables. See Equation 3.4.2 for definition of these propensity scores.

on a scalar rather than a vector, we report the difference rather than the more familiar F-test.

The top panel of Table 3.1 shows that for all men, the share of YLL due to opioids in 2009 increases from 2.7% to 4.8% after accounting for spillover deaths attributed to opioids using Model 1, but only 3.7% using Model 2, which controls for non-random variation in opioid mortality. The hypothesis tests reported in columns 5-7 show that accounting for the nonrandom variation in opioid mortality significantly decreases the estimated YLL by 0.111 years, though there still remains a significant aggregate impact of opioid spillovers. For 2017, while there are more YLL due to opioids, the impact of spillovers on the mortality differential is more muted. While accounting for nonrandom selection into opioid mortality levels decreases the estimated spillover effect somewhat, this decrease is not significant and we cannot rule out null spillover effects on the aggregate YLL measure.

Turning to the corresponding results for women, we see that opioids have a smaller

Figure 3.7: Observed Opioid Overdose Mortality and Model-implied Opioid-related Mortality, Non-Hispanic White Women



Note: The opioid related estimates use the method presented in Equation 3.4.3. Model 1 uses coefficients from Equation 3.4.1 with no controls for economic conditions. Model 2 uses coefficients from Equation 3.4.1 and re-weights the specification by the stabilized inverse propensity score, where propensity scores are calculated using change in opioid mortality as the outcome and changes in economic characteristics as explanatory variables. See Equation 3.4.2 for definition of these propensity scores.

impact on women's mortality, with opioid-coded deaths representing 2.2% of all YLL. Accounting for opioids' impact on other-cause death using Model 1 nearly doubles this impact. However, the Model 2 results that account for non-random selection significantly reduces estimated impacts, resulting in a total impact that is 37% larger than opioid-coded deaths alone.[15]

We also calculated the effects of adjustment on estimates of life expectancy using standard cause-elimination techniques. Table 3.1 Panel B reports the life-expectancy at age 25, as well as the counterfactual life-expectancy at age 25 after eliminating opioid overdose as a risk factor. Specifically, the first column reports actual life-expectancy, the second reports counterfactual life-expectancy after eliminating deaths coded as opioid poisoning. The third

[15]Note that the Hausman test here cannot estimate a standard error for the scalar difference between the model implied estimates, because the weighted estimate results in slightly tighter standard errors, a violation of the null hypothesis that both estimates are consistent. If both estimates were consistent, then the unweighted estimate would also be efficient.

and fourth columns eliminate all opioid-related deaths as outlined in Equation 3.4.3, again using standard life-table methods. Columns 5-7 again report the hypothesis testing of interest: whether accounting for opioid spillovers significantly changes the estimated impact of opioid cause-elimination on life expectance (columns 5 and 6) and whether weighting to account for nonrandom selection into opioid mortality levels significantly im-pacts the estimation of these spillover effects (column 7). Standard errors are again estimated using the Delta method, as the life expectancy after eliminating a cause is a (nonlinear) combination of estimates from the model. Column 7 is again essentially the scalar version of a Hausman test between Model 1 and 2.

Eliminating only deaths coded as opioid poisoning increases life expectancy for all white men by .23 years in 2009, a 0.43% increase. After accounting for spillover effects, eliminating both opioid and opioid-spillover deaths increases life expectancy by 0.40 years using Model 1 and 0.30 years using Model 2, a 0.76% or 0.34% increase, respectively. Turning to the hypothesis tests in columns 5-7, we again see that accounting for opioids' impact on other-cause mortality does significantly increase life expectancy relative to only opioid-coded deaths, but the weighting approach—i.e. accounting for non-random selection into opioid mortality levels—reduces this impact. In 2017, life expectancy for all white men increased by 0.60 years (1.1%) after eliminating opioid-coded deaths, and the removing all opioid-caused deaths implied by our statistical approach did not impact this change, regardless of whether we account for nonrandom geographic variation in the growth of opioid use.

For women, eliminating deaths coded as opioid poisoning increased life expectancy at age 25 0.15 years (0.25% increase) in 2009, but accounting for spillovers as well increased life expectancy 0.28 years (0.49%) using Model 1 and 0.20 years (0.34%) using Model 2. Again, the hypothesis tests in columns 5-7 show a similar pattern as above: accounting for effects of opioids on other-cause mortality significantly increases the impact on life expectancy, but this impact is significantly mitigated (though not eliminated) when we also account for the non-random variation in opioid mortality.[16] In 2017, the elimination of opioid-coded deaths added 0.31 years (0.55%) while accounting for spillovers as well increased life expectancy by 0.37 years (0.65%) using Model 1 and 0.31 years (0.54%) using Model 2, with none of these changes being significant.

[16]Again, because our weighted model results in tighter standard errors, the Hausman-style test of difference between model estimates cannot estimate a standard error, as the null hypothesis that the unweighted estimate is consistent and efficient is necessarily violated.

Overall, the results of these exercises highlight that opioid's impact on non-opioid mortality—i.e. the "spillover" effect of opioids—was significant in 1990-2009, but the magnitude of this impact was significantly dampened by accounting for non-random geographic variation in opioid mortality related to the socioeconomic decline of particular geographic areas over this time period. Furthermore, we find no evidence of a similar effects in the 2009-2017 period.

One explanation for the apparent shift in the spillover effects between the early period of the opioid epidemic and the later period is the nature of opioids themselves—as noted in Figures 3.2 and 3.3, the early period of opioid mortality was largely driven by the use of prescription opioids, while the later period was largely driven by illicit opioid use. It is possible that these different types of opioids haver different physiological effects, e.g. the use of illicit opioids may result in higher probability of overdose death, precluding the opioid-associated death captured by the spillover effect measures.

Appendix Figure B.1 reports the results of this exercise for white men overall. We find that the spillover effects identified in for the earlier period using Model 1, 1990-2009, are entirely explained by prescription opioid mortality. In the second period, where we find no evidence of spillover effects using Model 1 but some evidence of spillover effects at older ages using Model 2—we see that there is little evidence of spillover effects for either opioid type using Model 1 but evidence that the measured spillovers in older age-bins using Model 2 are driven by prescription opioids.

Appendix Figure B.2 report the corresponding results for white women. Similar to men, we find that the spillover effects identified by Model 1 for women in the first period, 1990-2009, are explained by prescription opioid mortality. The results for Model 2 suggest that the spillover effect of prescription opioid mortality is more muted in the early period.[17]

[17]In the later period, we estimate large negative total mortality impact of prescription opioids. However, the weights are defined on overall opioid mortality as the exposure variable. Hence, we exercise caution in interpreting the estimates by opioid-type when we use the weighted sample. See Appendix B for full results.

Table 3.1: Summary Measures of Opioid Impact on Mortality for Non-Hispanic White Men and Women in 2009 and 2017

A. *Years of Life Lost (YLL) between ages 25–84 due to opioid use*

	Total YLL	YLL to Opioid Overdose	Total YLL to Opioids (Model 1)	Total YLL to Opioids (Model 2)	Model 1 - Opioid OD	Model 2 - Opioid OD	Model 2 -Model1
			Non-Hispanic White Men				
2009	9.714	0.265	0.466	0.355	0.201	0.090	-0.111
(s.e.)			(0.039)	(0.040)	(0.039)	(0.040)	(0.009)
2017	9.946	0.684	0.694	0.671	0.010	-0.013	-0.023
(s.e.)			(0.019)	(0.239)	(0.019)	(0.239)	(0.238)
			Non-Hispanic White Women				
2009	6.668	0.150	0.290	0.205	0.140	0.055	-0.085
(s.e.)			(0.028)	(0.022)	(0.028)	(0.022)	(NA)
2017	6.764	0.327	0.385	0.323	0.058	-0.004	-0.062
(s.e.)			(0.013)	(0.394)	(0.013)	(0.394)	(0.394)

B. *Actual Life Expectancy at age 25 and Counterfactual Life Expectancy after eliminating opioid risk*

	Actual LE	Eliminate Opioid Overdose	Eliminate total opioid mort. (Model 1)	Eliminate total opioid mort. (Model 2)	Model 1 - Opioid OD	Model 2 - Opioid OD	Model 2 -Model1
			Non-Hispanic White Men				
2009	52.634	52.862	53.035	52.939	0.173	0.077	-0.096
(s.e.)			(0.034)	(0.035)	(0.034)	(0.035)	(0.008)
2017	52.429	53.026	53.035	53.016	0.009	-0.010	-0.019
(s.e.)			(0.017)	(0.186)	(0.017)	(0.186)	(0.185)
			Non-Hispanic White Women				
2009	57.271	57.416	57.550	57.467	0.134	0.051	-0.083
(s.e.)			(0.028)	(0.022)	(0.028)	(0.022)	(NA)
2017	56.908	57.221	57.276	57.216	0.055	-0.005	-0.060
(s.e.)			(0.012)	(0.360)	(0.012)	(0.360)	(0.360)

Estimates calculate total mortality attributable to opioids for each age-bin using Equatin 3.4.3 and aggregate to a summary measure, YLL and LE at age 25. Standard errors calculated using Delta method. Column 7 (Model 2-Model 1) reports a scalar-version of a Hausman test, under the null hypothesis that the unweighted approach (Model 1) is efficient)

3.6 Conclusion

Any analysis of the contribution of any specific cause of death to overall mortality or to life expectancy relies on assumptions about competing risks. Typically, and largely out of necessity, authors assume competing risks are independent of each other. This is an exceedingly strong and not particularly plausible assumption. Preston and his colleagues have proposed an alternative to this approach. Under the assumption that the geographic variation in a particular cause of death is exogenous to other causes, they develop a technique that allows them to identify path interdependence. Within the context of the opioid epidemic, Glei and Preston (2020) find evidence that many causes of death rose in the same places that opioid overdoes rose, and their estimates imply that the impact of opioid use on overall life expectancy was roughly twice the direct effect.

In this paper we have examined two assumptions implicit in the Glei and Preston calculations. First, given existing literature, it seemed likely the stagnating local economic conditions would have an impact on both chronic disease mortality and opioid overdoes mortality. This would then represent a correlated omitted variable which would bias the Glei and Preston estimates upward. Secondly, we thought it plausible that the magnitude of any spillover effects would vary over time, based in part on the heterogeneity in the specific drugs involved in overdose deaths.

Indeed, when we control for deteriorating economic conditions using re-weighting techniques pioneered by Robins (1999) and Robins et al. (2000), the estimated impact of opioid use on non-opioid coded mortality decreased significantly. Whereas summary measures of the total impact of opioid use on mortality that did not account for nonrandom geographic exposure to the opioid crisis were 76% larger than opioid coded deaths alone for white non-Hispanic men and 93% larger for white non-Hispanic women over 1990-2009—estimates that are broadly consistent with those of Glei and Preston—our reweighted estimates that account for economic deterioration decrease this estimate by over half, with reweighted estimates only 34% and 37% larger than opioid-coded mortality alone, respectively. Furthermore, we find little evidence of any meaningful impact of opioid use on total mortality beyond opioid-coded mortality over the period 2009-2017.

While we still find estimates that the total impact of opioid use on mortality in the United States is significantly understated by opioid-coded deaths alone, failing to account for endogenous selection into high opioid and high other-cause mortality geographic regions significantly overstates these impacts, potentially resulting in estimates upwards of twice as large relative to estimates that control for such selection. Several high-profile hypotheses have argued for a relationship between economic deterioration and the associated declines in access to medical care, increases in biophysiological stress and coping, increased mortality differentials, or "deaths of despair" Finally, we find that there is significant heterogeneity in these impacts over time, roughly coinciding with two "waves" in the opioid epidemic (Centers for Disease Control and Prevention (2022b)).

We believe that the methodology that has been developed by Preston and his colleagues is potentially useful; however, we would also argue that researchers using their methodology should do so with some caution, checking for both population heterogeneity and possible correlated omitted factors.

APPENDICES

APPENDIX A

Appendix to Chapter 1

A.1 Additional Policy Detail

The Higher Education Act of 1965 created the "Guaranteed Loans Program," which would eventually become the Federal Family Education Loans (FFEL) Program. Under this program, private banks would finance student loans, which were in turn guaranteed by the Federal government. In 1992, the amendments to the Higher Education Act allowed for the issuance of Direct Loans, which were financed directly by the Federal government. The two programs had substantively similar terms–i.e. the same means-tested Subsidized loans and more generally available Unsubsidized loans, the same limits on annual and lifetime borrowing, and the same interest rates–with higher education institutions selecting which program they participated in.

A crucial difference in treatment of FFEL loans and Direct Loans was eligibility for Income Driven Repayment. Under the ICR plan introduced in 1994, *only* Direct Loans were eligible. Throughout the late 1990s and early 2000s, the FFEL program dominated the federal student loans program, with total outstanding dollars in the FFEL program standing at over 2.5x the amount outstanding in the Direct Loans program (Department of Education (2002), National Student Loan Data System (NSLDS) (2022)), limiting the effectiveness of the ICR plan. IBR, introduced in 2009, explicitly allowed for either FFEL loans or Direct Loans. The FFEL program was ended in 2010, though existing FFEL loans continued to be serviced.

Table A.1 reports the federal loan limits for both dependent and independent students.

Table A.1: Borrowing Limits on Federal Student Loans

Academic Year	First Year	Year of Study Second Year	Third Year and above	Aggregate Limit
		Panel A. Dependent Students		
2006-2007	\$2,625	\$3,500	\$5,500	\$23,000
2007-2008	\$3,500	\$4,500	\$5,500	\$23,000
2008-2019	\$5,500 no more than \$3,500 sub.	\$6,500 no more than \$4,500 sub.	\$7,500 no more than \$5,500 sub.	\$31,000 no more than \$23,000 sub.
		Panel B. Independent Students		
2006-2007	\$6,625 no more than \$2,625 sub.	\$7,500 no more than \$3,500 sub.	\$10,500 no more than \$5,500 sub.	\$46,000 no more than \$23,000 sub.
2007-2008	\$7,500 no more than \$3,500 sub.	\$8,500 no more than \$4,500 sub.	\$10,500 no more than \$5,500 sub.	\$46,000 no more than \$23,000 sub.
2008-2019	\$9,500 no more than \$3,500 sub.	\$10,500 no more than \$4,500 sub.	\$12,500 no more than \$5,500 sub.	\$57,500 no more than \$23,000 sub.

For 2006-2008, the total limit for Federal loans, inclusive of subsidized and unsubsidized loans, coincided with the limit on subsidized loans alone for dependent students. Beginning in 2008, the total limit for loans was increased, but the cap for subsidized loans remained at 2007-2008 levels. For all years, the limit for subsidized loans did not differ between dependent and independent students, though independent students had larger total loan maximums. Source, CFR (2021) and Hegji (2021).

A.2 Model Detail

A.2.1 Model Preliminaries

Agents in the model are indexed by $i \in I$ and start with a $(1 + J) \times 1$ vector of abilities A_i with $A_{i,j} + \xi_{i,j}$ being the j^{th} entry. Each ability, $A_{i,j} + \xi_{i,j}$ is associated with a particular major, indexed $j \in J$. Additionally, while $A_{i,j}$ is known, $\xi_{i,j}$ is not, with $E(\xi_{i,j}) = 0$. Agents also begin with wealth, $y_{i,0} \in \mathbb{R}_+$.

In the first period, agents select a major, $j \in J$, including an outside option, $j = 0$. Define the variable $d_{i,j,1} = 1$ if individual i selects major j and $d_{i,j,1} = 0$ otherwise. If they select major $j \neq 0$, they pay some set pecuniary cost, which may depend on their initial wealth, $x(y_{i,0})$. This is similar to the "expected family contribution" or net cost of school in reality. For my purposes, I do not consider variation in costs for different programs of study.[1] In this first period agents receive flow utility $u(c_s)$ and pay non-pecuniary cost of education which is decreasing in their major-specific ability, $v(A_{i,j} + \xi_{i,j})$.[2] At the end of the period, agents learn their ability in their major with certainty, i.e. they learn $\xi_{i,j}$ for their selected $j \in J$. In the second period, agents repeat this investment decision, but may switch major after paying a switching cost. A corresponding variable $d_{i,j,2}$ denotes major selection in the second period. In the investment periods, if the agent is unable to pay for pecuniary costs with their initial wealth, they take on loans, l_i.[3]

In the third period, agents enter the labor market with the ability associated with their second period major–$A_i = A_{i,j} + \xi_{i,j}$ for second period selection of j. Agents select among occupations $k \in K$, where each occupation has (i) a occupation-specific wage-growth joint distribution, $F_k(w, g)$, and a occupation-major specific "match quality", $\phi_{j,k}$. Conditional on the agent's draw of (w, g) and their major, they earn $wA_{i,k}$ in the first period, $gwA_{i,k}$ in the second and $g^2wA_{i,k}$ in the third, where $A_{i,k} = \phi_{j,k}A_i$.[4] If agents borrowed in school, they repay according to some known repayment plan that may depend

[1] In my empirical analysis, I use data from a single institution which has a single posted price for full-time attendance for students within each college who have similar residency status (in-state or out-of-state).

[2] For $j = 0$, this cost is 0. For simplicity, I do not allow those who select the outside option to enter the labor market early, so there is no opportunity cost to school, per se. Instead, the benefit to not attending is captured by the avoiding the non-pecuniary and pecuniary costs of education.

[3] I arbitrarily assume that loans from period t are $l_{i,t} = \max\{x(y_{i,0}) - \frac{y_{i,0}}{2}, 0\}$ and total loans are $l_i = l_{i,1} + l_{i,2}$. Because my analysis focuses on decisions of college graduates, this assumption is not crucial–e.g. I could instead have students exhaust wealth before borrowing at all. In reality, family resources are an important source of college financing (Sallie Mae (2021)).

[4] Here, growth is determined by one parameter, g, but this easily extends to a case where growth can differ over time.

on the total amount borrowed, l_i, or total income, y_t. Call this repayment amount R_t, where repayment amount can differ in each of the working periods.

A.2.2 Labor Market

On the labor market, let β be the discount rate, and working period utility be $U(\{c_t\}_{t=3}^{5}) = \sum_{s=t-3} \beta^s u(c_t)$ subject to $c_t \leq g^{t-3}wA_{i,k} - R_t - s_t + (1+r)s_{t-1}$, where s_t is possible savings/borrowing in period t at rate $(1+r)$. For simplicity, I re-frame the working period so $t = 1$ is the first period of working and $t = 3$ the last.

If there are no liquidity constraints, and assuming $\beta = \frac{1}{1+r}$, then standard assumptions lead to full consumption smoothing in the working period, with optimal consumption, conditional on wage and growth draw:

$$c^* = \frac{(1+r)^2 + g(1+r+g)}{\sum_{s=0}^{2}(1+r)^s} wA_{i,k} - \frac{(1+r)^2}{\sum_{s=0}^{2}(1+r)^s} R_1 - \frac{(1+r)}{\sum_{s=0}^{2}(1+r)^s} R_2 + \frac{(1+r)^3}{\sum_{s=0}^{2}(1+r)^s} s_0$$

Here, I only consider repayment plans that last two periods of working life. If a repayment plan lasts the entirety of working life (e.g. a graduate tax), then there is an additional term: $\frac{1}{\sum_{s=0}^{2}(1+r)^s} R_3$.

Under fixed repayment:

$$R_1 = (1+R)l_i$$
$$R_2 = 0$$

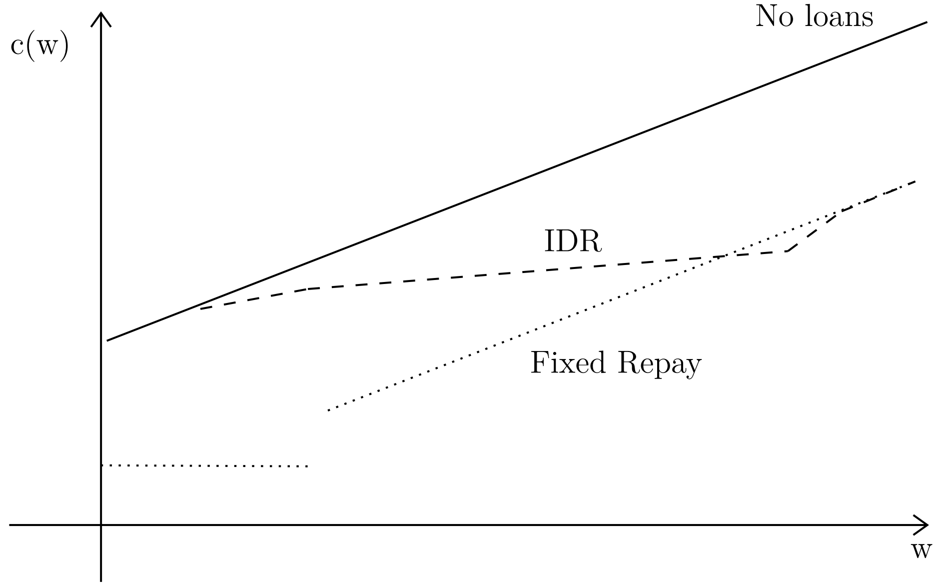
Additionally, there is default risk. If total lifetime earnings are insufficient to cover debt, $\frac{(1+r)^2+g(1+r+g)}{(1+r)^2} wA_{i,k} < (1+R)l_i + \underline{c}$, then the borrower defaults, pays default penalty $\kappa(l_i)$ which increases with loan balance, and consumes $\underline{c}$ for all working periods.

Under IDR

$$R_1 = \min\{\max\{\iota(wA_{i,k} - b), 0\}, (1+R)l_i\}$$
$$R_2 = \min\{\max\{\iota(gwA_{i,k} - b), 0\}, \max\{(1+R)^2 l_i - (1+R)\iota(wA_{i,k} - b), 0\}\}$$

Where ι is the repayment rate and b is the income disregard. There is no default risk, since incomes below b require no repayment, and if earnings are high enough to repay the total amount owed including interest, no additional repayments are required. Note that, for high

Figure A.1: Optimal Consumption Rule



Optimal consumption level for different wage draws, w, assuming fixed income growth, g, and initial savings, s_0, under no repayment, fixed repayment, and IDR.

enough earnings, the repayment under IDR coincides with that of fixed repayment–all loans are repaid in the first period. For low earnings, borrowers only partially repay loans, and the IDR results in less overall payments than fixed repayment. For a region where IDR results in partial repayment or full repayment *in the second period*, it's possible that the total resources used to repay the debt are greater under IDR, assuming that $(1+r) < (1+R)$.[5] Figure A.1 depicts the optimal consumption rules for different income draws under no repayment, fixed repayment, and IDR assuming fixed income growth and common initial savings, s_0.

Consider occupations with non-stochastic growth rates, but stochastic wages distributed $w \sim F_k(w)$.[6] Expected utility for an agent with major j from selecting occupation k is:

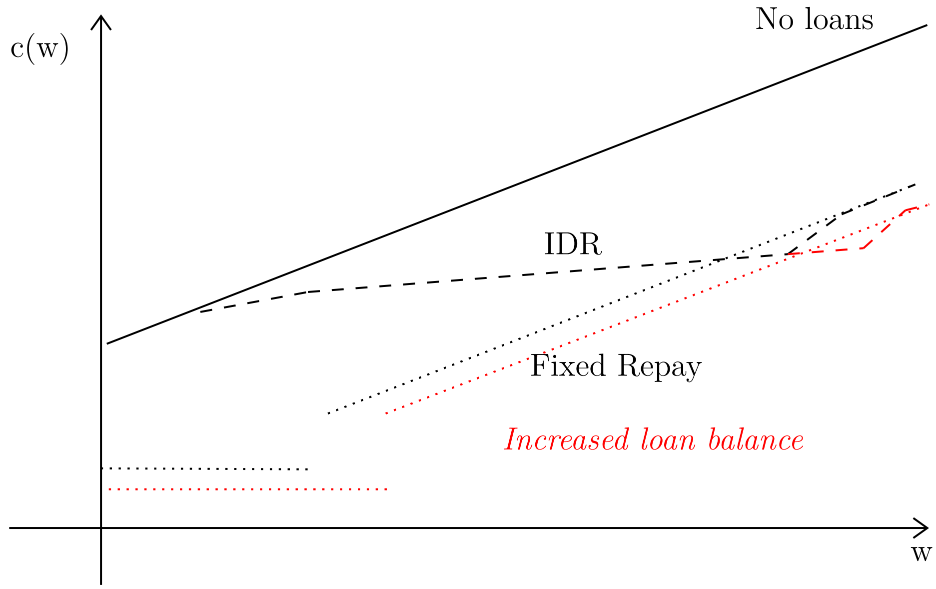
$$EU_{j,k} = (1+\beta+\beta^2)\left\{F(\hat{w_{def}})[u(\underline{c}) - \kappa((1+R)l_i)] + (1 - F(\hat{w_{def}}))E_{w>\hat{w_{def}}|F_k} u(c^*(w))\right\}$$

Where $\hat{w_{def}}$ is the wage draw which triggers default and $E_{w>\hat{w_{def}}|F_k}$ is the expectation with respect to w conditional on $w > \hat{w_{def}}$ according to distribution F_k. Note that, under IDR,

[5] If $(1+r) \geq (1+R)$ then IDR is strictly better than fixed repayment under all income draws. This is because the IDR extends the repayment period, allowing borrowers access to cheaper borrowing than market rates. With equality, the extended repayment time simply acts as an alternative borrowing mechanism at the same price as the market borrowing rate.

[6] Occupations with stochastic growth rates and wages result in similar conclusions, but the distribution of interest is that of the random variable $Z = \frac{(1+r)^2+g(1+r+g)}{1+(1+r)+(1+r)^2)} w$ instead.

Figure A.2: Optimal Consumption Rule, with an increase in l_i



Change in optimal consumption level for different wage draws, w, assuming fixed income growth, g, and initial savings, s_0, under no repayment, fixed repayment, and IDR.

$F(\hat{w_{def}}) = 0$.

To understand the effect of loans on expected utility, consider:

$$\frac{\partial EU_{j,k}}{\partial l_i} = (1+\beta+\beta^2)\Bigg\{ - f_k(\hat{w_{def}})\frac{\partial \hat{w_{def}}}{\partial l_i} - F_k(\hat{w_{def}})\kappa'((1+R)l_i)(1+R) + (1-F(\hat{w_{def}}))E_{w>\hat{w_{def}}|F_w}[u'(c^*(w))c^{*\prime}]\Bigg\}$$

Figure A.2 graphically depicts the change in consumption rules given an increase in loan balance, $ł_i$.

First, consider the fixed repayment loans. Note that an increase in loans *reduces* consumption at all income levels. With risk averse agents with decreasing absolute risk aversion (DARA), this will decrease expected utility from "riskier" wage distributions more than expected utility from less risky wage distributions.[7] To see this explicitly, consider two distributions F_k and F_g such that $EU_{j,k} = EU_{j,g}$ and suppose k is "riskier" than g.[8] Then an increase in loan balances leads to $EU'_{j,k} < EU'_{j,g}$, where $E(\cdot)'$ is the expected utility after the

[7]"Risky" here is in the Rothschild and Stiglitz (1970) sense.

[8]In particular, k is a mean translation of g_0, which is a mean-preserving spread of g. In order for $EU_{j,k} = EU_{j,g}$, it must be that $E_{F_k}(wA_{i,k}) > E_{F_g}(wA_{i,g})$.

loan increase. This follows from Kimball (1990). Intuitively, an increase in loans under fixed repayment is identical to a mandatory reduction in precautionary saving. Thus, a risk avers agent with DARA will select the less risky choice.

Now, consider IDR loans. An increase in loan balances under IDR only affects utility for relatively high wage draws, because higher loan balances require more time to pay off under the IDR.[9] This mutes the effect of loan balances on risk-taking, suggesting that increases in loan balances have smaller effects on risk-taking behavior under an IDR than under fixed repayment. Note, too, that the IDR increases utility for low-wage draws, relative to fixed repayment loans, holding loan balance fixed. This suggests that occupations that have higher weight on low wage draws look especially attractive under IDR relative to fixed repayment.

A natural question here is "is the shift in behavior induced by IDR efficient?" In the context of this model presented thus far, the answer is *no*, at least without considering externalities from particular occupations. Recall, IDR plans offer two major advantages over conventional, fixed repayment plans: (i) consumption smoothing, including avoidance of default risk and (ii) subsidies in the form in debt forgiveness at the end of the repayment window. Here, the first mechanism is redundant, since there is a functioning credit market. Hence, the behavioral response observed here is due completely to the subsidies for risk. These subsidies would only be efficient in the case that they internalized social benefits. Presidents Clinton and Obama both explicitly noted the role IDR plans could play in encouraging borrowers to select socially beneficial occupations. In a radio address announcing what would become the ICR plan, President Clinton (1993) argued that it would "be liberating [for those] with big debts and then feel driven into careers with higher pays but lower satisfaction" citing teaching as an example. Similarly, President Obama (2014) highlighted increased numbers of teachers and social workers as a goal of the PAYE plan. However, while such subsidies *could* be efficient, it is unlikely that they would be well-targeted.

As noted in Quiggin (2014), the consumption smoothing benefit of IDR plans exists in cases with incomplete markets. Consider instead a similar setting as above, but agents cannot borrow: $s_t \geq 0 \ \forall t$. Because incomes generally increase, $g > 1$, this suggests agents will be hand-to-mouth consumers in the labor market. Hence, conditional on the wage draw

[9]In the context of this model, this translates to a higher wage necessary to eventually pay of the loan in the second period, and even higher wages necessary to pay off the loan in the first period.

w, agents utility becomes:

$$U(w) = \begin{cases} u(wA_{i,k} - R_1) + \beta u(gwA_{i,k} - R_2) - \beta^2 u(g^2 wA_{i,k}) & \text{if } wA_{i,k} \geq R_1 + \underline{c} \\ (1+\beta+\beta^2)u(\underline{c}) - \kappa((1+R)l_i) & \text{otherwise} \end{cases}$$

Here, I maintain the assumption that default occurs and results in low consumption across all three periods. Note that the default risk is *higher* here than in the case with borrowing, since agents cannot borrow from future income to pay today's debt–recall that default only occurs under fixed repayment. Additionally, under fixed repayment, the same forces outlined above occur, with loans discouraging risky wage distributions. On the other hand, IDR plans allow for partial consumption-smoothing, relative to the fixed repayment case, and eliminate default risk.

Note that, under IDR, a borrower who does not fully repay in the first period but does fully repay in the second period has consumption profile $c_1 = wA_{i,k} - \iota(wA_{i,k} - b)$ and $c_2 = gwA_{i,k} - (1+R)^2 l_i + (1+R)\iota(wA_{i,k} - b)$. Essentially, the IDR allows consumers to defer debt and save at rate $(1+R)$. This allows for partial consumption smoothing, though not complete, since the consumption profile will still grow. Note that for consumers who do not fully repay, there is still subsidy value from the IDR as well. Hence, in the absence of functioning credit markets, the IDR can act as a mechanism to allow partial consumption smoothing as well.

A.2.3 Investment Rule

Suppose that agents have idiosyncratic preferences over occupations distributed according to a Type-I extreme value function, $\eta_{i,k} \sim EV(1)$. I define $P_{j,k}$ as the *probability* an inidividual with major j selects occupation k:

$$P_{j,k} = \frac{\exp\{EU_{j,k}\}}{\sum_{k' \in K} \exp\{EU_{j,k'}\}}$$

Then the *optimal investment* problem (in the second period) is:

$$\max_{j \in J} EU_{j,2} = \max_{j \in J} \left\{ E_\xi \left[\sum_{k \in K} P_{j,k} EU_{j,k} + u(c_s) - v(A_{i,j} + \xi_{i,j}) - \delta(d_{i,j,1} \neq 1) \right] \right\}$$

Note that the repayment plans affect the investment decision through *expected labor market outcomes.* Under fixed repayment loans, risky investments, i.e. those that are more highly associated with risky labor market outcome–measured through $\phi_{j,k}$, a particular major's "match quality" with occupation k–are less attractive relative to no loans or IDR.[10] Here E_ξ is the expectation with respect to the unknown element of ability, $\xi_{i,j}$. Note that uncertainty over abilities only serves to exacerbate the uncertainty in labor market outcomes, which again suggests that risk averse agents with DARA will avoid majors (i) over which they are especially uncertain of their abilities and (ii) that are associated with more labor market uncertainty.

For example, suppose there are two majors, h and j, and two occupations, k and l. Suppose that the wage distribution for k is a mean-translation of a mean-preserving spread of l such that a non-borrower with $A_{i,k} = A_{i,l}$ is indifferent between these occupations. Further, suppose that $\phi_{h,k} > \phi_{h,l}$ and $\phi_{j,k} < \phi_{j,l}$ with $\phi_{h,k} + \phi_{h,l} = \phi_{j,k} + \phi_{j,l}$. Ignoring uncertainty over ability, non-borrowers sort into their higher ability major, and a non-borrower with $A_{i,j} = A_{i,h}$ is equally likely to select each major.[11] However, *borrowers* of the same level ability are more likely to sort into major j relative to major l, simply due to the labor market uncertainty. That is, the loan repayment decreases expected returns to occupation k more than occupation l, and thus the major more strongly associated with occupation k becomes less attractive relative to the major more strongly associated with occupation l.

Under an IDR plan, however, this effect is muted–especially for occupations with high probability of low-wage draws, either $w < b$ or $w < \hat{w_{def}}$, i.e. below the income disregard (in which case the utility from IDR is identical to a non-borrower's) or in the default region (in which case the utility from fixed repayment includes the default penalty, which is eliminated under IDR). Hence, borrowers under IDR are *more likely* to sort into the "riskier" major. Additionally, the IDR allows for more flexibility in the trade-off between immediate remunerative returns and longer run gains–which is especially apparent if the labor market has imperfections in the credit market which limit borrowers' ability to consumption smooth in the absence of the IDR mechanism.

A.2.4 Additional Considerations

This model only considers pecuniary returns to education/occupations. However, Clinton (1993) specifically cited job satisfaction as a motivating reason to allow for flexible repayment. Simple extensions to this model that account for occupation- or major-specific amenities–e.g. if an occupation with more uncertainty over wage realizations *also* has non-pecuniary benefits

[10] In the extreme case where majors are associated with only one occupation, this problem simplifies to the same problem considered in the labor market section above.

[11] Uncertainty over abilities has a similar effect on optimal decisions as uncertainty in the labor market.

$\alpha_k > 0$–then the fixed repayment loans discourage selection into occupations with these non-pecuniary benefits but risky wage returns as loan balances increase.[12]

Furthermore, note that borrowing in the model is exogenously determined–that is, l_i is determined by family resources, y_i and the pecuniary costs of education $x(y_i)$. However, IDR expansion makes borrowing itself relatively cheaper, from the perspective of the borrower. Insofar as students have the ability to alter their financing decisions to adjust how much the borrow, there could be endogenous selection into treatment. Cadena and Keys (2013) note that a significant minority of low-income students who are eligible for subsidized loans turn them down. Furthermore, in 2012 37.3% of students who applied for financial aid reported turning down a federal loan, and 24.4% of borrowers reported accepting less than the maximum amount they were offered (National Center for Education Statistics (NCES) (2012)). This suggests a sizable minority of students *could* adjust their borrowing behavior in response to the 2013 PAYE expansion. In Section 1.5, I describe a subsample of students who are limited in their ability to endogenously respond to the policy change, which addresses these specific endogeneity concerns.

A.3 Linking Procedure

This project was reviewed by the University of Michigan's IRB and determined to fall under a Category 4 Exemption.[13] This project falls under a research exemption for purposes of FERPA.[14]

The linking procedure between the university that provided the administrative records and Experian Credit Bureau followed the below procedure:

- First, the university created a random, study-specific identifier. They then provided Experian with personally identifiable information (PII) necessary for linking and the random identifier.

[12]This follows directly from noting that α_k, if measured in monetary units, is analogous to the risk premium. This is because it acts as a mean translation of the distribution of wage draws. Again, following Kimball (1990), as loan balances increase, a given value for α_k is insufficient to compensate the increased risk for occupation k.

[13]IRB Study HUM00179860; Category 4 exemption: Use of secondary data collected for non-research purposes; Information recorded such that identity cannot be readily ascertained; Investigator does not contact or re-identify subjects

[14]Use of academic data to improve instruction. Major selection is one of the most important decisions students make. Understanding *why* students declare for particular majors can help institutions tailor coursework for students with particular needs and allocate resources to academic departments according to need.

- Experian queried their data and created an anonymized, individual-level dataset that contained the random identifier and credit bureau records.

- The university then created an anonymized dataset with the random identifier, academic records, and data from the National Student Clearinghouse.

- The two anonymized datasets were provided to the researcher. Data was stored in encrypted format. At no point did the researcher have access to PII, isolating research process from PII. Data were only used for approved research purposes.

A.4 Labor Market Outcomes by Major

Table A.2 contains labor market statistics associated with each major. These measures are calculated using the ACS for the year of student's entrance. For students who begin prior to 2009, I use ACS data from 2009. All statistics are early-career outcomes–i.e. for ages 25-29–with the exception of income growth, which compares average income between early-career and mid-career (i.e. 25-29 and 30-34).

Table A.2: Labor Market Outcomes by Major

Major	Early Career						Average Prime-Age
	Avg. Income ($2019)	Income S.D.	Unemp Rate	Not-in-LF	Income Growth	Share<FPL	Earnings (millions of $2019)
Natural Resources	$41,400	$24,900	0.062	0.060	0.393	0.068	$1.67
Architecture	$47,700	$27,500	0.077	0.103	0.304	0.081	$1.94
Area, Ethnic, Cultural, & Gender Studies	$40,500	$33,400	0.062	0.130	0.664	0.116	$1.86
Communication/Journalism	$44,800	$31,400	0.049	0.066	0.356	0.054	$1.70
Computer Science	$65,700	$46,800	0.045	0.107	0.272	0,065	$2.45
Education	$40,600	$21,000	0.030	0.087	0.203	0.042	$1.52
Engineering	$68,300	$39,500	0.038	0.093	0.322	0.061	$2.69
Engineering Tech.	$58,800	$42,000	0.052	0.095	0.217	0.053	$2.18
Foreign Languages	$41,100	$32,300	0.054	0.141	0.454	0.114	$1.72
English	$41,400	$33,400	0.057	0.112	0.413	0.074	$1.75
Liberal Arts & Sciences	$41,000	$29,100	0.044	0.095	0.299	0.067	$1.82
Biological/Biomedical Sci.	$46,700	$36,800	0.039	0.187	0.673	0.135	$2.42
Math	$56,300	$49,000	0.042	0.106	0.425	0.060	$2.37
Interdisciplinary Studies	$44,100	$36,000	0.049	0.135	0.397	0.098	$1.83
Parks, Recreation, Leisure	$42,100	$30,500	0.044	0.100	0.366	0.086	$1.64
Philosophy	$44,000	$40,400	0.071	0.146	0.482	0.099	$1.82
Physical Sci.	$48,500	$37,500	0.039	0.143	0.499	0.099	$2.36
Psychology	$40,900	$27,500	0.049	0.112	0.384	0.078	$1.69
Public Administration	$39,700	$28,200	0.043	0.087	0.246	0.069	$1.63
Social Sci.	$53,100	$46,900	0.054	0.103	0.486	0.075	$2.24
Visual & Perf. Arts	$36,700	$29,100	0.068	0.093	0.311	0.090	$1.50
Health	$54,300	$33,200	0.024	0.096	0.216	0.059	$1.91
Business	$55,300	$41,400	0.042	0.077	0.340	0.048	$2.18
History	$45,100	$38,300	0.067	0.103	0.492	0.085	$2.03
Total	$50,600	$36,300	0.047	0.106	0.385	0.076	$2.11

Note: ACS data from student's entry year used to quantify labor market outcomes by major. Income statistics conditional on being employed

A.5 Additional Summary Statistics

Table A.3 reports summary statistics by both graduation cohort and borrower status. These characteristics are determined *prior* to enrollment. All measures are drawn from administrative records, with the exception of household income, which is drawn from ACS data for the students' ZIP-Code.

Table A.3: Summary Statistics by Graduation Cohort and Borrower Status

Grad. Cohort	Instate		Female		White		Underrep. Minority	
	NonBor.	Bor.	NonBor.	Bor.	NonBor.	Bor.	NonBor.	Bor.
2009	0.579	0.655	0.527	0.496	0.721	0.718	0.090	0.177
2010	0.543	0.689	0.462	0.505	0.701	0.731	0.095	0.191
2011	0.531	0.688	0.491	0.517	0.690	0.768	0.085	0.146
2012	0.580	0.758	0.495	0.524	0.739	0.759	0.069	0.159
2013	0.586	0.721	0.498	0.522	0.779	0.769	0.058	0.145
2014	0.498	0.698	0.460	0.496	0.777	0.762	0.065	0.152
2015	0.508	0.696	0.490	0.512	0.701	0.715	0.057	0.161
2016	0.524	0.684	0.496	0.515	0.705	0.704	0.060	0.151
2017	0.523	0.677	0.487	0.516	0.714	0.730	0.058	0.164
2018	0.492	0.657	0.480	0.512	0.729	0.753	0.068	0.164
2019	0.534	0.646	0.499	0.533	0.698	0.708	0.090	0.165
	Parent Col.+		Mean HH Inc. (ZIP-1,000s of 2019$)		HS GPA			
	NonBor.	Bor.	NonBor.	Bor.	NonBor.	Bor.		
2009	0.895	0.759	$129	$103	3.735	3.733		
2010	0.900	0.752	$130	$103	3.724	3.723		
2011	0.909	0.739	$131	$101	3.742	3.745		
2012	0.916	0.741	$130	$97	3.753	3.740		
2013	0.922	0.749	$129	$101	3.762	3.758		
2014	0.929	0.758	$134	$102	3.742	3.738		
2015	0.936	0.775	$134	$100	3.775	3.784		
2016	0.946	0.795	$131	$99	3.797	3.803		
2017	0.934	0.784	$128	$101	3.809	3.804		
2018	0.932	0.810	$130	$101	3.807	3.811		
2019	0.941	0.846	$125	$101	3.824	3.829		

A.6 Identification with Compositional Changes

A.6.1 Preliminaries

Consider the standard potential outcomes framework:

$$Y = (1-D)(1-T)Y(0,0) + D(1-T)Y((1,0) + (1-D)TY(0,1) + DTY(1,1)$$

Where $Y(D,T)$ is outcome in time T for group D. Here, T indexes time period and D indexes treatment status. For the time being, I consider a standard 2X2 setting, i.e. $D \in \{0,1\}$ and $T \in \{0,1\}$. I am interested in the Average Treatment on the Treated, i.e.

$$ATT = E[Y(1,1) - Y(0,1)|D = 1, T = 1]$$

Obviously, I do not observe $Y(0,1)|D = 1$, necessitating additional assumptions for identification. In particular, I assume that the unobserved counterfactual outcome, $E(Y(0,1)|D = 1, T = 1, X)$ is captured by the observed trend from the untreated group plus the treated group's starting value, i.e. the conditional parallel trends assumption:

$$\begin{aligned} &E(Y(0,1)|X, D = 1, T = 1) \\ &= E(Y(1,0)|X, D = 1, T = 0) + E(Y(0,1)|X, D = 0, T = 1) - E(Y(0,0)|X, D = 0, T = 0) \end{aligned} \tag{A1$'$}$$

I also make a common support assumption:

$$P(D = 1|T = t) > 0 \text{ and } P(D = 1|X, T = t) < 1 \text{ almost surely } \forall t \in \{0,1\} \tag{A2}$$

Here, note that the common support assumption must hold *for all time periods*. Because the propensity score may differ over time, and the weights proposed below essentially reweight the estimate to match the treatment group *in the treated period*, common support across time is necessary for the weights to be well defined.[15] Furthermore, I assume that data are drawn from a mixture distribution:

Let $P(T = 1) = \lambda$. Conditional on $T = t$, data are i.i.d. from the distribution (Y(t),D,X) s.t. the sampling follows the mixture distribution:

[15] In particular, to ensure the weight used to integrate over the conditional ATTs and recover the population ATT is not an out-of-sample prediction, relative to the covariates X.

$$P_M(Y=y,D=d,X=x,T=t) = \lambda t P(Y(1)=y,D=d,X=x) \\ +(1-\lambda)(1-t)P(Y(0)=y,D=d,X=x) \tag{A3}$$

Note that assumptions A1′, A2, and A3 are identical to those of Abadie (2005), except I make explicit that compositional shifts are not ruled out, i.e. $P(D=1|X,T=0) \neq P(D=1|X,T=1)$ is allowed.

I make one additional assumption that will identify the population ATT from the conditional ATT:

$$X \perp T \implies F(X|T=t) = F(X) \tag{A4}$$

That is, the population characteristics, X, do not change over time, although I allow the particular composition of treatment and control groups to change (i.e. $F(X|D=1,T=1) \neq F(X|D=1,T=0)$)

A.6.2 Identification under Compositional Change

Below, I restate Abadie (2005)'s Lemma 3.2, which is the building block for ATT identification with repeated cross sections. However, I alter the weights to explicitly allow for compositional shifts; in particular, $P(D=1|X,T=0) \neq P(D=1|X,T=1)$ is allowed.

Lemma 1. *If Assumptions A1′, A2, and A3 hold, then*

$$ATT|X = E[Y(1,1) - Y(0,1)|D=1,T=1,X] = E_M[\omega Y|X]$$

where $E_M[\cdot]$ is the expectation w.r.t. the mixture distribution $P(Y,D,X,T)$ and

$$\omega = \frac{T-\lambda}{\lambda(1-\lambda)} \\ \times \frac{D - TP(D=1|X,T=1) - (1-T)P(D=1|X,T=0)}{TP(D=1|X,T=1)P(D=1|X,T=0) + (1-T)P(D=1|X,T=0)P(D=0|X,T=0)}$$

For expositional simplicity in the below proof, I define:

$$\rho = \\ \frac{D - TP(D=1|X,T=1) - (1-T)P(D=1|X,T=0)}{TP(D=1|X,T=1)P(D=1|X,T=0) + (1-T)P(D=1|X,T=0)P(D=0|X,T=0)}$$

Hence, $\omega = \frac{T-\lambda}{\lambda(1-\lambda)}\rho$.

Proof.

$$
\begin{aligned}
E_M(\omega Y|X) &= E_M(E_M(\omega Y|X,T)|X) \\
&= E_M(\lambda E(\omega Y|X,T=1) + (1-\lambda)E(\omega Y|X,T=0)|X) \\
&= E_M(E(\rho Y|X,T=1) - E(\rho Y|X,T=0)|X) \\
&= E_M(P(D=1|X,T=1)E(\rho Y|X,T=1,D=1) \\
&+ P(D=0|X,T=1)E(\rho Y|X,T=1,D=0) \\
&- \{P(D=1|X,T=0)E(\rho Y|X,T=0,D=1) \\
&+ P(D=0|X,T=0)E(\rho Y|X,T=0,D=0)\}|X) \\
&= \{E(Y|X,T=1,D=1) - E(Y|X,T=1,D=0)\} \\
&- \{E(Y|X,T=0,D=1) - E(Y|X,T=0,D=0)\}
\end{aligned}
$$

Note: by Assumption A1′, $\{E(Y|X,T=1,D=1) - E(Y|X,T=1,D=0)\} - \{E(Y|X,T=0,D=1) - E(Y|X,T=0,D=0)\} = E(Y(1,1) - Y(0,1)|X,D=1,T=1)$, which completes the proof □

Thus, by allowing the weights suggested by Abadie (2005) to vary with time, $P(D=1|X,T=t)$, I recover the identification of the conditional ATT. To identify the population ATT, integrate over the distribution of X:

$$
\begin{aligned}
&E[Y(1,1) - Y(0,1)|D=1,T=1] \\
&= \int E(Y(1,1) - Y(0,1)|X,D=1,T=1)dF(X|D=1,T=1) \\
&= \int E_M(\omega Y|X)dF(X|D=1,T=1) \\
&= \int E_M(\omega Y|X)\frac{P(D=1|X,T=1)}{P(D=1|T=1)}dF(X|T=1) \\
&= \int E_M(\omega Y|X)\frac{P(D=1|X,T=1)}{P(D=1|T=1)}dF(X) \\
&= E_M\left[\frac{P(D=1|X,T=1)}{P(D=1|T=1)}\omega Y\right]
\end{aligned}
$$

Where I use Assumption A4 to substitute $F(X)$ for $F(X|T=1)$.[16]

The intuition here is I first reweight observations by ω, allowing for direct comparison

[16]Note: if there are population compositional changes such that $F(X|T=1) \neq F(X)$, I conjecture that an additional reweighting in the spirit of DiNardo et al. (1996) would complete the identification of the population ATT.

between the treatment and control groups within one time period. I then reweight by $\frac{P(D=1|X,T=1)}{P(D=1|T=1)}$ to make treatment groups comparable across time periods.

A.6.3 Extension to Multiple Time Periods and Continuous Treatment

With multiple treatment (or pre-treatment) periods, the above assumptions identify $ATT(t) = E(Y(1,t) - Y(0,t)|D = 1, T = t)$ for $t \in \{-T_{\text{pre}}, \ldots, -1, 1, \ldots, T_{\text{post}}\}$ where there are T_{pre} pre-treatment periods and T_{post} post-treatment periods. The proof for this follows the same steps above for each time period separately, all relative to the base period $t = 0$.

Following Abadie (2005), consider a continuous treatment $D \in \mathbb{R}_+$ instead of a binary treatment. Again, suppose $D = 0$ is the not-treated value. Then, under a similar parallel trends assumption for each level of treatment (and each time period, if there are multiple time periods),

$$\begin{aligned}
&E(Y(d,0) + Y(0,t) - Y(0,0)|X, D = d, T = t) \\
&= E(Y(d,0)|X, D = d, T = 0) + E(Y(0,t)|X, D = 0, T = t) - E(Y(0,0)|X, D = 0, T = 0)
\end{aligned} \tag{A1$''$}$$

corresponding average treatment effects on the treated parameters for each treatment and time period, $ATT(d,t)$, can be recovered.[17] In particular, substitute ω presented above with the following weight:

$$\begin{aligned}
\omega^d = &\frac{\mathbb{1}(T = t) - \lambda_t}{\lambda_t(1 - \lambda_t)} \\
&\times \frac{\mathbb{1}(D = d) - \mathbb{1}(T = t)f(D = d|X, T = t) - \mathbb{1}(T = 0)f(D = d|X, T = 0)}{A}
\end{aligned}$$

Where

$$\begin{aligned}
A = &[\mathbb{1}(T = t)f(D = d|X, T = t) + \mathbb{1}(T = 0)f(D = d|X, T = 0)] \\
&\times[\mathbb{1}(T = t)f(D = d|X, T = t) + \mathbb{1}(T = 0)f(D = d|X, T = 0)]
\end{aligned}$$

and $f(\cdot)$ is the conditional probability density function for each treatment level d, $\mathbb{1}(\cdot)$ is the indicator function $= 1$ if the argument is true and $= 0$ otherwise, and $\lambda_t = \frac{P(T=t)}{P(T=t)+P(T=0)}$. Similarly, in order to integrate the conditional ATT to recover the population ATT, use the treated population's conditional density over the unconditional density, both from the treated

[17] Note that this is the *weak* parallel trends assumption, as presented in Callaway et al. (2021).

period in question:

$$E(Y(d,t) - Y(0,t)|D = d, T = t) = E_M\left[\frac{f(D = d|X, T = t)}{f(D = d|T = t)}\omega' Y\right]$$

As noted above, these weighted expectations are calculated separately for each pre-treatment and each post-treatment period separately, all relative to the base period $T = 0$. In essence, this approach separately estimates each level of treatment relative to the no-treatment control separately *and* each time periods relative to the base period separately. On one hand, this approach makes the standard difference-in-differences assumptions, requiring no stronger assumptions that are potentially less clear.[18] On the other hand, this is a data-intensive process, since comparisons are only between treatment groups and the base group–not treatment groups to other treatment groups–and treatment periods to the base period–not treatment periods to other treatment periods.

A.6.4 Comparison to Abadie (2005) and Hong (2013)

Here, I briefly comment on comparisons between this approach and those of Abadie (2005) and Hong (2013).

First, Abadie (2005) does not rule out flexible estimation of propensity scores.[19] In particular, his method can allow for parametric estimation of the propensity score *fully saturated by time variable*. That is, estimating $\hat{\pi}(X) = \pi(X\hat{\gamma}_t)$ where π is the known function determining treatment status and $\hat{\pi}$ is the estimated propensity score via maximum likelihood estimation. Note that the estimated parameters, $\hat{\gamma}$, are allowed to vary by time $t \in \{0, 1\}$. Under this parametric estimation approach, the propensity scores recovered by the Abadie (2005) approach and my proposed approach are the same. However, the *weights* used in identification differ. This is because the final integration term used to identify population ATT is $\frac{P(D=1|X)}{E(D=1)}$ in the Abadie (2005) setting and $\frac{P(D=1|X,T=1)}{E(D=1|T=1)}$ in my setting. Intuitively, my approach explicitly reweights observations to "look like" the treated *in the treatment period*, whereas the Abadie (2005) approach reweights to "look like" the treated only.[20]

Hong (2013) uses a different tack to identify the ATT under compositional changes. In particular, he begins with a conditional independence assumption which implies the

[18] See, e.g., Callaway et al. (2021) for discussion on the stronger assumptions that are implicitly (or explicitly) invoked when more structure is placed on the problem.

[19] On the contrary, Abadie (2005) presents flexible parametric and non-parametric techniques for estimating the propensity score.

[20] Obviously, if there are no compositional changes over time, there is no difference.

parallel trends assumption.[21] If the distribution of covariates is independent of treatment status and time conditional on propensity score, $X \perp (D,T)|P$, then Hong (2013) shows that the ATT is identified using estimators for unobserved outcomes, calculated by matching on the multidimensional propensity score:

$$A\hat{T}T = \sum_{i \in G_{D=1,T=1}} Y_i - \hat{E}[Y_j|D_j = 1, T_j = 0, P_{0,i}, P_{1,i}] - \hat{E}[Y_j|D_j = 0, T_j = 1, P_{0,i}, P_{1,i}] + \hat{E}[Y_j|D_j = 0, T_j = 0, P_{0,i}, P_{1,i}]$$

Where $\hat{E}$ are the conditional estimators of the outcome conditional on the propensity scores $P_{t,i} = P(D = 1|T = t, X_i)$. Hong (2013)'s approach is an extension of the "outcome regression" approach of Heckman et al. (1997) and Heckman et al. (1998).

In contrast, my approach–following Abadie (2005)–makes use of the law of total expectations and reweighting of outcomes to recover an estimate of the ATT. This avoids explicitly modeling the conditional estimators of the outcomes. This approach is relatively computationally tractable, and has simple extensions to multi-level/continuous treatment exposures and multiple time periods.

It should be noted that there are advantages to both approaches. Furthermore, as noted by Sant'Anna and Zhao (2020), in the unidimensional case, an approach that combines the outcome regression approach of Heckman et al. (1997), Heckman et al. (1998) and Abadie (2005) is doubly robust, in the sense that it is robust to either misspecifications in the outcomes modeling or the propensity score estimation.

[21] In particular, the untreated potential outcomes are independent of the treatment and time period conditional on covariates, X, i.e. $\{Y(D = 1, T = 0); Y(D = 0, T = 0); Y(D = 0, T = 1)\} \perp (D,T)|X$

A.7 Propensity Score Distributions

Figure A.3: Estimated Propensity Scores using standard weighting

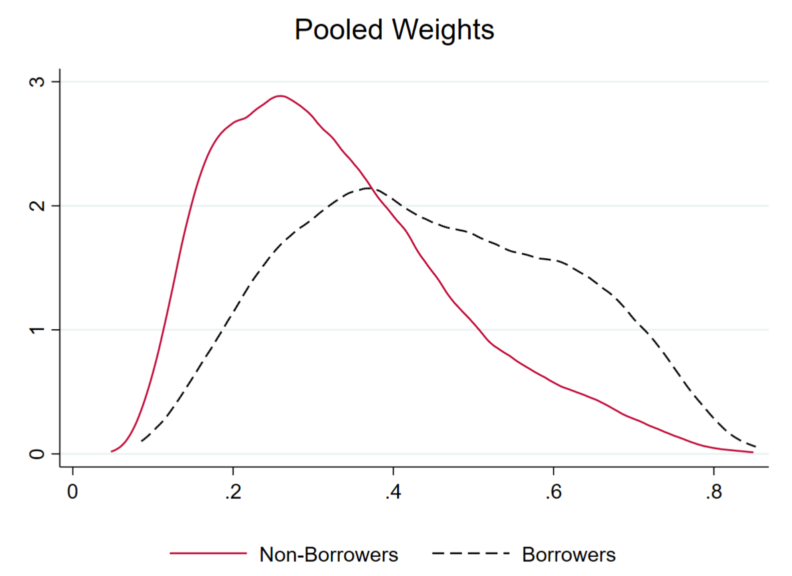
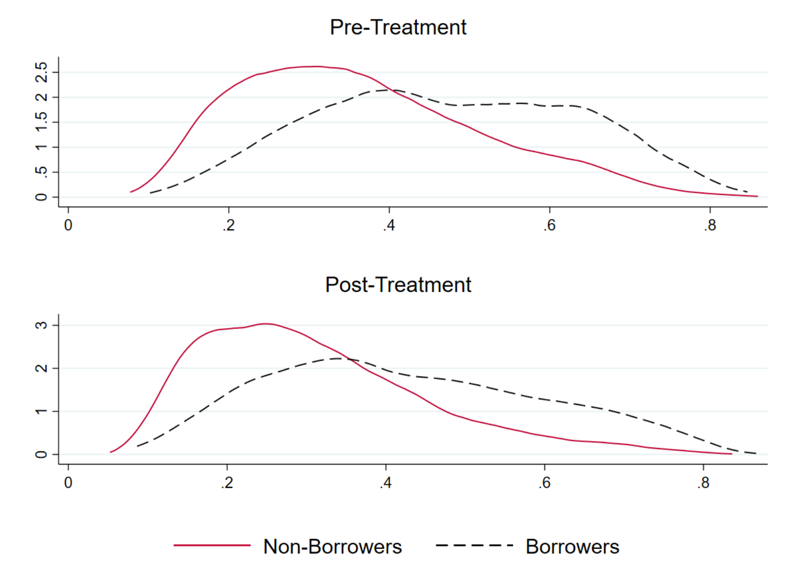
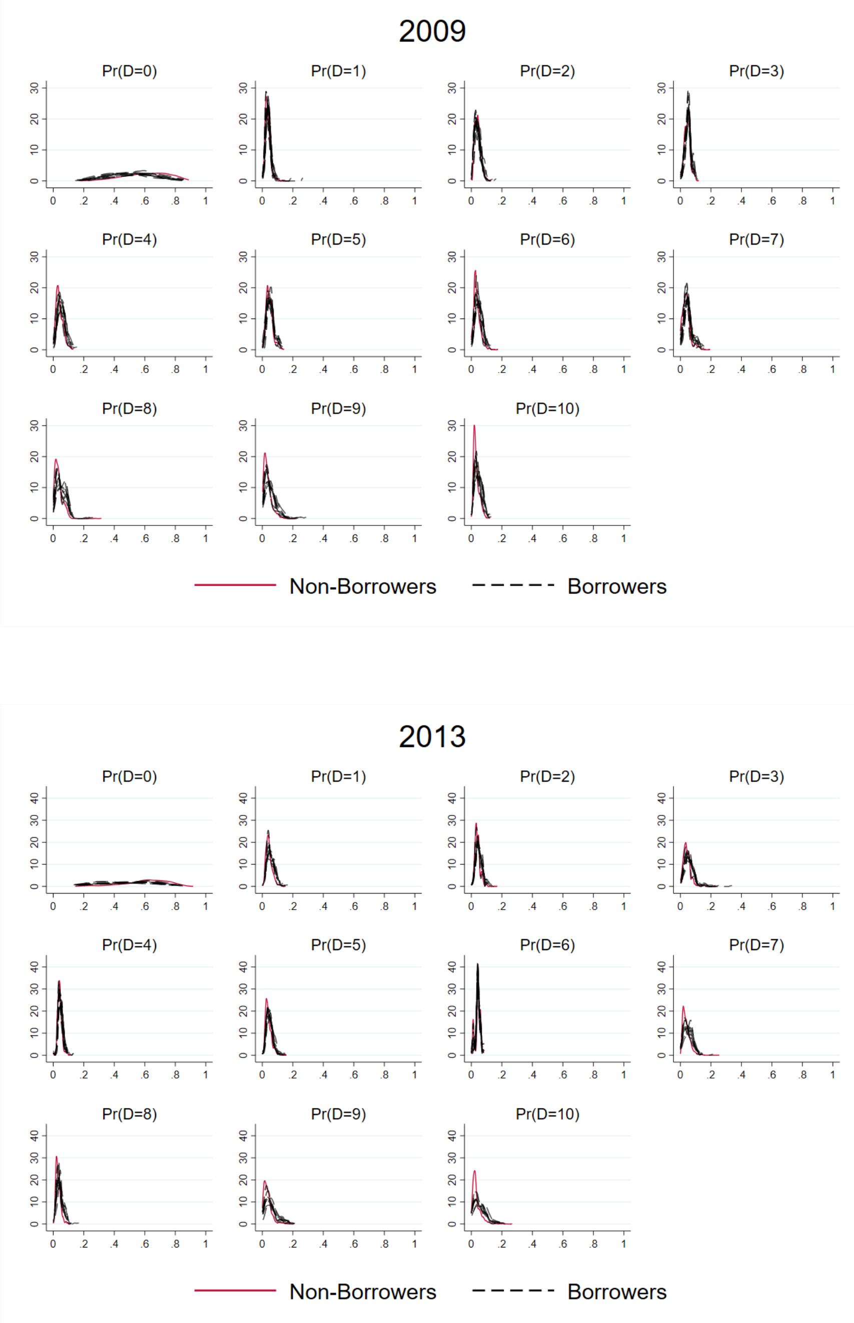


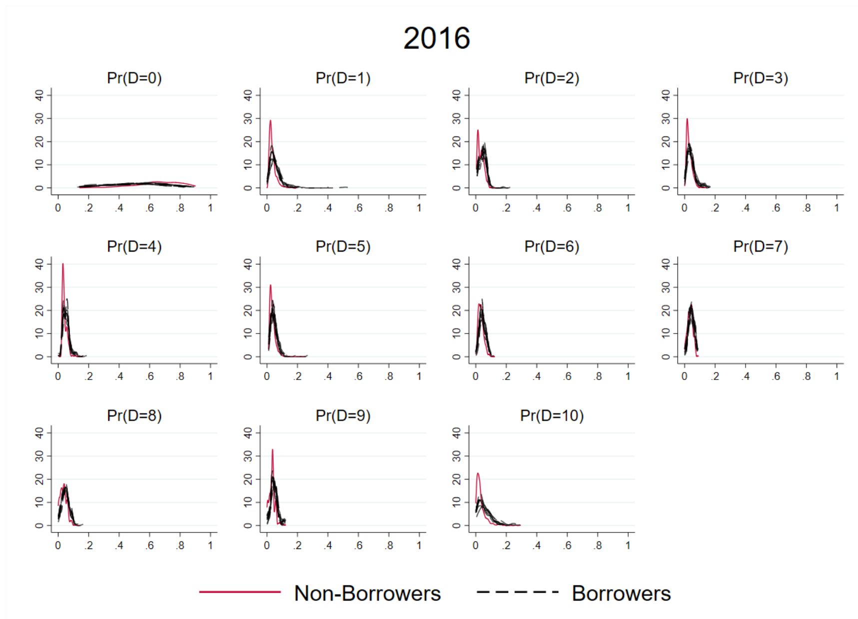
Figure A.4: Estimated Propensity Scores using compositional shift weighting

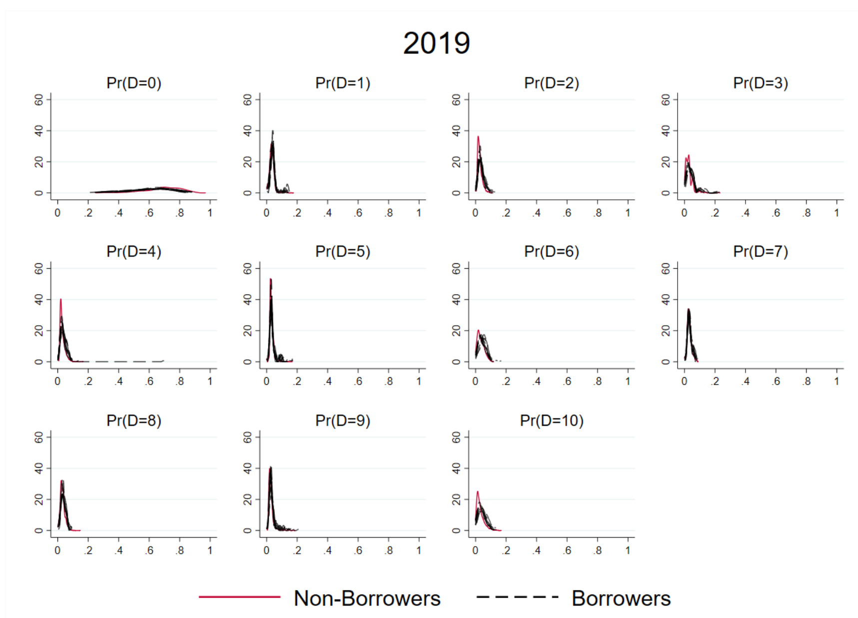


Propensity scores estimated using a logit model of borrowing status on parental education, the number of other dependents in the students' household, a quadratic in household income (proxied by home ZIP-code income), an indicator for whether the student is white, an indicator for residency status, an indicator for if they enter college with enough credits to be considered a second- or above-year student, and their high school GPA.

Figure A.5: Estimated Propensity Scores using compositional shift weighting, multilevel treatment (select years)







Propensity scores estimated using a multinomial logit model of borrowing status on parental education, the number of other dependents in the students' household, a quadratic in household income (proxied by home ZIP-code income), an indicator for whether the student is white, an indicator for residency status, an indicator for if they enter college with enough credits to be considered a second- or above-year student, and their high school GPA. Treatment exposure split into 11 categories: non-borrowers (D=0) and deciles of borrower (D=$\{1, \ldots, 10\}$).

A.8 Multinomial Logit Analysis of Major Choice

An alternative way of framing major selection is to use an unordered choice model among majors–whereas the main results I present order major by their associated labor market outcomes, using the motivating theory as a guide, here I instead use a multinomial choice model to estimate how the IDR expansion in 2013 affected major selection.

Let j index each major and let $Y^j_{i,c}$ be the latent utility associated with major j for individual i of graduation cohort c. Then the equation

$$Y^j_{i,c} = \alpha^j_0 + \alpha^j_d \mathbb{1}(D_{i,c} = 1) + \alpha^j_t \mathbb{1}(T_{i,c} = 1) + \gamma^j \mathbb{1}(D_{i,c} = 1, T_{i,c} = 1) + \varepsilon^j_{i,c}$$

Where $D_{i,c}$ denotes the treatment exposure–that is, borrowing status–and $T_{i,c}$ denotes time period, with $T = 0$ denoting the period prior to the policy expansion (pre-treatment period), i.e. graduation cohorts 2009-2013, and $T = 1$ the period after the policy expansion (treatment period), i.e. 2014-2019. This characterizes the effect of the policy expansion on the latent utility of individual i. The multinomial logit model estimates the probabilities for each major j:

$$Pr(Y^j_{i,c} > 0) = \frac{\exp\{\alpha^j_0 + \alpha^j_d \mathbb{1}(D_{i,c} = 1) + \alpha^j_t \mathbb{1}(T_{i,c} = 1) + \gamma^j \mathbb{1}(D_{i,c} = 1, T_{i,c} = 1)\}}{1 + \sum_{k \neq 24} \exp\{\alpha^k_0 + \alpha^k_d \mathbb{1}(D_{i,c} = 1) + \alpha^k_t \mathbb{1}(T_{i,c} = 1) + \gamma^k \mathbb{1}(D_{i,c} = 1, T_{i,c} = 1)\}}$$

CIP code 24 (liberal arts) is selected as the excluded major in order to estimate the model, an arbitrary choice that has no effect on the probability calculations below.

Table A.4 reports the results from the multinomial logit model. The coefficients reported here are the beta coefficients from the model, that is the marginal change in the log-odds ratio (relative to the base major) with respect to the explanatory variable. Unlike the linear model, where the coefficient γ captures the treatment effect of interest, here we cannot interpret γ as the causal parameter of interest, which is generally true of such nonlinear models (Puhani (2012)). Instead, the parameter of interest is the *change* in probability induced by borrowing status in the treatment period relative to a counterfactual probability. Analogous to the difference-in-differences model that uses change in non-borrower behavior over time to estimate the counterfactual behavior of the borrower behavior in the second period, here I use the change in non-borrower probability to estimate the counterfactual probability in the second period, similar to the approach outlined by Puhani (2012). That is, I report $\hat{P}(Y^j > 0|D = 1, T = 1) - \hat{P}(Y^j > 0|D = 1, T = 0) - \{\hat{P}(Y^j > 0|D = 0, T = 1) - \hat{P}(Y^j > 0|D = 0, T = 0)\}$ with probability estimates based off the multinomial logit model.

Table A.5 reports the two cross differences, $\hat{P}(Y^j > 0|D = 1, T = 1) - \hat{P}(Y^j > 0|D = 1, T = 0)$ and $\{\hat{P}(Y^j > 0|D = 0, T = 1) - \hat{P}(Y^j > 0|D = 0, T = 0)\}$, as well as the difference between these cross-differences, which can be interpreted as the shift in probability among borrowers induced by the policy expansion. After policy expansion, there were notable relative decreases in borrowers selecting Computer Science, Business, Social Science, and Architecture as majors, while there were relative increases in History, Biological/Biomedical Sciences, Psychology, and Physical Sciences.[22] However, not all of these shifts are significant at common statistical inference thresholds. However, I view these results of unordered major selection as suggestive that the effect does *not* appear to be driven by selection towards or away from particular "groups" of majors, e.g. STEM majors. I address this point in Section 1.6 as well.

Table A.6 reports the cross differences in $\hat{P}(Y^j > 0|D = 1, T = t, X) - \hat{P}(Y^j > 0|D = 0, T = t, X)$ for each time period t. Notably, these cross differences are (i) *within* time period and *across* borrower type and (ii) control for additional covariates: namely parental education, a quadratic in household income, an indicator for if the student is white, the student's high school GPA, and residency status. This table is reported to highlight that, even after the policy expansion, borrowers and non-borrowers have different likelihood of selecting particular majors. While not conclusive, this suggests that credit constraints remain an important consideration in major selection.

[22]"Relative" here refers to the borrowers' counterfactual behavior, inferred by the difference in cross-differences.

Table A.4: Multinomial Logit Model Estimation–Men

Major	Treatment Period (α_t^j)	Borrowing Status (α_d^j)	Interaction Effect (γ^j)
Natural Resources	-0.157	0.001	-0.155
	(0.203)	(0.300)	(0.377)
Architecture	-0.661**	0.955***	-0.443
	(0.324)	(0.355)	(0.405)
Area, Ethnic, Culture & Gender Studies	-0.563**	-0.107	0.230
	(0.256)	(0.310)	(0.383)
Communication /Journalism	-0.141	-0.381	0.407
	(0.196)	(0.281)	(0.339)
Computer Science	0.964***	0.083	-0.258
	(0.141)	(0.188)	(0.220)
Education	-0.595**	0.488	-0.128
	(0.279)	(0.318)	(0.488)
Engineering	-0.120	0.210	-0.031
	(0.142)	(0.184)	(0.228)
Engineering Tech.	0.131	0.395	-0.218
	(3.888)	(4.849)	(4.956)
Foreign Languages	-0.467**	0.108	0.316
	(0.232)	(0.231)	(0.334)
English	-0.661***	0.052	0.057
	(0.207)	(0.219)	(0.340)
Liberal Arts & Sciences	0	0	0
	(·)	(·)	(·)
Biological /Biomedical Sci.	-0.151	0.078	0.094
	(0.131)	(0.202)	(0.222)
Math	-0.198	-0.218	0.234
	(0.195)	(0.280)	(0.331)
Interdisciplinary Studies	1.095***	-0.099	-0.188
	(0.237)	(0.310)	(0.336)
Parks, Recreation, Leisure	0.082	0.002	-0.255
	(0.144)	(0.252)	(0.274)
Philosophy	-0.494**	-0.164	0.119
	(0.240)	(0.509)	(0.546)

Physical Sci.	-0.211	0.054	0.325
	(0.219)	(0.264)	(0.329)
Psychology	-0.342**	0.112	0.166
	(0.144)	(0.187)	(0.247)
Public Admin.	-0.041	-0.298	-0.263
	(0.259)	(0.357)	(0.418)
Social Sci.	-0.047	-0.030	-0.157
	(0.126)	(0.162)	(0.224)
Visual	-0.289**	0.281	0.094
& Perf. Arts	(0.141)	(0.227)	(0.240)
Health	0.595*	0.705	-0.382
	(0.335)	(0.430)	(0.565)
Business	0.158	-0.492**	-0.145
	(0.153)	(0.196)	(0.281)
History	-0.871***	-0.237	0.456
	(0.163)	(0.232)	(0.279)

$n = 17,895$; $*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .10$; Bootstrapped standard errors reported.

Table A.5: Cross Differences in Estimated Probabilities–Men

Major	1^{st} Cross Diff. (A)	2^{nd} Cross Diff. (B)	Diff. in Cross Diffs. (B-A)
Natural Resources	-0.159 p.p. (0.188)	-0.234 p.p. (0.357)	-0.075 p.p. (0.353)
Architecture	-0.377 p.p.** (0.162)	-1.271 p.p.*** (0.314)	-0.894 p.p.*** (0.346)
Area, Ethnic, Culture & Gender Studies	-0.398 p.p.** (0.173)	-0.186 p.p. (0.158)	0.212 p.p. (0.246)
Communication /Journalism	-0.197 p.p. (0.277)	0.409 p.p. (0.269)	0.606 p.p. (0.384)
Computer Science	8.805 p.p.*** (0.473)	6.655 p.p.*** (0.484)	-2.150 p.p.*** (0.636)
Education	-0.339 p.p.** (0.151)	0.574 p.p.** (0.229)	-0.235 p.p. (0.319)
Engineering	-2.491 p.p.** (1.025)	-2.100 p.p.** (1.012)	0.401 p.p. (1.427)
Engineering Tech.	0.011 p.p. (0.285)	-0.002 p.p. (0.314)	-0.012 p.p. (0.433)
Foreign Languages	-0.726 p.p.** (0.368)	-0.164 p.p. (0.345)	0.562 p.p. (0.427)
English	-1.380 p.p.*** (0.316)	-1.208 p.p.*** (0.414)	0.172 p.p. (0.592)
Liberal Arts & Sciences	0.006 p.p. (0.151)	0.086 p.p. (0.213)	0.080 p.p. (0.256)
Biological /Biomedical Sci.	-1.052 p.p.*** (0.403)	0.114 p.p. (0.616)	1.166 p.p.* (0.687)
Math	-0.469 p.p. (0.345)	0.236 p.p. (0.348)	0.704 p.p. (0.532)
Interdisciplinary Studies	2.011 p.p.*** (0.269)	1.460 p.p.*** (0.348)	-0.551 p.p. (0.382)
Parks, Recreation, Leisure	0.279 p.p. (0.343)	-0.290 p.p. (0.427)	-0.569 p.p. (0.563)
Philosophy	-0.265 p.p.* (0.140)	-0.148 p.p. (0.238)	0.118 p.p. (0.308)

Physical Sci.	-0.347 p.p.	0.386 p.p.	0.733 p.p.*
	(0.282)	(0.373)	(0.430)
Psychology	-1.565 p.p.***	-0.588 p.p.	0.977 p.p.
	(0.474)	(0.422)	(0.718)
Public Admin.	-0.047 p.p.	-0.197 p.p.	-0.150 p.p.
	(0.276)	(0.229)	(0.316)
Social Sci.	-0.622 p.p.	-1.758 p.p.***	-1.137 p.p.
	(0.862)	(0.653)	(1.151)
Visual	-1.177 p.p.***	-0.712 p.p.	0.465 p.p.
& Perf. Arts	(0.323)	(0.685)	(0.829)
Health	0.221 p.p.**	0.174 p.p.	-0.048 p.p.
	(0.107)	(0.214)	(0.247)
Business	2.310 p.p.***	0.684 p.p.	-1.626 p.p.
	(0.746)	(0.757)	(1.058)
History	-2.031 p.p.***	-0.779 p.p.*	1.252 p.p.**
	(0.380)	(0.398)	(0.560)

$*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .10$; Probabilities estimated using multinomial logit model of choice. Standard errors calculated using Delta-method based on bootstrapped s.e. reported in multinomial logit model, Table A.4. The first cross difference (column A) is $\{\hat{P}(Y^j|D = 0, T = 1) - \hat{P}(Y^j|D = 0, T = 0)\}$ while the second cross difference (column B) is $\{\hat{P}(Y^j|D = 1, T = 1) - \hat{P}(Y^j|D = 1, T = 0)\}$.

Table A.6: Cross Differences in Estimated Probabilities–Men; Within time period, across borrower status

Major	1st Cross Diff. (T=0)	2nd Cross Diff. (T=1) (T=1)	Diff. in Cross Diffs. ((T=1)-(T=0))
Natural Resources	-0.093 p.p. (0.251)	-0.189 p.p. (0.197)	-0.096 p.p. (0.311)
Architecture	0.871 p.p.*** (0.275)	0.178 p.p. (0.154)	-0.693 p.p.** (0.306)
Area, Ethnic, Culture & Gender Studies	-0.295 p.p.** (0.203)	-0.066 p.p. (0.163)	0.229 p.p. (0.252)
Communication /Journalism	-0.499 p.p.* (0.264)	0.129 p.p. (0.254)	0.628 p.p.* (0.357)
Computer Science	0.530 p.p. (0.593)	-1.797 p.p.*** (0.686)	-2.327 p.p.*** (0.887)
Education	0.190 p.p. (0.206)	0.062 p.p. (0.157)	-0.128 p.p. (0.253)
Engineering	3.580 p.p.*** (1.083)	3.014 p.p.*** (0.846)	-0.565 p.p. (1.341)
Engineering Tech.	0.028 p.p. (0.073)	0.009 p.p. (0.063)	-0.018 p.p. (0.095)
Foreign Languages	-0.157 p.p. (0.327)	0.382 p.p. (0.260)	0.538 p.p. (0.407)
English	-0.324 p.p. (0.382)	-0.026 p.p. (0.267)	0.298 p.p. (0.453)
Liberal Arts & Sciences	-0.221 p.p. (0.188)	-0.370 p.p. (0.253)	-0.149 p.p. (0.308)
Biological /Biomedical Sci.	0.464p.p. (0.685)	1.089 p.p.** (0.535)	0.625 p.p. (0.846)
Math	-0.663 p.p.* (0.386)	0.011 p.p. (0.301)	0.673 p.p. (0.478)
Interdisciplinary Studies	0.004 p.p. (0.243)	-0.267 p.p. (0.348)	-0.271 p.p. (0.416)
Parks, Recreation, Leisure	-0.181 p.p. (0.419)	-0.729 p.p.** (0.355)	-0.548 p.p. (0.534)
Philosophy	-0.124 p.p. (0.189)	-0.019 p.p. (0.136)	0.105 p.p. (0.227)

Physical Sci.	-0.160 p.p.	0.446 p.p.	0.606 p.p.
	(0.324)	(0.283)	(0.419)
Psychology	-0.069 p.p.	0.950 p.p.**	1.020 p.p.
	(0.556)	(0.442)	(0.689)
Public Admin.	-0.109 p.p.	-0.277 p.p.	-0.168 p.p.
	(0.285)	(0.207)	(0.344)
Social Sci.	-0.180 p.p.	-1.171 p.p.	-0.991 p.p.
	(0.873)	(0.715)	(1.097)
Visual	1.108 p.p.**	1.946 p.p.***	0.838 p.p.
& Perf. Arts	(0.506)	(0.466)	(0.669)
Health	0.150 p.p.	0.006 p.p.	-0.144 p.p.
	(0.143)	(0.157)	(0.207)
Business	-3.329 p.p.***	-3.971 p.p.***	-0.642 p.p.
	(0.834)	(0.627)	(1.023)
History	-0.521 p.p.	0.660 p.p.**	1.181 p.p.**
	(0.419)	(0.291)	(0.496)

$*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .10$; Probabilities estimated using multinomial logit model of choice. Standard errors calculated using Delta-method based on bootstrapped s.e. Multinomial logit model controls for parental education, a quadratic in household income (proxied using average household income in home ZIP), high school GPA, residency status, and an indicator for if the student is white. The first cross difference (Column T=0) is $\{\hat{P}(Y^j|D=1,T=0) - \hat{P}(Y^j|D=0,T=0)\}$ while the second cross difference (Column T=1) is $\{\hat{P}(Y^j|D=1,T=1) - \hat{P}(Y^j|D=0,T=1)\}$.

APPENDIX B

Appendix to Chapter 3

B.1 Competing Risks

Competing risks models assign an outcome—in the case of mortality analysis, a death—to a particular cause, or "risk." After an observation succumbs to a particular risk, it is no longer at susceptible to any other, competing risk. With additional assumptions, researchers may also estimate the impact of risk-elimination—that is, the counterfactual mortality in a world where the risk of one (or more) causes of death drops to zero. Depending on the research subject, alternative methods are appropriate.

The accounting approach of Years of Life Lost (YLL) suggested by Andersen et al. (2013) uses cumulative incidence of deaths to measure the total impact of a particular cause of death. YLL measures the total years between two ages that are lost to a death. For example, a death at age 25 contributes 60 years of life lost between ages 25 and 85 whereas a death at age 60 contributes 15 years of life lost. This is a measure of total cost of mortality by noting that a younger death loses more potential years of life than an older death. The YLL due to cause i between ages 25 and 84 (with 5-year age bins) is given by:

$$ {}_{60}YLL^{i}_{25} = \sum_{a=25}^{80} {}_{5}YLL^{i}_{a} = \sum_{a=25}^{80} 5 \left\{ \sum_{b \leq a} [{}_{5}D^{i}_{b}] - .5\,{}_{5}D^{i}_{a} \right\} $$

Where a indexes age bin, b is all age-bins up to age bin a, and i is cause of death. Here, D is the number of lifetable deaths rather than observed deaths. The overall YLL is simply the sum of all specific-cause YLL:

$$ {}_{60}YLL_{25} = \sum_{i \in I} {}_{60}YLL_{25}^{i} $$

Where I is the exhaustive set of mutually exclusive causes of death. This is inverse of the lifetable measure of years of life lived between age 25 and 85:

$$ {}_{60}e_{25} = \sum_{a=(25-29)}^{(80-84)} {}_{5}L_{a} \; ; \; {}_{60}e_{25} + {}_{60}YLL_{25} = 60 $$

The YLL approach is a simple and intuitive approach that measures the impact of particular causes of death on overall mortality, requiring no additional assumptions beyond the standard lifetable assumptions. It allows for decomposition of changes in mortality over space, time, or different populations, allowing for measurement of changes in impact of a particular cause of death. For example, Geronimus et al. (2019) use the YLL approach to decompose the changes in educational gaps in mortality over time, showing that 29% of increased mortality differential between high- and low-educated non-Hispanic White Men is due to increases in opioid and other drug overdoses, alcoholic liver disease, and suicide—collectively "deaths of despair"—while the same group of causes contributes 2% of the increase in mortality differential for non-Hispanic Black men.

While the YLL approach is attractive for its versatility and simplicity, it does not tell us what would occur if a particular cause of death were eliminated. If a particular cause of death is eliminated, the competing risk model suggests that one of the remaining risks will cause the death of an individual who actually died from the eliminated risk. Exactly how the counterfactual mortality rates are calculated depends on the assumptions of the cause-elimination method.

Standard methods assume that the causes of death are independent of one another, allowing straightforward calculation of counterfactual mortality in the absence of the eliminated cause. Preston et al. (2001) outline three methods to calculate the counterfactual lifetable in the absence of an eliminated cause. Each method substitutes a counterfactual lifetable measure after removing one cause and re-calculates the lifetable using these counterfactual measures. The first simply substitutes remaining cause mortality as the counterfactual mortality level: ${}_{5}m_{a}^{*} = \frac{Deaths - Deaths^{i}}{Population}$, where $Deaths^{i}$ are cause i deaths and $Deaths$ are total deaths. The second two substitute counterfactual measures of probability of surviving from age a to a+5 using ${}_{5}p_{a}^{*} = \exp\{-5\frac{Deaths - Deaths^{i}}{Population}\}$, or, as suggested by

Chiang (1968), ${}_5p_a^* = {}_5p_a^{\frac{Deaths^i}{Deaths}}$. In any case, after making this substitution, the counterfactual lifetable where cause i is eliminated is computed in the standard way.

In general, the drawback of these methods is the relatively strong assumption of independence between risks. For example, if opioid use is driven in part by economic distress, as suggested by the "deaths of despair" hypothesis (Case and Deaton (2020)), then if opioid use were eliminated, the remaining "deaths of despair" causes—alcoholic liver disease, non-opioid drug overdose, and suicide—may be expected to increase by more than suggested by the standard cause-elimination models that impose independence among causes. Similarly, if opioid use is one symptom of broader high-stress coping, then the "weathering" hypothesis would similarly suggest non-independent relationships between particular causes (Geronimus et al. (2006)).

Alternative cause-elimination methods attempt to measure the relationship between particular causes, for example by using the multiple cause data from death certificates. About 80% of death certificates list more than one cause of death, only one of which is identified as the underlying cause in the National Vital Statistics (Centers for Disease Control and Prevention (2022a)). Manton et al. (1976) describe various methods that make use of multiple cause of death data to identify non-independence among causes of death in cause-elimination models. These methods, however, require that death certificates accurately capture the interrelationships between causes. For example, if prolonged opioid use negatively impacts physiological processes, leading to eventual death seemingly unrelated to the opioid use, then these methods would understate the dependence between causes. Indeed, Ray et al. (2016) identify significant excess mortality among opioid users in a sample of Tennessee Medicaid recipients, more than two-thirds of which were not coded as opioid related.

B.2 Calculating total effects of opioid use

Several papers by Preston and co-authors (Fenelon and Preston (2012); Preston et al. (2010); Glei and Preston (2020)) use an alternative method to identify deaths caused—either directly or indirectly—by a particular cause of death. They use geographic variation in a coded cause of death to identify underlying causes: for example, lung cancer deaths act as a proxy for smoking and drug overdose as a proxy for drug use. The geographic variation in these deaths is used to attribute additional, other-cause deaths to these underlying causes using nonlinear regression analysis.

Specifically, suppose opioid use is a directly related to a measure of opioid mortality. Following Preston et al. (2010), this can be stated as $M_{OD} = \beta_{\theta}^{OD}\theta$, where θ is a measure of opioid use rate, which varies with opioid usage patterns for a particular population. If opioid usage also affects risk of other causes of death, then standard proportional hazard modelling suggests:

$$\ln(M_{other}) = \sum \beta_i X_i + \beta_{\theta}^{other}\theta = \sum \beta_i X_i + \frac{\beta_{\theta}^{other}}{\beta_{\theta}^{O}D} M_{OD}$$

Note that this assumption implies that the measure of opioid usage that determines opioid mortality has a proportional impact on other-cause mortality. We cannot separately identify the individual channels of opioid usage on (i) opioid overdose deaths and (ii) other-cause deaths, but we can use opioid mortality as a proxy for opioid use to identify the scaled impact of opioid use on other-cause mortality. Note that this does recover the relationship between the opioid use mortality risk and other cause mortality risk, but this is conceptually different from the specific channel we describe whereby prolonged opioid use may cause physiological stress on other systems that cause non-opioid poisoning death.

We allow this to vary by age-group, which leads to our empirical modelling equation:

$$M_{CZ,age,year}^{other} = \exp\{\sum_{age \in A} \beta_{age}^{OD} M_{CZ,age,year}^{OD} + \phi_{year,age} + \phi_{CZ,age} + \varepsilon_{CZ,age,year}\}$$

Where $\phi_{year,age}$ are age-specific year fixed effects and $\phi_{CZ,age}$ are age-specific commuting-zone fixed effects. This estimation equation allows other cause mortality to vary flexible across time and space and assumes that the spatiotemporal variation in opioid mortality is an exogenous determinant of other-cause mortality, which can be used to identify the overall impact of opioid usage on mortality. The assumption that the impact of opioid use on other-cause mortality is constant across time and subpopulations is tested in the main text. Finally, the assumption that all other (non-spillover) deaths are independent of opioid-related deaths is captured in the assumption that $cov(\varepsilon_{CZ,age,year}, M_{CZ,age,year}^{OD}) = 0$.

Using this model ,we can estimate the share of other-cause mortality that is attributable to opioid usage.

$$ShM_{other}^{opioid} = (1 - \exp\{-\beta_{age}^{OD} M_{CZ,age,year}^{OD}\})$$

Which implies the total mortality attributable to opioid use is:

$$M_{CZ,age,year}^{\text{total opioid}} = M_{CZ,age,year}^{OD} + M_{CZ,age,year}^{other}(1 - \exp\{-\beta_{age}^{OD} M_{CZ,age,year}^{OD}\})$$

If the *remaining* other-cause mortality is independent of this total opioid-related mortality, then the standard methods for calculating counterfactual mortality figures outlined above are valid. While this assumption is still strong, it is arguably less restrictive than the assumption that deaths from all other causes are independent of opioid-coded overdose deaths. Thus, we calculate the YLL due to all opioid mortality as well as the counterfactual life-expectancy measures after eliminating all opioid related death.

B.3 Balance of pre-determined covariates and the Generalized Propensity Score

Tables B.3-B.6 report balance on the same Commuting Zone-level economic characteristics on which we estimate our generalized propensity scores. Tables B.3 and B.5 show the results of an unweighted balance test across quintiles while Tables B.4 and B.6 show the results of the re-weighted balance across quintiles. Quintiles are defined based on stratification of commuting zones by the exposure variable, which is change in opioid mortality rate between 1990 and 2009 for Tables B.3 and B.4 and change in opioid overdose rate between 2009 and 2017 for Tables B.5 and B.6.

These are blocking-based diagnostics, which show the balance between CZs in a particular quintile of the exposure variable, either the change in opioid overdose mortality from 1990-2009, $\Delta M^{OD}_{1990-2009}$, or the change in opioid overdose mortality from 2009-2017, $\Delta M^{OD}_{2009-2017}$. The unweighted balance table simply compares CZs within a quintile of the exposure variable to those not in the quintile, while the GPS-weighted balance compares CZs within a quintile of the exposure variable to those with similar generalized propensity scores. The algorithm to estimate the GPS-weighted balance is similar to that outlined in Hirano and Imbens (2004) and Bia and Mattei (2008), which is also restated in Austin (2019)

1. The continuous exposure variable, $\Delta M^{OD}_{1990-2009}$ or $\Delta M^{OD}_{2009-2017}$, is discretized into quintiles, T_1, T_2, T_3, T_4, T_5. Here, we account for population weights when discretizing our exposure variable.

2. We select the $j^{th} \in \{1, 2, 3, 4, 5\}$ exposure strata. For each stratum, we calculate the median of the exposure variable, again accounting for population weights, and label this t_j; i.e. $t_j = median(\Delta M^{OD}_{t_1-t_2} | obs. \in T_j)$

3. For each CZ, we evaluate the GPS relative to the median of the exposure strata under consideration, $GPS(t_j, X)$

4. We block the resulting $GPS(t_j, X)$ into quintiles, labeled $B_1^j, B_2^j, B_3^j, B_4^j, B_5^j$, accounting for CZ population weights

5. We create a binary exposure variable that captures whether an observation actually falls within the j^{th} exposure stratum or not, $Z_j = \mathbb{1}(obs. \in T_j)$

6. For each block, B_k^j, we calculate the difference in means for our covariates between observations that are actually in the exposure stratum and those that are not, i.e. compare $Z_i = 1$ to $Z_i = 0$. Note that this compares the difference in means for those that are *actually* in the exposure stratum to those that are not, but among those with similar *predicted* probabilities of falling within this stratum. Again, this difference in means (and corresponding standard errors) are weighted by population weights

7. For the j^{th} exposure stratum, T_j, we now have five differences in means and five standard errors (for each of our covariates), each corresponding to a blocking stratum of $GPS(t_j, X)$. We take the weighted sum of these difference in means, using the population of those who are actually in the exposure stratum and that blocking stratum (in both T_j and B_k^j) relative to the total population in the exposure stratum (in T_j) to calculate weights. Again, the calculation of these shares accounts for population weights.

8. We repeat steps 2-7 for each of the remaining exposure stratum

The results of this blocking-based diagnostic algorithm is reported as the GPS-weighted balance.

B.4 Differences between Glei and Preston (2020) and our approach

It should be noted that our estimation model differs from that of Glei and Preston (2020), who model other-cause mortality as:

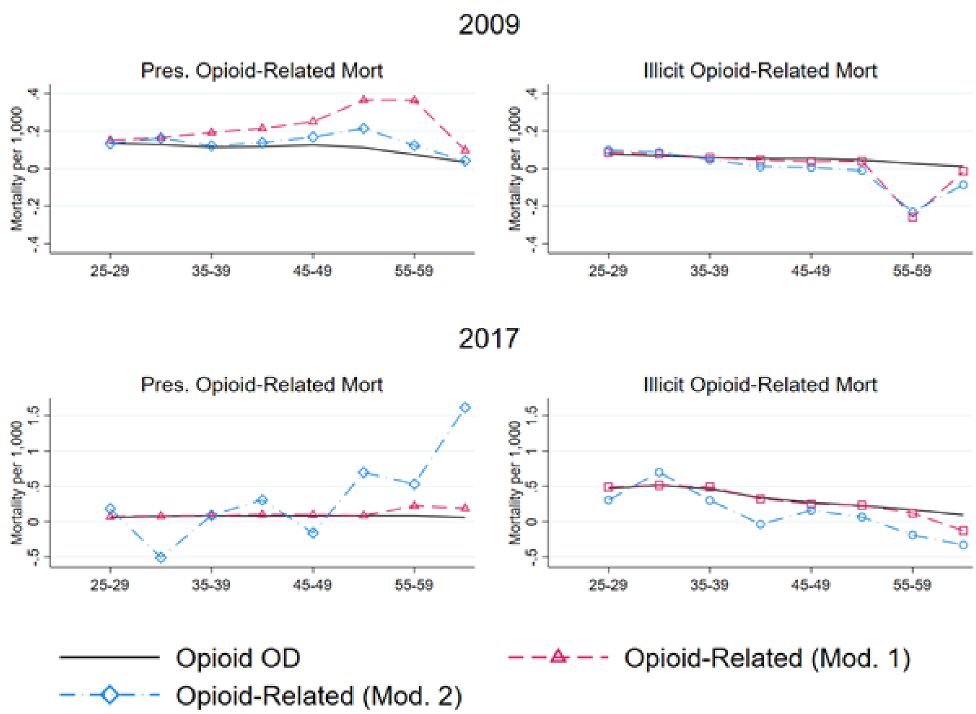
$$M^{other}_{State,age,year} = \exp\{\sum_{age \in A} \beta^{Drug}_{age} M^{DrugOD}_{State,age,year} + \phi_{age} + \phi_{State} + \beta_{time} T\}$$

Besides the differences in the fixed effects, which we allow to vary with age-bin rather than including age-bin separately, and their use of a linear time trend rather than time fixed effects, our method also differs in timeframe—we examine 1990, 2009, and 2017 while Glei and Preston examine 1999-2016—and level of observation—we examine commuting zones while Glei and Preston examine State level data.

Despite the difference in timeframe and level of observation, Tables B.9 and B.10 report that our data result in estimates that correspond with Glei and Preston's when we estimate a model that closely corresponds to theirs. Similarly, using year fixed effects rather than a time trend does not significantly impact results. The inclusion of age-specific commuting-zone and age-specific year fixed effects does decrease the magnitude of the coefficients on age-bins less than 80 for men and less than 65 for women, but the qualitative result that there are significant spillover effects of opioid usage on other-cause mortality remains.

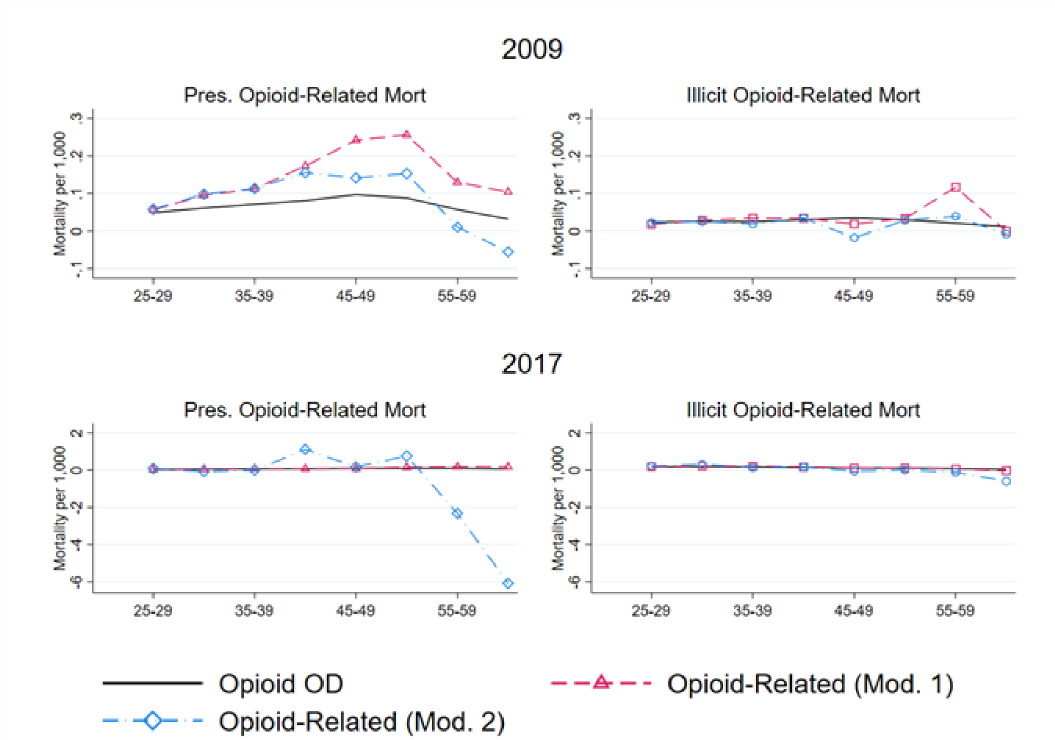
B.5 Figures and tables

Figure B.1: Observed Opioid Overdose Mortality and Model-implied Opioid-related Mortality by opioid type, Non-Hispanic White Men



Note: The opioid related estimates use the method presented in Equation 3.4.3. Model 1 uses coefficients from Equation 3.4.1 with no controls for economic conditions. Model 2 uses coefficients from Equation 3.4.1 and re-weights the specification by the stabilized inverse propensity score, where propensity scores are calculated using change in opioid mortality as the outcome and changes in economic characteristics as explanatory variables. See Equation 3.4.2 for definition of these propensity scores. Specifications include separate measures for illicit and prescription opioid mortality.

Figure B.2: Observed Opioid Overdose Mortality and Model-implied Opioid-related Mortality by opioid type, Non-Hispanic White Women



Note: The opioid related estimates use the method presented in Equation 3.4.3. Model 1 uses coefficients from Equation 3.4.1 with no controls for economic conditions. Model 2 uses coefficients from Equation 3.4.1 and re-weights the specification by the stabilized inverse propensity score, where propensity scores are calculated using change in opioid mortality as the outcome and changes in economic characteristics as explanatory variables. See Equation 3.4.2 for definition of these propensity scores. Specifications include separate measures for illicit and prescription opioid mortality.

Table B.1: Opioid mortality by age, year, sex, and opioid type (per 1,000)

A. *Non-Hispanic White Men*

	Total Opioid Mort.				Illicit Opioid Mort.			Presciption Opioid Mort.		
	1990	2000	2009	2017	2000	2009	2017	2000	2009	2017
25-29	0.007	0.048	0.211	0.527	0.023	0.077	0.471	0.025	0.134	0.056
30-34	0.013	0.053	0.196	0.590	0.023	0.069	0.514	0.031	0.128	0.076
35-39	0.018	0.070	0.173	0.543	0.031	0.059	0.462	0.039	0.114	0.081
40-44	0.011	0.087	0.171	0.422	0.038	0.054	0.342	0.049	0.117	0.080
45-49	0.005	0.078	0.180	0.353	0.033	0.054	0.272	0.045	0.126	0.081
50-54	0.003	0.039	0.157	0.309	0.016	0.045	0.225	0.023	0.112	0.084
55-59	0.002	0.018	0.101	0.252	0.007	0.027	0.169	0.010	0.074	0.083
60-64	0.001	0.008	0.046	0.150	0.003	0.011	0.093	0.005	0.034	0.058
65-69	0.001	0.005	0.020	0.069	0.002	0.004	0.038	0.003	0.015	0.03
70-74	0.002	0.003	0.011	0.026	0.001	0.002	0.01	0.003	0.009	0.016
75-79	0.001	0.003	0.006	0.014	0.001	0.001	0.004	0.002	0.005	0.010
80-84	0.001	0.005	0.007	0.011	0.003	0.001	0.002	0.002	0.005	0.009

B. *Non-Hispanic White Women*

	Total Opioid Mort.				Illicit Opioid Mort.			Presciption Opioid Mort.		
	1990	2000	2009	2017	2000	2009	2017	2000	2009	2017
25-29	0.002	0.016	0.073	0.203	0.006	0.024	0.172	0.010	0.049	0.031
30-34	0.002	0.022	0.086	0.237	0.007	0.025	0.188	0.015	0.061	0.049
35-39	0.003	0.031	0.095	0.242	0.012	0.025	0.175	0.019	0.071	0.067
40-44	0.002	0.039	0.112	0.208	0.014	0.03	0.137	0.025	0.081	0.071
45-49	0.001	0.036	0.133	0.203	0.013	0.035	0.116	0.023	0.097	0.087
50-54	0.001	0.023	0.118	0.195	0.008	0.03	0.106	0.016	0.088	0.089
55-59	0.001	0.016	0.079	0.168	0.005	0.021	0.077	0.010	0.057	0.091
60-64	0.002	0.010	0.044	0.097	0.004	0.012	0.038	0.006	0.032	0.059
65-69	0.001	0.005	0.023	0.047	0.002	0.008	0.016	0.003	0.015	0.032
70-74	0.001	0.004	0.013	0.025	0.001	0.003	0.007	0.003	0.010	0.019
75-79	0.002	0.004	0.009	0.015	0.001	0.003	0.003	0.003	0.007	0.012
80-84	0.001	0.003	0.007	0.011	0.001	0.002	0.002	0.002	0.005	0.009

Data on opioid overdose mortality by opioid type not available in 1990.

Table B.2: All other-cause mortality (Excl. opioid overdose) by age, year, and sex (per 1,000)

	A. *Non-Hispanic White Men*			
	1990	2000	2009	2017
25-29	1.424	1.134	1.196	1.245
30-34	1.707	1.347	1.317	1.504
35-39	2.162	1.837	1.703	1.894
40-44	2.786	2.695	2.467	2.443
45-49	4.113	4.093	4.018	3.710
50-54	6.557	5.989	6.088	5.871
55-59	10.968	9.419	8.789	9.216
60-64	17.521	14.912	12.384	12.892
65-69	26.246	23.042	18.911	18.196
70-74	40.973	36.464	28.678	26.976
75-79	64.356	58.173	45.860	42.016
80-84	99.242	91.849	75.870	70.164
	B. *Non-Hispanic White Women*			
	1990	2000	2009	2017
25-29	0.508	0.483	0.512	0.560
30-34	0.648	0.664	0.680	0.809
35-39	0.902	0.998	0.991	1.113
40-44	1.385	1.515	1.520	1.553
45-49	2.277	2.255	2.448	2.392
50-54	3.801	3.575	3.592	3.714
55-59	6.231	5.851	5.08	5.650
60-64	9.794	9.475	7.792	7.778
65-69	14.66	15.085	12.513	11.606
70-74	23.198	22.673	19.852	18.492
75-79	36.607	37.341	32.436	29.991
80-84	61.589	62.853	54.992	52.220

Table B.3: Balance Table for $\Delta M^{OD}_{1990-2009}$; Unweighted

	1^{st} Quin.	2^{nd} Quin.	3^{rd} Quin.	4^{th} Quin.	5^{th} Quin.
Poverty Rate	-0.0030	-0.0052	-0.0041	0.0056	0.0067
(s.e.)	(0.0081)	(0.0069)	(0.0082)	(0.0058)	(0.0068)
Assistance Rate	-0.0021	-0.0023	0.0021	0.0027	0.0000
(s.e.)	(0.0020)	(0.0026)	(0.0038)	(0.0027)	(0.0020)
Non-College	0.0490***	-0.0402**	-0.0302**	0.0035	0.0175
(s.e.)	(0.0171)	(0.0193)	(0.0150)	(0.0153)	(0.0168)
Non-Employed	0.0055	-0.0232**	-0.0101	0.0039	0.0247**
(s.e.)	(0.0080)	(0.0101)	(0.0103)	(0.0087)	(0.0098)
Work Disability	-0.0010	-0.0117**	-0.0083*	0.0045	0.0167***
(s.e.)	(0.0056)	(0.0052)	(0.0042)	(0.0043)	(0.0041)
Divorce/Sep.	-0.0144***	-0.0002	-0.0023	0.0098***	0.0070**
(s.e.)	(0.0027)	(0.0034)	(0.0038)	(0.0027)	(0.0033)
Chg. in Pop.	-0.1596***	0.0377	-0.0174	0.0574	0.0796
(s.e.)	(0.0336)	(0.0488)	(0.0474)	(0.0659)	(0.0552)
Chg. in Pov.	0.0196	-0.0158	-0.0040	0.0132	-0.0125
(s.e.)	(0.0191)	(0.0156)	(0.0164)	(0.0171)	(0.0168)
Chg. in Assist.	0.0749	0.0295	0.0568	-0.1270**	-0.0317
(s.e.)	(0.0556)	(0.0509)	(0.0609)	(0.0496)	(0.0581)
Chg. in Col.	-0.0214*	0.0247	-0.0173	0.0138	-0.0032
(s.e.)	(0.0115)	(0.0211)	(0.0144)	(0.0139)	(0.0124)
Chg. in Non-Emp.	-0.0157	0.0227*	-0.0015	0.0018	-0.0091
(s.e.)	(0.0096)	(0.0117)	(0.0082)	(0.0098)	(0.0107)
Chg. in Work Dis.	0.0119	-0.0262	-0.0048	-0.0041	0.0248*
(s.e.)	(0.0116)	(0.0182)	(0.0147)	(0.0145)	(0.0147)
Chg. in Div./Sep.	0.0424	-0.0399	-0.0234	-0.0150	0.0363**
(s.e.)	(0.0309)	(0.0326)	(0.0246)	(0.0215)	(0.0181)

$*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .1$. Rates are measured from 1990 Census data. Changes are change in natural log of each rate between 1990-2007 except for population, which uses 1980-2007.

Table B.4: Balance Table for $\Delta M^{OD}_{1990-2009}$; GPS Weighted

	1^{st} Quin.	2^{nd} Quin.	3^{rd} Quin.	4^{th} Quin.	5^{th} Quin.
Poverty Rate	0.0071	-0.0018	-0.0073	-0.0013	-0.0055
(s.e.)	(0.0149)	(0.0122)	(0.0149)	(0.0121)	(0.0126)
Assistance Rate	0.0014	-0.0021	0.0008	0.0016	-0.0023
(s.e.)	(0.0044)	(0.0040)	(0.0052)	(0.0044)	(0.0039)
Non-College	0.0661**	-0.0407	-0.0395	-0.0050	0.0066
(s.e.)	(0.0307)	(0.0355)	(0.0268)	(0.0307)	(0.0293)
Non-Employed	0.0256*	-0.0166	-0.0131	0.0003	0.0071
(s.e.)	(0.0145)	(0.0153)	(0.0165)	(0.0164)	(0.0176)
Work Disability	0.0091	-0.0059	-0.0098	-0.0009	0.0019
(s.e.)	(0.0072)	(0.0076)	(0.0064)	(0.0079)	(0.0068)
Divorce/Sep.	-0.0068	0.0039	-0.0033	0.0053	0.0000
(s.e.)	(0.0054)	(0.0058)	(0.0059)	(0.0043)	(0.0048)
Chg. in Pop.	-0.1202*	0.0937	0.0341	0.0499	0.0104
(s.e.)	(0.0688)	(0.0851)	(0.0790)	(0.1139)	(0.0866)
Chg. in Pov.	0.0120	-0.0249	0.0013	0.0078	0.0069
(s.e.)	(0.0370)	(0.0296)	(0.0298)	(0.0311)	(0.0366)
Chg. in Assist.	-0.0010	0.0054	0.0802	-0.0826	0.0647
(s.e.)	(0.1134)	(0.0957)	(0.1175)	(0.1029)	(0.1087)
Chg. in Col.	-0.0151	0.0326	-0.0100	0.0122	-0.0141
(s.e.)	(0.0232)	(0.0278)	(0.0215)	(0.0216)	(0.0184)
Chg. in Non-Emp.	-0.0202	0.0211	0.0029	-0.0017	-0.0014
(s.e.)	(0.0208)	(0.0224)	(0.0172)	(0.0187)	(0.0212)
Chg. in Work Dis.	0.0403	-0.0258	-0.0075	-0.0233	0.0104
(s.e.)	(0.0249)	(0.0269)	(0.0275)	(0.0273)	(0.0301)
Chg. in Div./Sep.	0.0428	-0.0331	-0.0240	-0.0204	0.0097
(s.e.)	(0.0515)	(0.0529)	(0.0393)	(0.0388)	(0.0326)

$*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .1$. Rates are measured from 1990 Census data. Changes are change in natural log of each rate between 1990-2007 except for population, which uses 1980-2007.

Table B.5: Balance Table for $\Delta M^{OD}_{2009-2017}$; Unweighted

	1^{st} Quin.	2^{nd} Quin.	3^{rd} Quin.	4^{th} Quin.	5^{th} Quin.
Poverty Rate	0.0225***	0.0061	0.0040	-0.0090*	-0.0246***
(s.e.)	(0.0080)	(0.0068)	(0.0074)	(0.0051)	(0.0070)
Assistance Rate	0.0037	-0.0016	-0.0018	0.0000	-0.0005
(s.e.)	(0.0024)	(0.0027)	(0.0038)	(0.0023)	(0.0023)
Non-College	-0.0154	-0.0354**	0.0133	0.0174	0.0193
(s.e.)	(0.0190)	(0.0175)	(0.0196)	(0.0138)	(0.0160)
Non-Employed	0.0179*	-0.0204**	-0.0017	-0.0018	0.0042
(s.e.)	(0.0092)	(0.0093)	(0.0120)	(0.0077)	(0.0098)
Work Disability	0.0148***	-0.0069	-0.0036	-0.0028	-0.0025
(s.e.)	(0.0050)	(0.0061)	(0.0051)	(0.0047)	(0.0050)
Divorce/Sep.	0.0088***	0.0037	-0.0038	0.0007	-0.0097***
(s.e.)	(0.0029)	(0.0033)	(0.0027)	(0.0027)	(0.0033)
Chg. in Pop.	0.0636	0.1203***	0.0761	-0.0948**	-0.1643***
(s.e.)	(0.0424)	(0.0421)	(0.0613)	(0.0426)	(0.0395)
Chg. in Pov.	-0.0366**	-0.0248	-0.0091	0.0377**	0.0338*
(s.e.)	(0.0157)	(0.0154)	(0.0157)	(0.0164)	(0.0192)
Chg. in Assist.	-0.0186	0.0280	-0.0670	-0.0036	0.0641
(s.e.)	(0.0609)	(0.0520)	(0.0482)	(0.0474)	(0.0646)
Chg. in Col.	0.0399***	0.0561***	-0.0156	-0.0312***	-0.0489***
(s.e.)	(0.0113)	(0.0197)	(0.0124)	(0.0092)	(0.0100)
Chg. in Non-Emp.	-0.0159	0.0134	0.0084	0.0064	-0.0111
(s.e.)	(0.0098)	(0.0129)	(0.0117)	(0.0078)	(0.0110)
Chg. in Work Dis.	-0.0071	-0.0412**	-0.0186	0.0331**	0.0328***
(s.e.)	(0.0121)	(0.0184)	(0.0179)	(0.0138)	(0.0114)
Chg. in Div./Sep.	0.0020	-0.0337	0.0145	-0.0033	0.0189
(s.e.)	(0.0263)	(0.0346)	(0.0275)	(0.0265)	(0.0191)

$*** = p-val < .01$; $** = p-val < .05$; $* = p-val < .1$. Rates are measured from 1990 Census data. Changes are change in natural log of each rate between 1990-2007 except for population, which uses 1980-2007.

Table B.6: Balance Table for $\Delta M^{OD}_{2009-2017}$; GPS Weighted

	1^{st} Quin.	2^{nd} Quin.	3^{rd} Quin.	4^{th} Quin.	5^{th} Quin.
Poverty Rate	0.0067	-0.0137	-0.0015	0.0047	-0.0098
(s.e.)	(0.0163)	(0.0135)	(0.0143)	(0.0091)	(0.0105)
Assistance Rate	0.0055	-0.0031	-0.0018	0.0001	-0.0004
(s.e.)	(0.0040)	(0.0047)	(0.0050)	(0.0042)	(0.0042)
Non-College	0.0104	-0.0198	0.0027	-0.0027	-0.0083
(s.e.)	(0.0396)	(0.0324)	(0.0327)	(0.0284)	(0.0276)
Non-Employed	0.0241	-0.0232	-0.0063	-0.0022	0.0003
(s.e.)	(0.0203)	(0.0184)	(0.0186)	(0.0140)	(0.0168)
Work Disability	0.0136	-0.0134	-0.0068	-0.0008	0.0002
(s.e.)	(0.0104)	(0.0095)	(0.0081)	(0.0077)	(0.0088)
Divorce/Sep.	0.0026	-0.0036	-0.0013	0.0055	-0.0046
(s.e.)	(0.0055)	(0.0044)	(0.0052)	(0.0047)	(0.0055)
Chg. in Pop.	-0.0694	0.0083	0.0898	-0.0035	-0.0404
(s.e.)	(0.0775)	(0.0633)	(0.0972)	(0.0692)	(0.0678)
Chg. in Pov.	-0.0065	0.0135	0.0033	0.0171	0.0037
(s.e.)	(0.0313)	(0.0315)	(0.0353)	(0.0332)	(0.0343)
Chg. in Assist.	0.0260	0.0900	-0.0481	-0.0300	0.0476
(s.e.)	(0.1058)	(0.0999)	(0.1078)	(0.0999)	(0.0959)
Chg. in Col.	-0.0028	0.0164	-0.0093	0.0060	-0.0072
(s.e.)	(0.0166)	(0.0167)	(0.0215)	(0.0110)	(0.0104)
Chg. in Non-Emp.	-0.0211	0.0134	0.0143	0.0097	-0.0028
(s.e.)	(0.0206)	(0.0205)	(0.0174)	(0.0146)	(0.0196)
Chg. in Work Dis.	0.0238	-0.0203	-0.0129	0.0162	0.0071
(s.e.)	(0.0241)	(0.0329)	(0.0303)	(0.0254)	(0.0209)
Chg. in Div./Sep.	0.0155	-0.0254	-0.0073	-0.0177	0.0063
(s.e.)	(0.0507)	(0.0436)	(0.0373)	(0.0455)	(0.0417)

$*** = p - val < .01$; $** = p - val < .05$; $* = p - val < .1$. Rates are measured from 1990 Census data. Changes are change in natural log of each rate between 1990-2007 except for population, which uses 1980-2007.

Table B.7: Poisson Model results for Non-Hispanic Whites, 1990-2009; Oucome: All Other-Cause Mort. (excl. opioids)

	Non-Hispanic White Men		Non-Hispanic White Women	
	Standard	IPW	Standard	IPW
Opioid mort. (per 1,000) X age 25-29	0.148	0.0776	0.143	0.228
	(0.116)	(0.121)	(0.215)	(0.277)
Opioid mort. (per 1,000) X age 30-34	0.302**	0.380*	0.765***	0.767**
	(0.152)	(0.198)	(0.269)	(0.299)
Opioid mort. (per 1,000) X age 35-39	0.511***	0.420*	0.597***	0.500**
	(0.154)	(0.255)	(0.204)	(0.224)
Opioid mort. (per 1,000) X age 40-44	0.380**	0.223	0.644***	0.522***
	(0.156)	(0.190)	(0.194)	(0.184)
Opioid mort. (per 1,000) X age 45-49	0.211**	0.0665	0.492***	0.0569
	(0.102)	(0.132)	(0.188)	(0.198)
Opioid mort. (per 1,000) X age 50-54	0.327**	0.103	0.459**	0.182
	(0.143)	(0.128)	(0.197)	(0.160)
Opioid mort. (per 1,000) X age 55-59	0.149	-0.222	0.432	-0.0550
	(0.196)	(0.165)	(0.373)	(0.231)
Opioid mort. (per 1,000) X age 60-64	0.137	-0.0603	0.188	-0.288
	(0.180)	(0.211)	(0.334)	(0.275)
Wald stat. of joint relevance (8 df)	33.21	10.95	42.01	22.85
[p-value]	[0.000]	[0.205]	[0.000]	[0.004]
Wald stat. of equality of coeff. (IPW vs. Standard)		-12.02		33.06
[p-value]		[1.000]		[0.000]
Observations (Year X Age X CZ cell)	13,798	13,798	13,174	13,174
Number of Commuting Zones	882	882	878	878

$*** = p < .01$; $** = p < .05$; $* = p < .1$. Standard errors clustered at CZ by Year level. Columns 1 and 3 (Standard) estimate Equation 3.4.1 and only weight by the average population share over the three years (1990, 2009, and 2017). Columns 2 and 4 (IPW) weight by the product of the average population share and the stabilized IPWs, estimated using Equation 3.4.2.

Table B.8: Poisson Model results for Non-Hispanic Whites, 2009-2017; Oucome: All Other-Cause Mort. (excl. opioids)

	Non-Hispanic White Men		Non-Hispanic White Women	
	Standard	IPW	Standard	IPW
Opioid mort. (per 1,000) X age 25-29	0.0504	0.0444	0.152*	0.488**
	(0.0381)	(0.0391)	(0.0906)	(0.200)
Opioid mort. (per 1,000) X age 30-34	0.00352	0.0331	0.159***	0.0665*
	(0.0241)	(0.0523)	(0.0595)	(0.0376)
Opioid mort. (per 1,000) X age 35-39	0.0310	-0.150***	0.186***	-0.235***
	(0.0240)	(0.00782)	(0.0506)	(0.0282)
Opioid mort. (per 1,000) X age 40-44	-0.00991	-0.0222	0.118**	0.0366
	(0.0243)	(0.0561)	(0.0518)	(0.198)
Opioid mort. (per 1,000) X age 45-49	-0.0140	0.0100	-0.0137	-0.159***
	(0.0202)	(0.00882)	(0.0420)	(0.0257)
Opioid mort. (per 1,000) X age 50-54	0.00407	0.0675	0.0964**	0.0487
	(0.0261)	(0.0658)	(0.0479)	(0.0383)
Opioid mort. (per 1,000) X age 55-59	0.0224	-0.105***	0.0589	0.0693
	(0.0316)	(0.00797)	(0.0489)	(0.250)
Opioid mort. (per 1,000) X age 60-64	-0.0539	0.132*	0.0766	-0.137
	(0.0466)	(0.0767)	(0.0902)	(1.751)
Wald stat. of joint relevance (8 df)	5.949	549.3	35.03	118.2
[p-value]	[0.653]	[0.000]	[0.000]	[0.000]
Wald stat of equality of coeff. (IPW vs. Standard);		-71.72		-122.2
[p-value]		[1.000]		[1.000]
Observations (Year X Age X CZ cell)	13,720	13,720	13,180	13,180
Number of Commuting Zones	882	882	878	878

$*** = p < .01$; $** = p < .05$; $* = p < .1$. Standard errors clustered at CZ by Year level. Columns 1 and 3 (Standard) estimate Equation 3.4.1 and only weight by the average population share over the three years (1990, 2009, and 2017). Columns 2 and 4 (IPW) weight by the product of the average population share and the stabilized IPWs, estimated using Equation 3.4.2.

Table B.9: Comparison to Glei and Preston (2020) model. White Men (includes all years, 1990, 2000, 2008-2017)

	Preston Model	+ Year FE	+Year FE +YearXAge FE +CZXAge FE
Opioid mort. (per 1,000) X age 25-29	0.125***	0.118***	0.0421***
	(0.0176)	(0.0180)	(0.0154)
Opioid mort. (per 1,000) X age 30-34	0.172***	0.162***	0.0759***
	(0.0132)	(0.0134)	(0.0213)
Opioid mort. (per 1,000) X age 35-39	0.196***	0.185***	0.0992***
	(0.0135)	(0.0136)	(0.0255)
Opioid mort. (per 1,000) X age 40-44	0.248***	0.239***	0.0908***
	(0.0149)	(0.0152)	(0.0202)
Opioid mort. (per 1,000) X age 45-49	0.259***	0.254***	0.0515***
	(0.0185)	(0.0192)	(0.0108)
Opioid mort. (per 1,000) X age 50-54	0.319***	0.309***	0.0450***
	(0.0133)	(0.0141)	(0.00979)
Opioid mort. (per 1,000) X age 55-59	0.330***	0.305***	0.0324***
	(0.0127)	(0.0130)	(0.00992)
Opioid mort. (per 1,000) X age 60-64	0.303***	0.254***	0.0188
	(0.0166)	(0.0157)	(0.0159)
Opioid mort. (per 1,000) X age 65-69	0.220***	0.157***	0.0468***
	(0.0180)	(0.0167)	(0.0176)
Opioid mort. (per 1,000) X age 70-74	0.142***	0.106***	0.0227
	(0.0238)	(0.0229)	(0.0215)
Opioid mort. (per 1,000) X age 75-79	0.0649***	0.0445**	0.0289
	(0.0231)	(0.0224)	(0.0178)
Opioid mort. (per 1,000) X age 80-84	-0.0236	-0.0289	0.0225
	(0.0213)	(0.0209)	(0.0154)
Age FE	X	X	X
CZ Fe	X	X	X
Linear Time Trend	X		
Year FE		X	X
CZ X Age FE			X
Year X Age FE			X
Observations (Year X Age X CZ cell)	108,996	108,996	108,697

Note: we compare these figures to those of Model I in Glei and Preston (2020)'s Table 1.

Table B.10: Comparison to Glei and Preston (2020) model. White Women (includes all years, 1990, 2000, 2008-2017)

	Preston Model	+ Year FE	+Year FE +YearXAge FE +CZXAge FE
Opioid mort. (per 1,000) X age 25-29	0.498***	0.494***	0.102***
	(0.0318)	(0.0322)	(0.0333)
Opioid mort. (per 1,000) X age 30-34	0.553***	0.552***	0.211***
	(0.0263)	(0.0264)	(0.0259)
Opioid mort. (per 1,000) X age 35-39	0.459***	0.460***	0.160***
	(0.0204)	(0.0204)	(0.0233)
Opioid mort. (per 1,000) X age 40-44	0.528***	0.536***	0.150***
	(0.0174)	(0.0174)	(0.0217)
Opioid mort. (per 1,000) X age 45-49	0.411***	0.423***	0.0823***
	(0.0161)	(0.0166)	(0.0174)
Opioid mort. (per 1,000) X age 50-54	0.415***	0.427***	0.0985***
	(0.0141)	(0.0142)	(0.0155)
Opioid mort. (per 1,000) X age 55-59	0.386***	0.388***	0.113***
	(0.0162)	(0.0154)	(0.0163)
Opioid mort. (per 1,000) X age 60-64	0.139***	0.134***	0.0486**
	(0.0194)	(0.0186)	(0.0190)
Opioid mort. (per 1,000) X age 65-69	-0.0424*	-0.0279	0.0396*
	(0.0219)	(0.0206)	(0.0202)
Opioid mort. (per 1,000) X age 70-74	0.0408*	0.0588***	0.0874***
	(0.0210)	(0.0202)	(0.0205)
Opioid mort. (per 1,000) X age 75-79	0.0204	0.0454*	0.0634***
	(0.0241)	(0.0234)	(0.0215)
Opioid mort. (per 1,000) X age 80-84	0.0212	0.0377*	0.0786***
	(0.0223)	(0.0218)	(0.0171)
Age FE	X	X	X
CZ Fe	X	X	X
Linear Time Trend	X		
Year FE		X	X
CZ X Age FE			X
Year X Age FE			X
Observations (Year X Age X CZ cell)	109,008	109,008	107,208

Note: we compare these figures to those of Model I in Glei and Preston (2020)'s Table 1.

Table B.11: Poisson Model results for Non-Hispanic Whites by Opioid Type, 1990-2009; Oucome: All Other-Cause Mort. (excl. opioids)

	Non-Hispanic White Men		Non-Hispanic White Women	
	Standard	IPW	Standard	IPW
Illicit Opioid mort.				
X age 25-29	0.102	0.231	-0.496	-0.170
	(0.194)	(0.195)	(0.475)	(0.538)
Illicit Opioid mort.				
X age 30-34	0.0957	0.199	0.222	0.0321
	(0.218)	(0.178)	(0.545)	(0.518)
Illicit Opioid mort.				
X age 35-39	-0.00575	-0.117	0.384	-0.242
	(0.366)	(0.453)	(0.557)	(0.393)
Illicit Opioid mort.				
X age 40-44	-0.0569	-0.323	0.0935	0.104
	(0.249)	(0.352)	(0.334)	(0.364)
Illicit Opioid mort.				
X age 45-49	-0.0798	-0.215	-0.188	-0.611
	(0.254)	(0.316)	(0.414)	(0.440)
Illicit Opioid mort.				
X age 50-54	-0.0203	-0.201	0.0349	-0.00891
	(0.289)	(0.279)	(0.458)	(0.457)
Illicit Opioid mort.				
X age 55-59	-1.188***	-1.075***	0.904	0.169
	(0.352)	(0.275)	(0.745)	(0.483)
Illicit Opioid mort.				
X age 60-64	-0.186	-0.714	-0.127	-0.217
	(0.493)	(0.550)	(0.369)	(0.419)
Presc. Opioid mort.				
X age 25-29	-0.00703	0.290	0.363	
	(0.126)	(0.0969)	(0.241)	(0.302)
Presc. Opioid mort.				
X age 30-34	0.224	0.204	0.849***	0.928***
	(0.143)	(0.197)	(0.256)	(0.317)
Presc. Opioid mort.				
X age 35-39	0.412***	0.0378	0.599***	0.615**
	(0.140)	(0.219)	(0.194)	(0.251)

Presc. Opioid mort. X age 40-44	0.344**	0.0732	0.774***	0.618***
	(0.160)	(0.169)	(0.224)	(0.214)
Presc. Opioid mort. X age 45-49	0.249**	0.0836	0.630***	0.189
	(0.117)	(0.132)	(0.195)	(0.242)
Presc. Opioid mort. X age 50-54	0.379***	0.151	0.544***	0.207
	(0.139)	(0.147)	(0.187)	(0.158)
Presc. Opioid mort. X age 55-59	0.453	0.0755	0.253	-0.160
	(0.285)	(0.240)	(0.333)	(0.268)
Presc. Opioid mort. X age 60-64	0.151	0.0171	0.291	-0.348
	(0.203)	(0.265)	(0.537)	(0.351)
Wald stat. of joint relevance Illicit Opioids (8 df)	12.18	21.47	3.607	2.889
[p-value]	[0.143]	[0.006]	[0.891]	[0.941]
Wald stat. of joint relevance for Pres. Opioids (8 df)	31.44	2.855	53.66	28.03
[p-value]	[0.000]	[0.943]	[0.000]	[0.000]
Wald stat. coeff. equality; Illicit and Pers. (8 df)	14.93	12.73	13.03	11.95
[p-value]	[0.061]	[0.122]	[0.111]	[0.153]
Observations	13,798	13,798	13,174	13,174
Number of Commuting Zones	882	882	878	878

$*** = p < .01$; $** = p < .05$; $* = p < .1$. Standard errors clustered at CZ by Year level. Separate measures for opioid overdose mortality by opioid type included in model. Columns 1 and 3 (Standard) estimate Equation 3.4.1 and only weight by the average population share over the three years (1990, 2009, and 2017). Columns 2 and 4 (IPW) weight by the product of the average population share and the stabilized IPWs, estimated using Equation 3.4.2.

Table B.12: Poisson Model results for Non-Hispanic Whites by Opioid Type, 2009-2017; Oucome: All Other-Cause Mort. (excl. opioids)

	Non-Hispanic White Men		Non-Hispanic White Women	
	Standard	IPW	Standard	IPW
Illicit Opioid mort. X age 25-29	0.0303	-0.262***	0.155	0.569***
	(0.0444)	(0.0100)	(0.0974)	(0.00753)
Illicit Opioid mort. X age 30-34	0.00275	0.253***	0.177***	0.908**
	(0.0239)	(0.0796)	(0.0670)	(0.362)
Illicit Opioid mort. X age 35-39	0.0332	-0.175***	0.199***	-0.158***
	(0.0251)	(0.00434)	(0.0628)	(0.00188)
Illicit Opioid mort. X age 40-44	-0.0218	-0.422***	0.155***	0.118*
	(0.0269)	(0.00805)	(0.0571)	(0.0659)
Illicit Opioid mort. X age 45-49	-0.0228	-0.111***	-0.0243	-0.581***
	(0.0215)	(0.00470)	(0.0611)	(0.0167)
Illicit Opioid mort. X age 50-54	0.00464	-0.119***	0.0319	-0.245***
	(0.0335)	(0.0185)	(0.0725)	(0.0130)
Illicit Opioid mort. X age 55-59	-0.0309	-0.227***	-0.0368	-0.429***
	(0.0363)	(0.0223)	(0.0699)	(0.0121)
Illicit Opioid mort. X age 60-64	-0.183***	-0.348***	-0.213	-2.054***
	(0.0643)	(0.0137)	(0.156)	(0.0616)
Presc. Opioid mort. X age 25-29	0.205**	1.900***	0.141	3.762***
	(0.0903)	(0.0566)	(0.198)	(0.256)
Presc. Opioid mort. X age 30-34	0.0134	-4.334***	0.0838	-2.904**
	(0.0845)	(1.385)	(0.145)	(1.313)
Presc. Opioid mort. X age 35-39	0.00439	0.0639**	0.148	-0.887***
	(0.0724)	(0.0303)	(0.111)	(0.0125)

Presc. Opioid mort.				
X age 40-44	0.0950	1.220***	0.0323	15.69
	(0.0694)	(0.0218)	(0.106)	(17.69)
Presc. Opioid mort.				
X age 45-49	0.0436	-0.780***	-0.00246	0.458***
	(0.0529)	(0.0317)	(0.0647)	(0.0228)
Presc. Opioid mort.				
X age 50-54	0.00194	1.308***	0.178**	2.256***
	(0.0556)	(0.0950)	(0.0701)	(0.0876)
Presc. Opioid mort.				
X age 55-59	0.185***	0.600***	0.178**	-3.906***
	(0.0673)	(0.121)	(0.0821)	(0.0695)
Presc. Opioid mort.				
X age 60-64	0.175**	2.220***	0.258**	-9.863***
	(0.0845)	(0.0521)	(0.112)	(0.200)
Wald stat. of joint relevance				
for Illicit Opioids (8 df)	12.83	6,412	29.43	16,706
[p-value]	[0.118]	[0.000]	[0.000]	[0.000]
Wald stat. of joint relevance				
for Pres. Opiods (8 df)	19.53	6,895	19.12	11,910
[p-value]	[0.012]	[0.000]	[0.014]	[0.000]
Wald stat. coeff. equality;				
Illicit and Pers. (8 df)	24.39	6,458	12.97	10,796
[p-value]	[0.002]	[0.000]	[0.113]	[0.000]
Observations	13,720	13,720	13,180	13,180
Number of Commuting Zones	882	882	878	878

$*** = p < .01$; $** = p < .05$; $* = p < .1$. Standard errors clustered at CZ by Year level. Separate measures for opioid overdose mortality by opioid type included in model. Columns 1 and 3 (Standard) estimate Equation 3.4.1 and only weight by the average population share over the three years (1990, 2009, and 2017). Columns 2 and 4 (IPW) weight by the product of the average population share and the stabilized IPWs, estimated using Equation 3.4.2.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abadie, A. (2005): "Semiparametric Difference-in-Differences Estimators," *Review of Economic Studies*, 72, 1–19.

Abraham, K. G., E. Filiz-Ozbay, E. Y. Ozbay, and L. J. Turner (2018): "Behavioral Effects of Student Loan Repayment Plan Options on Borrowers' Career Decisions: Theory and Experimental Evidence," *NBER Working Paper no. 24804*.

Acemoglu, D. and R. Shimer (1999): "Efficient Unemployment Insurance," *Journal of Political Economy*, 107, 893–928.

Addo, F. R. (2014): "Debt, Cohabitation, and Marriage in Young Adulthood," *Demography*, 51, 1688–1701.

Akers, E. J. (2012): "Excess Sensitivity of Labor Supply and Educational Attainment: Evidence from Variation in Student Loan Debt," Ph.D. thesis, Columbia University, New York, NY.

Altonji, J. G., P. Arcidiacono, and A. Maurel (2016): "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Manchin, and L. Woessmann, Amsterdam, The Netherlands: North-Holland, Elsevier, vol. 5, 305–396.

Andersen, P. K., V. Canudas-Romo, and N. Keiding (2013): "Cause-specific Measures of Life Years Lost," *Demographic Research*, 29, 1127–1152.

Andrews, R. J., S. A. Imberman, M. F. Lovenheim, and K. M. Stange (2022): "The Returns to College Major Choices: Average and Distributional Effects," *NBER Working Paper no. 30331*.

Austin, P. C. (2019): "Assessing Covariate Balance when using the Generalized Propensity Score with Quantitative or Continuous Exposures," *Statistical Methods in Medical Research*, 28, 1365–1377.

Avoletta, B. (2020): personal communication, April 8 and 17, 2020.

Barr, N., B. Chapman, L. Dearden, and S. Dynarski (2019): "The US College Loans System: Lessons from Australia and England," *Economics of Education Review*, 71, 32–48.

Beffy, M., D. Fougere, and A. Maurel (2012): "Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter?" *Review of Economics and Statistics*, 94, 334–347.

Bia, M. and A. Mattei (2008): "gpscore: A Stata package for the estimation of the dose-response function through adjustment for the generalized propenisty score," *Stata Journal*, 8, 354–373.

Biden, J. R. (2022): "Remarks by President Biden Announcing Student Loan Debt Relief Plan," Speech transcript; last accessed 31 May 2023.

Black, S. E., J. T. Denning, L. J. Dettling, S. Goodman, and L. J. Turner (2020): "Taking it to the Limit: Effects of Increased Student Loan Availability on Attainment, Earnings, and Financial Well-Being," *NBER Working Paper no. 27658*.

Bos, M., E. Breza, and A. Liberman (2018): "The Labor Market Effects of Credit Market Information," *Review of Financial Studies*, 31, 2005–2037.

Cadena, B. C. and B. J. Keys (2013): "Can Self-Control Explain Avoiding Free Money? Evidence from Interest-Free Loans," *Review of Economics and Statistics*, 95, 1117–1129.

Caetano, G., M. Palacios, and H. A. Patrinos (2019): "Measuring Aversion to Debt: An Experiment among Student Loan Candidates," *Journal of Family and Economic Issues*, 40, 117–131.

Callaway, B., A. Goodman-Bacon, and P. Sant'Anna (2021): "Difference-in-Differences with Continuous Treatment," Working paper; last accessed 31 May 2023.

Card, D. and A. Solis (2022): "Measuring the Effect of Student Loans on College Persistence," *Education Finance and Policy*, 17, 335–366.

Carneiro, P. and J. J. Heckman (2002): "The Evidence on Credit Constraints in Post-Secondary Schooling," *Economic Journal*, 112, 705–734.

Case, A. and A. Deaton (2015): "Rising Morbidity and Mortality in Midlife among White Non-Hispanic Americans in the 21st Century," *Proceedings of the National Academy of Sciences*, 112, 15078–15083.

——— (2017): "Mortality and Morbidity in the 21st Century," *Brookings Papers on Economic Activity*, 397–476.

——— (2020): *Deaths of Despair and the Future of Capitalism*, Princeton University Press.

Centers for Disease Control and Prevention (2022a): "Instructions for Classification of Underlying and Multiple Causes of Death–2022," Last accessed 17 August 2022.

——— (2022b): "Opioid Data Analysis and Resources: CDC's Response to the Opioid Overdose Epidemic," Last accessed 17 August 2022.

Cervantes, A., M. Creusere, R. McMillion, C. McQueen, M. Short, M. Steiner, and J. Webster (2005): "Opening the Doors to Higher Education: Perspectives on the Higher Education Act 40 Years Later," TG Research and Analytical Services research paper; last accessed 31 May 2023.

CFR (2021): "Title 36; Subtitle B; Chapter VI; Part 382–Federal Family Education Loan (FFEL) Program," .

Chapman, B. (2006): "Income Contingent Loans for Higher Education: International Reforms," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek and F. Welch, Amsterdam, The Netherlands: North-Holland, Elsevier, vol. 2, 1435–1503.

——— (2014): "Income Contingent Loans: Background," in *Income Contingent Loans: Theory, Practice, and Prospects*, ed. by B. Chapman, T. Higgins, and J. E. Stiglitz, Hampshire, England: Palgrave MacMillan, 12–28.

Charles, K. K., E. Hurst, and M. Schwartz (2018): "The Transformation of Manufacturing and the Decline in U.S. Employment," *NBER Working Paper no. 24468*.

Chaumont, G. and S. Shi (2022): "Wealth Accumulation, On-the-job Search and Inequality," *Journal of Monetary Economics*, 128, 51–71.

Chiang, C. L. (1968): *Introduction to Stochastic Processes in Biostatistics*, Wiley.

Clinton, W. J. (1993): "Radio Address to the Nation, 1 May 1993," Last accessed 31 May 2023.

Committee on Community-Based Solutions to Promote Health Equity in the United States (2017): "Communities in Action: Pathways to Health Equity," Washington, DC: National Academies Press.

Committee on Rising Midlife Mortality Rates and Socioeconomic Disparities (2021): "High and Rising Mortality among Working-Age Adults," Washington, DC: National Academies Press.

Cook, E. E. and S. Turner (2022): "Progressivity of Pricing at US Public Universities," *NBER Working Paper no. 29829*.

Cooper, D. H. and J. C. Wang (2014): "Student Loan Debt and Economic Outcomes," *Federal Reserve Bank of Boston, Current Policy Perspectives*.

Correia, S., P. G. aes, and T. Zylkin (2019): "ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects," *Stata Journal*, 20, 95–115.

Cox, D. (1962): *Renewal Theory*, Methuen Ltd.

Cox, J. C., D. Kreisman, and S. Dynarski (2020): "Designed to Fail: Effects of the Default Option and Informational Complexity on Student Loan Repayment," *Journal of Public Economics*, 192.

de Gayardon, A., C. Callender, K. Deane, and S. DesJardins (2018): "Graduate Indebtedness: Its perceived effects on behaviour and life choices–a literature review," *Centre for Global Higher Education (CGHE) working paper no. 38*.

Dean, A. and S. Kimmel (2019): "Free Trade and Opioid Overdose Death in the United States," *SSM–Population Health*, 8, 100409.

Degenhardt, L., C. Bucello, B. Mathers, C. Briegleb, H. Ali, M. Hickman, and J. McLaren (2011): "Mortality among regular or dependent users of heroin and other opioids: a systematic review and meta-analysis of cohort studies: Mortality among opioid users," *Additiction*, 106, 32–51.

Department of Education (2002): "Federal Student Loan Programs Data Book, Fiscal Years 1997-2000," Last accessed 31 May 2023.

——— (2022): "Student Loans Overview: Fiscal Year 2022 Budget Projections," Last accessed 31 May 2023.

Dillon, E. W. (2018): "Risk and Return Trade-Offs in Lifetime Earnings," *Journal of Labor Economics*, 36, 981–1021.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64, 1001–1044.

Dynarski, S. (2014): "An Economist's Perspective on Student Loans," *Brookings Economic Studies working paper*.

Federal Reserve Bank of New York, Center for Microeconomic Data (2022): "Household Debt and Credit Report," Last accessed 31 May 2023.

Fenelon, A. and S. Preston (2012): "Estimating Smoking-Attributable Mortality in the United States," *Demography*, 49, 797–818.

Field, E. (2009): "Educational Debt Burden and Career Choice: Evidence from a Financial Aid Experiment at NYU Law School," *AEJ: Applied Economics*, 1, 1–21.

Geronimus, A., M. Hicken, D. Keene, and J. Bound (2006): ""Weathering" and Age Patterns of Allostatic Load Scores among Blacks and Whites in the United States," *American Journal of Public Health*, 96, 826–833.

Geronimus, A. T., J. Bound, T. A. Waidmann, J. M. Rodriguez, and B. Timpe (2019): "Weathering, Drugs, and Whack-a-Mole: Fundamental and Proximate Causes of Widening Educational Inequity in the U.S. Life Expectancy by Sex and Race, 1990-2015," *Journal of Health and Social Behavior*, 60, 222–239.

Glei, D. and S. Preston (2020): "Estimating the Impact of Drug Use on US Mortality, 1999-2016," *PLOS ONE*, 15, e0226732.

Goldstein, E., C. Viboud, V. Charu, and M. Lipsitch (2012): "Improving the Estimation of Influenze-Related Mortality over a Seasonal Baseline," *Epidemiology*, 23, 829–838.

Gomes, T., M. Tadrous, M. Mamdani, J. Paterson, and D. Juurlink (2018): "The Burden of Opioid-Related Mortality in the United States," *JAMA Network Open*, 1, e180217.

Government Publishing Office (1994): "Federal Register," 59, thursday, December 1, 1994; last accessed 31 May 2023.

Hall, O., C. Trimble, S. Garcia, P. Entrup, M. Deaner, and J. Teater (2022): "Unintentional Drug Overdose Mortality in Years of Life Lost among Adolescents and Young People in the US from 2015 to 2019," *JAMA Pediatrics*, 176, 415.

Hampole, M. V. (2023): "Financial Frictions and Human Capital Investments," Ph.D. thesis, Northwestern University, Evanston, IL.

Heckman, J. and B. Honoré (1989): "The Identifiability of the Competing Risks Model," *Biometrika*, 76, 325–330.

Heckman, J. J., H. Ichimura, and P. Todd (1997): "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, 64.

——— (1998): "Matching as an Econometric Evaluation Estimator," *Review of Economic Studies*, 65.

Hedegaard, H., A. M. no, M. Spencer, and M. Warner (2021): "Drug Overdose Deaths in the United States, 1999-2020," National Center for Health Statistics.

Hegji, A. (2021): "Federal Student Loans Made Through the William D. Ford Federal Direct Loan Program: Terms and Conditions for Borrowers," *Congressional Research Service report: R45931*.

Herbst, D. (2023): "The Impact of Income-Driven Repayment on Student Borrower Outcomes," *AEJ: Applied Economics*, 15, 1–25.

Hershbein, B. and M. S. Kearny (2014): "Major Decisions: What Graduates Earn over their Lifetimes," Data project provided through The Hamilton Project, Washington DC.

Hirano, K. and G. Imbens (2004): "The Propensity Score with Continuous Treatments," in *Modelin and Causal Inference from Incomplete-Data Perspectives*, ed. by A. Gelman and X. Meng, Wiley, 73–84.

Hollingsworth, A., C. Ruhm, and K. Simon (2017): "Macroeconomic Conditions and Opioid Abuse," *Journal of Health Economics*, 56, 222–233.

Hong, S.-H. (2013): "Measuring the Effect of Napster on Recorded Music Sales: Difference-in-Differences Estimators under Compositional Changes," *Journal of Applied Econometrics*, 28, 297–324.

Ji, Y. (2021): "Job Search under Debt: Aggregate Implications of Student Loans," *Journal of Monetary Economics*, 117, 741–759.

Johnson, M. T. (2013): "Borrowing Constraints, College Enrollment, and Delayed Entry," *Journal of Labor Economics*, 31, 669–725.

Kaas, L. and S. Zink (2011): "Human Capital Investment with Competitive Labor Search," *European Economic Review*, 55, 520–534.

Karamcheva, N., J. Perry, and C. Yannelis (2020): "Income-Driven Repayment Plans for Student Loans," *CBO Working Paper 2020-02*.

Keane, M. P. and K. I. Wolpin (2001): "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, 42.

Kimball, M. (1990): "Precautionary Savings in the Large and Small," *Econometrica*, 58, 53–73.

Krishnan, K. and P. Wang (2019): "The Cost of Financing Education: Can Student Debt Hinder Entrepreneurship?" *Management Science*, 65, 4451–4949.

Lee, W., S. Park, D. Weinberger, D. Olson, L. Simonsen, B. Grenfell, and C. Vioud (2022): "Direct and Indirect Mortality Impact of the COVID-19 Pandemic in the United States, March 1, 2020 to January 1, 2022," *eLife*, 12, e77562.

Lochner, L. and A. Monge-Naranjo (2016): "Student Loans and Repayment: Theory, Evidence, and Policy," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Manchin, and L. Woessmann, Amsterdam, The Netherlands: North-Holland, Elsevier, vol. 5, 397–478.

Lochner, L., T. Stinebrickner, and U. Suleymanoglu (2021): "Parental Support, Savings, and Student Loan Repayment," *AEJ: Policy*, 13, 329–371.

Looney, A. (2022): "Biden's Income-Driven Repayment plan would turn student loans into untargeted grants," *Bookings Op-Ed*.

Looney, A. and C. Yannelis (2015): "A Crisis in Student Loans? How Changes in the Characteristics of Borrowers and in the Institutions They Attended Contributed to Rising Loan Defaults," *Brookings Papers on Economic Activity*.

Ma, J. and M. Pender (2021): "Trends in College Pricing and Student Aid 2021," New York: College Board research report.

Madrian, B. (2006): "The U.S. Health Care System and Labor Markets," *NBER Working Paper no. 11980*.

Manton, K., H. Tolley, and S. Poss (1976): "Life Table Techniques for Multiple-Cause Mortality," *Demography*, 13, 541–564.

Martin, S. M. (2022): "College Major Specificity, Earnings Growth, and Job Changing," Ph.D. thesis, University of Michigan, Ann Arbor, MI.

McEwen, B. (1998): "Protective and Damaging Effects of Stress Mediators," *New England Journal of Medicine*, 338, 171–179.

McEwen, B. and E. Stellar (1993): "Stress and the Individual. Mechanisms leading to Disease," *Archives of Internal Medicine*, 153, 2093–2101.

McPherson, M. S. and M. O. Schapiro (2006): "US Higher Education Finance," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek and F. Welch, Amsterdam, The Netherlands: North-Holland, Elsevier, vol. 2, 1403–1434.

Menzio, G. and S. Shi (2010): "Directed Search on the Job, Heterogeneity and Aggregate Fluctuations," *American Economic Review*, 100, 327–332.

Millet, C. M. (2003): "How Undergraduate Loan Debt Affects Application and Enrollment in Graduate or First Professional School," *Journal of Higher Education*, 74, 386–427.

Moen, E. R. (1997): "Competitive Search Equilibrium," *Journal of Political Economy*, 105, 385–411.

——— (1998): *Economica*, 65, 491–505.

Monnat, S. (2019): "The Contributions of Socioeconomic and Opioid Supply Factors to U.S. Drug Mortality Rates: Urban-Rural and Within-Rural Differences," *Journal of Rural Studies*, 68, 319–335.

——— (2022): "Demographic and Geographic Variation in Fatal Drug Overdoses in the United States, 1999-2020," *The Annals of the American Academy of Political and Social Science*, 703, 50–78.

Mueller, H. M. and C. Yannelis (2019): "The Rise in Student Loan Defaults," *Journal of Financial Economics*, 131, 1–19.

——— (2022): "Increasing Enrollment in Income-Driven Student Loan Repayment Plans: Evidence from the Navient Field Experiment," *Journal of Finance*, 77, 367–401.

Naimi, A. I., E. E. M. N. Auger, and J. S. Kaufman (2014): "Constructing Inverse Probability Weights for Continuous Exposures," *Epidemiology*, 25, 292–299.

National Center for Education Statistics (NCES) (1993, 2016, and 2019): "Baccalaureate and Beyond Survey," Last accessed 31 May 2023.

——— (2012): "Beginning Postsecondary Students (BPS) Survey," Last accessed 31 May 2023.

——— (2016): "National Postsecondary Student Aid Study (NPSAS)," Last accessed 31 May 2023.

——— (2021): "Integrated Postsecondary Education Data System (IPEDS)," Last accessed 31 May 2023.

National Center for Health Statistics (2020): "Detailed Mortality–All Counties (1990-2019)," As compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.

National Student Loan Data System (NSLDS) (2022): "Federal Student Loans Portfolio," Last accessed 31 May 2023.

Obama, B. H. (2014): "Remarks by the President on Opportunity for All: Making College More Affordable, 9 June 2014," Last accessed 31 May 2023.

Ottonello, P. (2018): "Capital Unemployment," Unpublished working paper.

Palacios, M. (2014): "Overemphasized Costs and Underemphasized Benefits of Income Contingent Financing," in *Income Contingent Loans: Theory, Practice, and Prospects*, ed. by B. Chapman, T. Higgins, and J. E. Stiglitz, Hampshire, England: Palgrave MacMillan, 207–215.

Patnaik, A., M. Wiswall, and B. Zafar (2021): "College Majors," in *The Routledge Handbook of the Economics of Education*, ed. by B. P. McCall, New York, NY: Routledge.

Paulozzi, L., C. Jones, K. A. Mack, and R. Rudd (2011): "Vital Signs: Overdoses of Prescription Opioid Pain Relievers–United States, 1999-2008," *Morbidity and Mortality Weekly Report*, 60, 1487–1492.

Pierce, J. and P. Schott (2020): "Trade Liberalization and Mortality: Evidence from U.S. Counties," *American Economic Review: Insights*, 2, 47–64.

Presidential Task Force on Education (1964): "Report of the President's Task Force," John W. Garnder, Chairman; Lyndon Baines Johnson (LBJ) Presidential Library, Task Force Reports, Box no. 1.

——— (1967): "Report of the President's Task Force," William C. Friday, Chairman; Lyndon Baines Johnson (LBJ) Presidential Library, Task Force Reports, Box no. 4.

Preston, S., D. Glei, and J. Wilmoth (2010): "A New Method for Estimating Smoking-Attributable Mortality in High-Income Countries," *International Journal of Epidemiology*, 39, 430–438.

Preston, S., P. Heuveline, and M. Guillot (2001): *Demography: Measuring and Modeling Population Processes*, Wiley-Blackwell.

Puhani, P. A. (2012): "The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear "Difference-in-Differences" Models," *Economics Letters*, 115, 85–87.

Quiggin, J. (2014): "Income Contingent Loans as a Risk Management Device," in *Income Contingent Loans: Theory, Practice, and Prospects*, ed. by B. Chapman, T. Higgins, and J. E. Stiglitz, Hampshire, England: Palgrave MacMillan, 39–48.

Ray, W., C. Chung, K. Murray, K. Hall, and C. Stein (2016): "Prescription of Long-Acting Opioids and Mortality in Patients with Chronic Noncancer Pain," *JAMA*, 315, 2415–2423.

Robins, J. (1999): "Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference," in *Statistical Models in Epidemiology: The Environment and Clinical Trials*, ed. by M. Halloran and D. Berry, Springer, The IMA Volumes in Mathematics and its Applications, vol 116, 95–133.

Robins, J., M. Hernán, and B. Brumback (2000): "Marginal Structural Models and Causal Inference in Epidemiology," *Epidemiology*, 11, 550–560.

Rothschild, M. and J. E. Stiglitz (1970): "Increasing Risk: I. A Definition," *Journal of Economic Theory*, 2, 225–243.

Rothstein, J. and C. E. Rouse (2011): "Constrained after College: Student Loans and Early-Career Occupational Choicses," *Journal of Public Economics*, 95, 149–163.

Rudd, R., L. Paulozzi, M. Bauer, R. Burleson, R. Carlson, D. Dao, and et al. (2014): "Increases in Heroin Overdose Deaths–28 States, 2010 to 2012," *Morbidity and Mortality Weekly Report*, 63, 849–854.

Ruggles, S., S. Flood, R. Goeken, M. Schouweiler, and M. Sobek (2022): "IPUMS USA: Version 12.0 [dataset]," .

Sallie Mae (2021): "How America Pays for College," Research report, last accessed 31 May 2023.

Sant'Anna, P. H. and J. Zhao (2020): "Doubly Robust Difference-in-Differences Estimators," *Journal of Econometrics*, 219, 101–122.

Seeman, T., B. McEwen, J. Rowe, and B. Singer (2001): "Allostatic Load as a Marker of Cumulative Biological Risk: MacArthur Studies of Successful Aging," *Proceedings of the National Academy of Sciences*, 98, 4770–4775.

Seltzer, N. (2020): "The Economic Underpinnings of the Drug Epidemic," *SSM-Population Health*, 12, 100679.

Shimer, R. (1996): "Contracts in Frictional Labor Markets," Unpublished manuscript.

Shireman, R. (2017): "Learn Now, Pay Later: A History of Income-Contingent Student Loans in the United States," *The Annals of the American Academy of Political and Social Science*, 672, 184–201.

Slavova, S., D. O'Brien, K. Creppage, D. Dao, A. Fondario, E. Haile, B. Hume, T. Largo, C. Nguyen, J. Savel, D. Wright, and Members of the Council of State and Territorial Epidemiologists Overdose Subcommittee (2015): "Drug Overdose Deaths: Let's Get Specific," *Public Health Reports*, 130, 339–342.

Solis, A. (2012): "Credit Access and College Enrollment," *Uppsala University Department of Economics Working Paper no. 2013:12*.

Spencer, M., A. M. no, and M. Warner (2022): "Drug Overdose Deaths in the United States, 2001-2021," *NCHS Data Brief no. 457*.

Stiglitz, J. E. (2014): "Remarks on Income Contingent Loans: How Effective Can They be at Mitigating Risk?" in *Income Contingent Loans: Theory, Practice, and Prospects*, ed. by B. Chapman, T. Higgins, and J. E. Stiglitz, Hampshire, England: Palgrave MacMillan, 31–38.

Taylor, A. N. and D. J. Sheffner (2016): "Oh, What a Relief it (Sometimes) Is: An Analysis of Chapter 7 Bankruptcy Petitions to Discharge Student Loans," *Stanford Law and Policy Review*, 27, 295–336.

Thomson, J., B. Butts, S. Camara, E. Rasnick, C. Brokamp, C. Heyd, R. Steuart, S. Callahan, S. Taylor, and A. Beck (2022): "Neighborhood Socioeconomic Deprivation and Health Care Utilization of Medically Complex Children," *Pediatrics*, 149, e2021052592.

Tsiatis, A. (1975): "A Nonidentifiability Aspect of the Problem of Competing Risks," *Proceedings of the National Academy of Sciences*, 72, 20–22.

USDA ERS–Documentation (n.d.): Last accessed 11 September 2018.

Williams, D. and C. Collins (1995): "U.S. Socioeconomic and Racial Differences in Health: Patterns and Explanations," *Annual Review of Sociology*, 21, 349–386.

Wiswall, M. and B. Zafar (2015): "Determinants of College Major Choice: Identification using an Information Experiment," *Review of Economic Studies*, 82, 791–824.

Wooldridge, J. (1999): "Distribution-free Estimation of Some Nonlinear Panel Data Models," *Journal of Econometrics*, 90, 77–97.

Wright, R., P. Kircher, B. Julîen, and V. Guerrieri (2021): "Directed Search: A Guided Tour," *Journal of Economic Literature*, 59, 90–148.

Zafar, B. (2013): "College Major Choice and the Gender Gap," *Journal of Human Resources*, 48, 545–595.