Financial Frictions and Human Capital Investments

Menaka V. Hampole*

November 26, 2022

CLICK HERE FOR MOST RECENT VERSION

Abstract

How do financial frictions affect the type of human capital investments that students make in college? To study this question, I build a novel dataset covering more than 700,000 U.S. students, merging commencement records with address histories, credit bureau records, and professional resumes. I document that students trade off initial earnings against lifetime earnings when choosing college majors and that students from low-income families are more likely to choose majors associated with higher initial earnings but lower lifetime earnings. I provide causal estimates of how student debt affects this trade-off using the staggered implementation of universal no-loan policies across 22 universities from 2001 to 2019. I find that students who are required to take on more student loans to finance their education choose majors with higher initial earnings but lower lifetime earnings. Furthermore, student debt affects students differentially depending on their family backgrounds: Students from low-income families display greater sensitivity to changes in student debt. Finally, I show that differences in student debt amounts lead to different job profiles and earnings later in life. Combined, these findings highlight the role of financial frictions in human capital investments and subsequent labor market trajectories.

Keywords: student debt, human capital, labor markets
JEL Codes: D14, I22, J24

*Northwestern University, Kellogg School of Management. E-mail: menaka.hampole@kellogg.northwestern.edu. I am particularly indebted to Anthony DeFusco, Janice Eberly, Carola Frydman, Gregor Matvos, Matthew Notowidigdo, and Dimitris Papanikolaou for extensive advice and guidance. For thoughtful suggestions, I would like to thank Julie Agnew, Scott Baker, Simcha Barkai, Efraim Benmelech, Zach Bleemer, Pawel Janas, Adam Jørring, Arvind Krishnamurthy, Annamaria Lusardi, David Matsa, Konstantin Milbradt, Olivia S. Mitchell, Andrii Parkhomenko (discussant), Paolo Paruolo (discussant), Paola Sapienza, Antoinette Schoar, Paolo Sodini, Adriana Troiano, Joanna Venator, and Constantine Yannelis (discussant), as well as numerous seminar participants at Northwestern University, 2020 Stanford GSB Rising Scholars Conference, 2021 Macro Finance Society, 2022 USC Marshall Ph.D. Conference in Finance, 2022 ICMM Workshop on Consumer Financial Management, 7th IZA Workshop on the Economics of Education, and 2022 CEPR European Conference on Household Finance. I gratefully acknowledge funding from the National Science Foundation (Award No. SES-2049909), the NBER Consumer Finance Fellowship, the Russell Sage Foundation (Grant No. 1905-15785), the Peter G. Peterson Foundation, the Financial Institutions and Markets Center, and the Kellogg School of Management.

1 Introduction

Human capital investments are some of the most important decisions individuals make. This includes both the decision of whether to invest in human capital and the specific *type* of human capital to invest in. As the U.S. economy has shifted from manufacturing to services, the return to specific types of human capital has risen (Autor, 2014; Deming, 2017). One of the primary processes through which a young person invests in specific types of human capital is through their choice of college major (Hemelt et al., 2021), and there are large variations in earnings across different college majors. In fact, differences in earnings across college majors are often much greater than the average earnings gap between high school and college graduates.[1]

At the same time, investing in human capital is increasingly expensive. Since 1980, tuition for four-year college has risen at five times the rate of inflation, and the majority of students rely on student loans to finance their investment (Dynarski et al., 2018). This rising cost presents financial barriers, especially for students from low-income families. Although the federal government offers billions of dollars of aid each year, in the form of both unsubsidized and subsidized loans—which has helped increase college attendance for students from low-income families—such students still tend to earn considerably less, later in life than students from higher-income families (Bartik and Hershbein, 2018).[2] To design better policies for allocation of student loans and financial aid we need to ask: How do financial frictions affect the *type* of human capital investments that students make? And how do the effects vary across the income distribution?

This paper sheds new light on these questions by studying how reliance on student loans affects a student's choice of college major, and thus their future earnings. I proceed in four steps. First, using a novel hand-collected dataset, I document that the choice of college major involves an intertemporal trade-off: Students trade off majors with higher initial earnings and lower lifetime earnings against majors with lower initial earnings and higher lifetime earnings. Second, I investigate how family income correlates with major choices along this trade-off. Specifically, I document that students from low-income families are more likely than students from high-income families to choose majors with higher initial earnings and lower lifetime earnings. Third, I provide new causal estimates of how financial frictions affect this trade-off. Using a quasi-experimental source of variation in the availability of non-debt

[1] See, for example, Arcidiacono (2004a), Hamermesh and Donald (2008), Altonji et al. (2012), Andrews et al. (2017), and Andrews and Stange (2019).

[2] Bachelor's degree holders from low-income backgrounds start their careers earning about two-thirds as much as those from higher-income backgrounds, but this ratio declines to one-half by mid-career (Hershbein, 2016).

student aid, I show that students who are required to take less debt to finance their education choose majors with lower initial earnings but higher lifetime earnings. Finally, I estimate that this effect is strongest for students from low-income families: no-loan policies, which eliminate the need to take on student loans, have a muted effect on major choices of wealthy students, but strongly affect those of low- and middle-income students, leading them to choose majors with lower initial earnings and higher lifetime earnings.

To understand the mechanism being explored, consider a simple consumption-savings problem in which a student chooses a major as well as a level of consumption in each period. The student faces a borrowing constraint, as they cannot borrow against future income. They also face mandatory debt repayments after college. There are two majors: major 1, which is associated with a high lifetime income but a low initial income, and major 2, which is associated with a low lifetime income but a high initial income. Therefore, major choice involves a trade-off between initial income and lifetime income. If the student has high initial wealth, they will behave like a permanent income agent and choose major 1. However, if the student has low initial wealth, they will behave myopically, as they cannot borrow against future income, and instead select major 2. This illustrates how financial frictions may lead to variations in human capital decisions across the income distribution.

For my analysis, I have combined four types of microdata to build a novel dataset that tracks students from college through their career paths. First, I hand-collect student-level data from university commencement programs to create a roster of over 700,000 students, covering 50 schools and 24 graduating classes.[3] Importantly, the commencement programs list each graduating student's major, allowing me to measure how major choices within an institution shift over time; this forms the basis for my key outcome variables. Second, for each student on the roster, I collect publicly available resumes LinkedIn, Doximity, and personal webpages, which let me link individuals to eventual occupational outcomes. Third, I merged the data with United States (USPS) records aggregated by a third-party provider to construct a history of postal addresses for each student and their household of origin (including their parents). Fourth, for each student and their parents, I obtained account-level credit bureau data from Experian PLC (a major credit bureau).

The resulting panel dataset allows me to identify the effects of student debt across several outcome variables while controlling for observable characteristics. Additionally, it also allows me to decompose the effects of student debt according to parental socioeconomic status, as measured by parents' credit bureau records and address-level characteristics. I supplement

[3]For comparison, Rothstein and Rouse (2011) analyze data from just one school, comprising 9,000 students; Scott-Clayton and Zafar (2019) study the recipients of West Virginia's PROMISE scholarship (a total of 30,000 students); Bettinger et al. (2019) study recipients of California's Cal Grants (roughly 50,000 students); and Chakrabarti et al. (2022) analyze a random sample of 50,000 students.

the dataset with data from the American Community Survey (ACS) on the future earnings profiles associated with various choices of college major.

In the first part of the paper, I document two stylized facts. The first is that, when choosing their college major, students trade off initial earnings against lifetime earnings. I demonstrate this as follows. Using publicly available data from the ACS, I residualize annual earnings for observable characteristics in order to construct earnings characteristics of college majors. I focus on four characteristics of each major: initial earnings out of college, average lifetime earnings, dispersion in earnings, and earnings growth evaluated at age 22. I observe that majors associated with high lifetime earnings (such as biology, the physical sciences, and the social sciences) have low initial earnings, while majors associated with high initial earnings (such as education, communication, and nursing) have low lifetime earnings.

Second, I document that students' major choices (and the associated trade-off between initial and lifetime earnings) are related to their families' socioeconomic status. On average, students from low-income, low-credit score families choose majors with higher initial wages but lower earnings growth, than students from high-income, high credit-score families. This heterogeneity in choice of wage profiles could be explained by several factors, including differences in college preparedness or differential exposure to specific types of majors and occupations. Therefore, to further understand the impact of financial frictions on major choice, I next turn to a causal framework.

In the main part of the analysis, I estimate the causal impact of student loans on college major choice. I do this by exploiting the staggered implementation of a specific type of financial aid policy called a *universal no-loan policy* (UNLP), which was adopted by 22 universities in the U.S. between 2001 and 2019. Pre-policy, the difference between how much a household can pay (effective family contribution) and cost of attending the university was financed with a combination of student loans and grants. When the policy was phased in, the difference between the cost of attending the university and the effective family contribution was replaced with expanded no-strings-attached grants. In the analysis, I follow a two-pronged approach. First, I exploit the staggered implementation of the policies within the sample of schools that eventually implemented the policy, across cohorts that were already enrolled at the time of the implementation to conduct a two-way fixed effect analysis. I only include students already enrolled at the time of the policy to mitigate selection bias due to the policy implementation. Second, I use a difference-in-difference strategy to compare students attending treated schools to a set of control schools that fall within the top 50 schools from US News, but which did not implement a UNLP during my sample period.

My research design has four empirical advantages. First, the staggered implementation of UNLPs over 18 years alleviates the risk of contemporaneous shocks. Second, UNLPs greatly

reduce student loan amounts, providing a strong first stage.[4] Third, UNLPs are implemented for a broad sample of students at an early stage of life, before large human capital decisions are made. My study is relevant to current policy proposals on free tuition, income share agreements, and expansion of financial aid for colleges.[5] Finally, my data sample includes more than 700,000 students, which allows me to obtain reliable estimates of how the effects of UNLPs vary across sub-samples.[6]

In the first stage, I find that the policy leads to a decline in students who take student loans by 38 percent (or 12 percentage points), while the average amount of student debt at graduation falls by between $5,000 (for students enrolled prior to the announcement) and $9,000 (for students who enrolled after the policy). This effect is large and is immediately present in the year of the policy implementation. Furthermore, the policy lead to a shift in students' major choices and the associated annual earnings: initial post-college annual earnings are between $300 and $600 lower. The implied earnings (implied through the choice of major) remain lower in their 20's and reverse in by age 32. From 32 and onwards, the implied earnings are higher, and in the late 30's and 40's, they are about $2,000 per year higher. Averaged over a lifetime, the annual earnings associated with students' post-UNLP major choices are $1,200 per year. higher than those associated with pre-UNLP major choices.

Finally, to provide a deeper perspective on how financial frictions affect human capital investments, I investigate heterogeneity in the effects of UNLPs across the income and credit score distribution. In particular, I show that UNLPs have a larger impact on major choices of students from low- and middle-income backgrounds and students whose parents have low credit scores: Students from families with below-median income or students whose parents have below-median credits scores see an effect that is about twice as large as the effect for students from families with an above-median income or an above-median credit score.

[4]For example, Rothstein and Rouse (2011) find that upon implementation of a UNLP at one university, the typical amount of student loan debt at graduation fell by over $11,000, from over $15,000 before the UNLP was implemented to $4,000 afterward.

[5]In contrast, Di Maggio et al. (2019) study low-income individuals (with average monthly income $2000) whose student debt is expunged at a later age (age 35 on average) (Di Maggio et al., 2019, p. 3, p. 49).

[6]For example, one closely related paper is Chapman (2016), which studies the impact of financial aid on earnings and reports the findings by bins of family income. However, the author does not find any statistical differences across the family income distribution. As mentioned in the paper, this is likely due to the small sample size and the fact that the income distribution in the sample differs significantly from the national average.

1.1 Related Literature

My paper builds on three main strands of literature. The first concerns the impact of financial frictions on *levels* of investment in human capital. Prior research has demonstrated that financial frictions and financial aid play a key role in college enrollment (Dynarski, 2003; Black et al., 2020), college completion (Denning et al., 2019a; Castleman and Long, 2016; Bettinger et al., 2019), and graduate school enrollment (Scott-Clayton and Zafar, 2019; Chakrabarti et al., 2022). For example, Gallagher et al. (2022) finds that individuals who are liquidity constrained following a natural disaster are less likely to enroll in college. This paper builds on these results by exploring the role of financial friction on the type of human capital investments that students make while in college, and hence their earnings. Specifically, I find that reducing the need to finance college with student loans has a material impact on major choice in college, especially for low-income students.

Second, my findings build on a large empirical literature on various aspects of the relationship between student loans and earnings.[7] There is a growing body of academic work studying the effects of student debt on labor market outcomes such as earnings (Minicozzi, 2005; Rothstein and Rouse, 2011; Di Maggio et al., 2019). However, the empirical evidence has so far been mixed and contradictory, and, as noted by Yannelis and Tracey (2022) in their recent review of the literature, there is significant heterogeneity in outcomes. For example, while some studies have found that lower student debt is associated with *higher* earnings, others have found it to be associated with *lower* earnings. Papers that have found that lower student debt is associated with higher earnings include Weidner (2016); Gervais and Ziebarth (2017); Di Maggio et al. (2019); Bettinger et al. (2019), and Scott-Clayton and Zafar (2019). Papers that have found it to be associated with lower earnings include Minicozzi (2005); Field (2009); Rothstein and Rouse (2011); Chapman (2016); Daniels and Smythe (2019); Denning et al. (2019a), and Luo and Mongey (2019). Papers that have found little to no effects on earnings include Goodman et al. (2021).

This wide range of results highlights a key challenge facing researchers. Human capital investments often involve an intertemporal trade-off: Investments in human capital cost both time and money and trades lower current earnings for higher expected future earnings. However, the prior literature typically estimates the effect of student loans on an outcome, such as earnings, at a *single point in time*, abstracting away from the intertemporal trade-off. This means that if, for example, varying levels of student debt cause students to choose

[7] Other studies have examined the effect of student debt on school enrollment and drop-out rates (Bettinger, 2004; Chakrabarti et al., 2022; Folch and Mazzone, 2022), marriage (Gicheva, 2016), business creation (Krishnan and Wang, 2018; Krishnan and Wang, 2019), home ownership (Houle and Berger, 2015; Goodman et al., 2021; Folch and Mazzone, 2022), and non-wage job amenities (Luo and Mongey, 2019).

different career trajectories with different earnings paths, then estimates or earnings will vary greatly depending on when they are measured. The wide range of estimates in the existing literature may also partly be due to the inherent empirical challenges associated with causal inference. For example, several papers rely on surveys with few observations, (e.g. the Survey of Consumer Finances, the Baccalaureate and Beyond Longitudinal Study, and the National Longitudinal Survey of Youth), study single-year policy changes that might be confounded by contemporaneous macroeconomic shocks, or restrict their attention to low-income or low-credit-score individuals. My paper helps explain the apparent inconsistencies in the previous literature by considering earnings growth across the entire career trajectory, as implied by major choice. Finally, I identify a previously unexplored mechanism for variables in earnings: major choice.[8]

Third, my paper contributes to the rich literature studying various determinants of college major choice. These include but are not limited to role of ability (Turner and Bowen, 1999; Altonji, 1993; Arcidiacono, 2004b; Stinebrickner and Stinebrickner, 2008), earnings uncertainty (Nielsen and Vissing-jorgensen, 2006; Bonin et al., 2007; Dillon, 2018; Saks and Shore, 2005), earnings expectations (Arcidiacono, 2004b; Stange, 2012; Stinebrickner and Stinebrickner, 2014; Zafar, 2011; Wiswall and Zafar, 2014, 2017), non-pecuniary benefits (Zafar, 2013; Boneva and Rauh, 2017), and natural disasters (Gallagher et al., 2022). I add to this literature by showing how major choice is related not just to mean earnings, but to an intertemporal trade-off between initial and later earnings. Moreover, I show that financial frictions (in the form of student debt) affect major choices, especially for low- and middle-income students—an effect that has previously been difficult to show. Prior work has shown that on average, low-income students choose majors associated with low mean earnings, such as social work, nursing, and education (Arpita et al., 2020); I note that these majors are also associated with higher initial earnings and are more likely to be chosen by students who are incurring more student debt.

The remainder of this paper is organized as follows. Section 2 I discuss the data sources and methods used in the analysis. In Section 3 I provide a descriptive analysis of the data and document two new stylized facts. Section 4 outlines the empirical research design; it includes institutional background on U.S. financial aid and UNLPs. Section 5 presents estimates of how UNLPS affect educational and occupational choice. Finally, Section 6 contains concluding remarks.

[8]Many studies find that increases in earnings come from higher completion rates on the extensive margin of students and do not focus on major choice (Denning et al., 2019b).

2 Data

My analysis relies mainly on data that was hand-collected for this paper, supplemented by publicly available data. I combine novel data from five sources: (i) university commencement records, (ii) United States Postal Service (USPS) records, (iii) credit bureau records, (iv) LinkedIn, Doximity, and personal website resume data, and (v) survey data.

2.1 Data Sources

In this section, I briefly describe the data sources, how each dataset is used, and how they were merged together. I provide more complete details on the merge and matching rates in the data appendix section B.

2.1.1 Commencement Programs

My dataset is structured around commencement programs obtained from university registrars or special collections archives. to create the dataset, I collected, digitized, and standardized commencement programs from 50 universities spanning 25 graduating cohorts from (1998 to 2022). Each university produces at least one commencement program per academic year.[9] Each commencement program includes a list of graduating bachelor's students by full name, degree, the field of study, hometown, and honors earned. The commencement programs are necessary for two purposes: to create a roster of students to track over time and to construct my main outcome variable of interest, major choice. An example commencement program can be found in Figure C1, in the Appendix.

2.1.2 Proprietary USPS Data

To decompose the effects of UNLPs according to parental wealth, I merge the roster of students from the commencement programs with postal address data from two proprietary data vendors, Intelius and BeenVerified, to create a history of geocoded addresses for each student.[10]

The merge is based on each student's full name, approximate age, hometown (as provided by the commencement program), and dates of residence in his or her hometown. To verify the current locations of international students, I use the locations in their publicly available

[9] Some universities produce one commencement program per academic session (i.e., quarter or semester).

[10] These databases contain data on U.S. individuals, including personal property records, bankruptcy records, address histories, and potential relatives. In addition, I use LexisNexis address finder to randomly check Intelius and BeenVerified results. See Cronqvist et al. (2012) and Pool et al. (2015) for details on this type of data.

resumes. For each domestic student, I identify the postal address where the student lived in the year before she started college. I create a *household* consisting of all individuals over the age of 18 who lived at this address in that year, classifying household members who are at least 15 years older than the student as *parents*.

2.1.3 Credit Bureau Data

My third unique data source is credit bureau records from Experian PLC, one of the three main credit bureaus. The data from Experian allows me to construct one of the key variables, amount of student loans taken on by each student during their undergraduate degrees. Experian provides information on households' balance sheets, giving the yearly history of all the borrowers' loans from 2004 to 2021 (both included). These include auto loans, mortgages, home equity lines of credit, student loans, and credit card balances (revolving). The data includes granular information about the main features of these loans, such as date opened, account type, credit limits, monthly scheduled payment, balance, and performance history. Versions of this data have been employed in other papers studying households' financial decisions. However, my proprietary version is unique in a few respects.

First, to match the borrowers (students) from the commencement programs with their credit records, Experian uses both the borrowers' names and their locations (collected from the USPS records described above). They then provide me with a matched, anonymized sample of students who attended universities that implemented a UNLP. Additionally, I obtain a 5% random sample of student's that attended comparable universities that did not implement a UNLP.

Second, using the households identified from the USPS records, Experian also matches each student's parents to their credit bureau records. This data is crucial for the analysis of heterogeneity in major choices based on family socioeconomic status. Additionally, I obtain credit bureau records of parents of student's from the 5% random sample of student's that attended comparable universities that did not implement a UNLP.

Lastly, I also obtain a random 1% sample of the U.S. population to show how my sample of students compares to the general population.

2.1.4 Online Resume Data

To analyze the effect of student debt on labor market outcomes, for each student in the commencement programs I have obtained an individual public resume from LinkedIn, Doximity, and publicly available resumes from professional websites. LinkedIn and Doximity an employment-oriented online platform where job seekers maintain profiles and employers

post jobs.[11] Each resume contains three types of information: (1) educational history, including school name, degree, major, graduation date, and honors (such as Phi Beta Kappa membership or Latin honors); (2) job history, including firm names, position titles, years of employment, locations, and descriptions of tasks; and (3) a list of technical and soft skills and areas of expertise. Each position title is assigned a 2010 Standard Occupational Classification (SOC) code using Occupational Information Network (O*NET), a labor market database maintained by the Bureau of Labor Statistics (BLS).

A concern with this data is that individuals may falsify their profiles. However, as discussed in Jeffers (2017), this is unlikely to be a prevalent problem, since the public availability of LinkedIn profiles means that individuals who make false claims about their schooling or employment are likely to be found out. A more salient concern is that of stale profiles. I address this issue in the following ways. First, I restrict the sample to individuals whose recorded employment history covers every year since they graduated from college. Second, I restrict the sample to individuals who have connections on LinkedIn. With these restrictions, I find that I am able to match, on average, 85% of each graduating class from the commencement reports to LinkedIn resume data. For more details on the match, please see data appendix B.

2.1.5 Supplemental datasets

I supplement the hand-collected data with a number of publicly available datasets.

American Community Survey (ACS): I use demographic survey data from the 2009–2019 ACS, extracted from the IPUMS 1% samples (Ruggles et al., 2017; Deming and Noray, 2018). Importantly for my study, the ACS survey includes demographic variables (e.g., race and gender), highest level of education (e.g., high school or four-year college), undergraduate field of study, occupation code, and annual earnings.

Bureau of Labor Statistics (BLS): I deflate earnings across years using the annual Consumer Price Index (CPI) provided by the Bureau of Labor Statistics (BLS).

Housing and Urban Development (HUD): I use spatial crosswalks provided by the Department of Housing and Urban Development (HUD). The crosswalks link Counties, ZIP codes, and Census Tracts.

[11] Similar data has previously been used by Lucca et al. (2014), Jeffers (2017), Bernstein et al. (2018), Krishnan and Wang (2018), and Egan et al. (2019).

Integrated Postsecondary Education Data System (IPEDS): I obtain data on individuals schools from the Integrated Postsecondary Education Data System (IPEDS). This data includes school-level variables on tuition, enrollment rates, graduation rates, SAT scores, class composition in terms of race and ethnicity, and average student loan take-up.

2.2 Merging Process

As described earlier, my baseline dataset is drawn from commencement programs. I merge the commencement program data with three supporting datasets: (1) resume data, (2) USPS records, and (3) credit bureau records.

I first merge the commencement program data with resumes. Here, for a pair of records to be considered a match, I require the name, university, and class year to be identical in both. Since women frequently change their last names upon marriage, in women's records, I require only the first names to match, and I verify possible name changes using USPS data. I verify all matches based on the student's field of study.

Next, I merge the commencement records with USPS records, as follows. From each student's full name (first, middle, and last) and approximate age (based on year of graduation), I identify possible matches in the USPS database. I then use the student's hometown (as given in the commencement program) and other location data from resumes to confirm a match to a specific USPS records. After obtaining the student's USPS record I create a household for them by identifying all individuals with the same last name who lived at the same postal address as the student at the same time. The time frame I consider is the entire period before the student begins college. This process is intended to capture the student's parents and likely siblings. I characterize as parents any household members who are at least 15 years older than the student.

Lastly, I provide Experian with the postal address histories obtained from the USPS data, which it uses to match a credit bureau record to each student and each parent.

2.3 Sample Representativeness

In order to gauge the representativeness of the sample, I compare the sample of UNLP students to a 1% sample from Experian and to the ACS as well. Additionally, for robustness, I compare the sample of UNLP students to a 5% random sample of non-UNLP students.[12] Table 2 provides summary statistics and co-variance balance tests. Panel (A) reports summary statistics at the school level. Sub-panel (i) compares the 22 UNLP schools with the

[12]Due to funding restriction, I've obtained credit bureau records for only a 5% random sample of non-UNLP students.

remaining top-50 schools that make up the control sample, and Sub-panel (ii) compares UNLP schools with all schools in the U.S. Comparing UNLP schools with non-UNLP control schools, we see that there is virtually no statistically significant difference between the two samples. The two samples are similar in terms of admission rate, test scores, student body, cost of attendance, and faculty compensation.[13] However, while UNLP schools are similar to non-UNLP control schools, they are quite different from the average U.S. university. UNLP schools have lower admissions rates, higher test scores, and higher tuition and fees. They also have a higher proportion of faculty-to-students and pay their faculty higher wages.

Table 2 Panel B reports demographic summary statistics comparing the regression sample with the full sample. The full sample includes all cohorts from both UNLP and non-UNLP schools, while the regression sample includes all cohorts from non-UNLP schools but only cohorts up until (and including) the first three treated cohorts among the UNLP schools. I restrict the regression sample to address a potential selection bias. (See Section 4.5 for a detailed discussion.) The demographic variables reported in the Panel are derived from an algorithm that probabilistically assigns gender and race. I describe this algorithm in detail in the Online Appendix Section C.V. Note that the algorithm applied to the student-level data assigns a higher fraction of students as "white" than what is officially reported in the IPEDs data. Part of this is due to the fact that the algorithm uses less granular ethnicity groups than the IPEDS data, e.g., it does not include Pacific Islanders.

In Figure C6, I plot histograms of the distribution of median household income by zip codes. Panel (A) plots the distribution of ZIP code incomes for the UNLP students' families in blue relative to a 1% random sample of the U.S. population. Panel (B) plots the distribution of zipcodes for the UNLP students' families relative to families from the random sample of students from non-UNLP control schools in gray. When comparing UNLP students to a 1% sample of US population, I find that on average, students that attend UNLP schools are wealthier than the U.S. population. Specifically, in the 1% ACS sample of U.S. population, the mean of the median-household income by zipcode is $67,243 and the median is $60,787. The 10th percentile is $38,795 and the 90th percentile is $104,467. In the UNLP school sample, the mean of the median-household income by zipcode is $85,182 and the median is $77,157. The 10th percentile is $44,924 and the 90% percentile is $136,343. In Panel B, when comparing UNLP students to a random sample of non-UNLP students, I find no difference in means. Specifically, in the 5% sample of non-UNLP schools, the mean of median-household income is $36,319 and the median is $78,641. The 10th percentile is $45,929 and the 90th

[13] This is not surprising since the control sample also includes highly selective universities, including MIT, Duke, and Caltech.

percentile is $138,220.

Similarly, in Figure C7, I plot histograms of the distribution of of Vantage scores. Panel (A) plots the distribution of Vantage Scores for the UNLP students' families relative to a 1% random sample from Experian. Similarly, I find that on average, student's from UNLP school's parents have better vantage scores. In the 1% sample of U.S. population, the mean vantage score is 699 (out of 850) and the median vantage score is 717. The 10th percentile is 535 and the 90th percentile is 817. In the UNLP school sample, the mean vantage score is 713 and the median vantage score if 732. The 10th percentile is 571 and the 90th is 813. In Panel (B), I plot the distribution of vantage scores for the UNLP parents relative to parents from a random sample of non-UNLP parents from the control schools. I find no difference. Specifically, in the 5th sample of non UNLP students parents, the mean vantage score if 715 and the median vantage score if 733. The standard deviation is 92. The 10th is 575 and the 90th percentile is 815.

2.4 Variable Construction

The main purpose of this paper is to investigate how student debt is related to college students' choice of majors and the associated future earnings. In this subsection, I detail how I calculate the expected earnings trajectory for each major.

To calculate the expected earnings trajectory, I use data on earnings by occupation and major from the 2009–2019 ACS, extracted from the IPUMS 1% samples (Ruggles et al., 2017; Deming and Noray, 2018). Importantly for my study, the ACS data includes each individual's undergraduate field of study, occupation, and annual wage. I classify fields of study according to the Classification of Instructional Programs (CIP) system developed by the National Center for Education Statistics (NCES), and occupations according to the SOC system, using the 2010 Census Bureau definitions of occupations. I include individuals who have a minimum of a bachelor's degree and are between the ages of 21 and 60, which yields a sample with 5,628,805 observations.

To quantify the effects of UNLPs on major choice, I ask how earnings differ across individuals with different college majors. The main variable of interest is total pre-tax wage and salary income—that is, money received as an employee—for the previous year (Ruggles et al., 2017).[14] To purge the data of variation unrelated to field of study, I regress the annual wage on age, age-squared, race, ethnicity, sex, and survey-year fixed effects:[15]

[14] The average pre-tax wage and salary income in the sample is $61,190, the 10th percentile has zero wage and salary income, and the 90th percentile is $126,000. Table D2 in the online appendix reports the average wage, standard deviation, and percentiles for each undergraduate field of study.

[15] Table D3 in the online appendix reports the regression coefficients. To avoid having to drop zeros, I add one dollar to the annual wage before taking the natural logarithm.

$$log(1 + wage_{it}) = \beta_1 Age_i + \beta_2 Age_i^2 + \mathbf{\Gamma} \boldsymbol{X}_i + \gamma_t + \epsilon_{it}. \tag{1}$$

Next, in order to avoid a Jensen's inequality problem, I calculate the difference between the predicted value (transformed to dollars) and the annual wage in dollars:

$$\varepsilon_{it} = wage_{it} - e^{log(\widehat{1+wage_{it}})} - 1. \tag{2}$$

Once I have the dollar residuals, I calculate the mean, μ_m, and standard deviation, σ_m, of the residuals for each major, m, separately:

$$\mu_m = \frac{1}{N_m} \sum_{i=1}^{N_m} \varepsilon_{it} \mathbb{I}_{\text{major}=m}, \tag{3}$$

$$\sigma_m = \sqrt{\frac{1}{N_m} \sum_{i=1}^{N_m} (\varepsilon_{it} - \mu_m)^2 \mathbb{I}_{\text{major}=m}}. \tag{4}$$

Table D1 reports the results for each major. I find that engineering, mathematics, business, and physical sciences have high mean residuals and high standard deviation of residuals, while fine arts, English, and education have low mean residuals and low standard deviation of residuals. Column 3 reports the coefficient of variation of the residuals, which is the standard deviation scaled by the mean.

As well as differing in mean and standard deviation, earnings across majors also differ in their life-cycle trajectories. For example, if a student majors in education or business, the earnings right out of college is high, but the wage growth is relatively low. On the other hand, a student who majors in biology and goes to medical school has a low wage initially after college but a much higher wage later on.

Figure **??** displays this relationship through a binned scatter plot of residualized annual earnings relative to age across the 10 most frequent college majors.

I first regress the annual wage and age on race, ethnicity, gender, and survey-year fixed effects, then generate the residuals from those regressions and add the sample mean of each variable back to its residuals.

To capture the trade-off between initial wage and later wage, I run the following regres-

sion:

$$log(1 + wage_{imt}) = \sum_m \beta_1^m \mathbb{I}_{\text{major}=m} Age_i + \sum_m \beta_2^m \mathbb{I}_{\text{major}=m} Age_i^2 + \mathbf{\Gamma} \boldsymbol{X}_i + \gamma_t + \varepsilon_{it}. \quad (5)$$

From the above regression, I recover the major-specific coefficients on age and age-squared. I then calculate the slope of wage with respect to age, evaluated at age 22. Table D1, Column 4, reports these slopes for each major. The slopes are positively correlated with the mean of the residuals. In other words, majors with high lifetime earnings, such as biology and engineering, also have steep earnings trajectories immediately after college.

3 Descriptive Analysis

In this section, I provide a descriptive analysis of the data, documenting two new stylized facts about the relationship between college majors, future earnings, and financial frictions. The first of these is that both expected initial earnings and expected lifetime earnings differ across majors; moreover, majors associated with high initial earnings (such as teaching and communications) lead to low lifetime earnings, while majors associated with low initial earnings (such as biology and the physical sciences). The second stylized fact is that family-level financial frictions affect students' major choices and consequently their future earnings trajectories, as implied by major choice.

3.1 The Intertemporal Trade-off in Major Choice

There are large differences in expected earnings of college majors over the life-cycle. Figure 1 illustrates these differences. This figure plots a binned scatter plot residualized annual wages relative to age (controlling for race, ethnicity, gender, and year, and adding back the sample mean of each variable back to its residuals) across the 15 most frequent college majors (using the 2-digit CIP classification). I highlight three majors, Education, Business, and Biology, with vastly different wage profiles. If a student majors in education or business, their earnings right out of college is high, but the wage growth is relatively low. We can compare this to a student who majors in biology and goes to medical school. This person has a low wage initially after college due to investment in human capital but the slope is much higher. Education has an average annual wage of $48,000, Business has an average annual wage of $68,000, while Biology has an average annual wage of $87,000. The difference between Biology and Education is higher than the equivalent difference between the average

student with a college degree and an individual with only a high school degree.[16] There are similar differences when we compare cumulative earnings and if we compare majors by their discounted present value of earnings.[17]

On the other hand, as Figure 1 shows, while biology majors have high lifetime earnings, they have lower initial earnings than business and education majors. In fact, education is associated with the highest initial earnings out of all the top 15 majors. This illustrates the first stylized fact that I document: that students trade of higher initial earnings against higher lifetime earnings when choosing majors.

To formalize this intertemporal trade-off between initial and life-time earnings, in Figure 2, I plot the average annual life-time earnings against the initial wages. In Panel (A), I plot the relationship for the top-15 majors. I find that there is a negative relationship between initial earnings out of college and mean lifetime annual earnings across majors. Specifically, students trade off initial earnings again mean annual earnings. This relationship holds for more granular levels of majors. In Panel (B), I plot the most frequent majors within the Social Sciences and find the same relationship. For example, student who study Sociology or Social Work on average have a high initial earnings but a low average life-time wage. On the other hand, students who study Political Science or Economics have on average a lower initial earnings but a higher average annual life-time earnings.

3.2 Family Background

The second stylized fact that I document is that the intertemporal trade-off is related to the financial and socioeconomic status of parents. I show that students who have parents with low credit scores, choose majors that are associated with high initial earnings out of college but low lifetime earnings. I show that this relationship holds both unconditionally and when controlling for demographic characteristics and school and year fixed effects.

Figure 3, I plot the relationship between students' major choices and their parents' credit scores. Panels (A) and (B) show binned scatter plots of the raw data, with the parents' credit scores displayed on the x-axis. On the y-axis, Panel (A) shows the initial earnings and Panel (B) the average annual lifetime earnings for the major chosen by the student. Panels (C) and (D) show binned scatter plots of residualized earnings against the parents' credit scores,

[16] The average annual wage for Biology minus Education is \$87,000 - \$48,000 = \$39,000. For College minus High School, the gap is \$64,000 - \$28,000 = \$38,000.

[17] The cumulative wage gap for Biology minus Education is \$3.6m - \$2.1m = \$1.5m. For College minus High School, the gap in cumulative wages is \$2.6m - \$1.2m = \$1.4m. The present value gap for Biology minus Education is \$1.6m - \$1.0m = \$1.6m. For College minus High School, the gap in cumulative wages is \$1.1m - \$0.5m = \$0.5m. When computing the present value of wages, I use the approach of Gourinchas and Parker (2002) and use a $\beta = 0.96$.

controlling for race, ethnicity, gender, school, and year fixed effects. Again, Panel (C) shows initial earnings and Panel (D) shows average annual lifetime earnings. In each panel, the dots represent 20 equal-sized bins based on, the variable on the x-axis, and the solid line represents a linear regression on the entire dataset. In Panels (C) and (D), the transparent bars indicate the 95% confidence intervals; with the t-statistics in parentheses, and the standard errors are clustered at the school level.

The raw correlations could potentially be driven by selection into different types of colleges. For example, Chetty et al. (2020) find that many of the nation's elite colleges have more children of the 1% than from families in the bottom 60% of family income. If students at the most selective colleges also choose higher earning majors, then selection into universities by family income could be driving my results (and not the choice of college major).

To address this issue and isolate the variation from school selection, I next turn to an analysis where I control for observable characteristics. Specifically, I control for gender, race, school and year fixed effects. The results are reported in Figure 3, Panels (C) and (D). By controlling for school and year fixed effects, I am comparing students who graduated from the same university, and I find the same strong relationship between family background and choice of college major. In other words, even conditional on attending the same college, students from low-income, low credit score families are more likely to choose low earning majors.[18]

The results are similar when I compare major choice relative to the parents' income. In Figure 4, I plot the relationship between the neighborhood that a student grew up in and the type of major that the student chose in college. Panels (A) and (B) plot binned scatter plots of the raw data, and both panels display the natural logarithm of the median income in the ZIP code where the parents' live on the x-axis. In Panel (A) the y-axis plots the initial wage for the major that the student chose, and Panel (B) plots the average lifetime wage of the major that student chose on the y-axis. Panels (C) and (D) plot binned scatter plots of residualized wages and income when controlling for race, ethnicity, gender, year, and school fixed effects. Panel (C) plot the initial wage relative to log income controlling for the fixed effects, and Panel (D) plot the average lifetime wage controlling for the fixed effects. I find that students from poorer neighborhoods, choose majors that have a high initial wage and lower mean lifetime earnings.

[18]In the Online Appendix, in Table D5, I report the full regression coefficients both with and without school fixed effects. In line with the results on selection reported by Chetty et al. (2020), we see that the coefficients are larger without school fixed effects. This indicates that wealthier parents (and parents with higher credit scores) send their children to more prestigious universities, where the students, on average, chose higher-earning majors.

My findings are consistent with descriptive evidence in Arpita et al. (2020). They show that first-generation college graduates who have the highest average parental education are more likely to choose higher return majors like physics while students who have lower average parental education choose lower return majors like education, social work, and education. However, these estimates also include institutional differences. For example, student's that attend public universities maybe fundamentally different from those that attend smaller private universities. My estimates are able to control for institutional differences but looking within university.

I have thus documented two stylized facts: First, I have documented that students trade off higher initial wage against higher life-time wage when choosing majors. College majors with higher initial wages have on average lower life-time wage and vice versa. Second, I find that this intertemporal trade-off is related to family background. Students who grew up in low-income neighborhoods or who's parents have low credit scores are more likely to choose majors with higher initial wages and lower life-time wages.

Of course, all the standard caveats of ordinary least squares regression apply. For example, the heterogeneity in the choice of college majors could be explained by several factors not related to financial frictions. These could include differences in college preparedness or differential exposure to specific types of majors and occupations. For example, it might be that students who grew up with parents who are lawyers or doctors have a preference for these fields of study, unrelated to the financial frictions that they face. For example, student's may be more likely to choose similar occupations as their parents (Xia, 2016). Therefore, to understand the causal impact of financial frictions on major choice, I next turn to an empirical framework that allows me to estimate the causal effect.

4 Empirical Design

In this section, I detail the empirical methodology that allows me to estimate the causal impact of financial frictions on major choice. I begin by describing the main benchmark regression specifications: an event study and a difference-in-differences model. Next, I employ two alternative models to assess the robustness of the empirical results: a regression model with an additional control group, and a regression model that adjusts for potential treatment heterogeneity across cohorts. Finally, I describe the model that tests whether the effect of student debt varies with family income.

4.1 Institutional Background

Over the past 40 years, student loans have played an increasingly important role in funding college attendance for students in the U.S. This is largely due to the unprecedented rise in college costs since the 1980s (average tuition prices for four-year colleges have increased at five times the rate of the Consumer Price Index), as well as shifts in federal spending policies. From the 1980s to the mid-2000s, because of concerns over federal budget deficits, Congress did not consistently increase funding for the Federal Pell Grant Program (the primary source of federal financial aid for low-income college students) (Mettler, 2014). As a result, the purchasing power of Pell Grants decreased; the current maximum award ($6,895) covers only a third of the average annual price of a public four-year college, or a fifth of the average annual price of a private four-year college. Since roughly 1980, loans have been replacing grants as the dominant form of federal financial aid (Cervantes et al., 2005; Mettler, 2014), as Congress has prioritized making college affordable for middle- and upper-income students over making it accessible to low-income students (Hearn, 2015).

The main question studied in this paper is whether the shift from grants to loans has created constraints on students' college major choices and labor market outcomes. As described later in this section, I am able to analyze the causal impact of student debt thanks to expanded financial aid programs recently adopted at a number of universities, which aim to eliminate or significantly reduce the need for student loans.[19] Between 1998 and 2022, over 75 public and private universities adopted such programs, which are commonly referred to as no-loan programs (NLPs).[20]

In general, universities offer financial aid to students to cover the difference between the cost of attendance (COA) and the expected family contribution (EFC) (i.e., the amount that the student's family can afford to pay for college). The COA includes tuition and fees, room and board, books, supplies, transportation, loan fees, and other school-related expenses. It may also include costs related to child and dependent care, a student's disabilities, or study-abroad programs. The EFC is calculated from family assets (the value of savings or investment accounts, home value, business assets, household income, etc.), family size, and the number of dependent children enrolled in college. This information is collected on the Free Application for Federal Student Aid (FAFSA), which students must complete shortly before college enrollment in order to receive financial aid from the federal government (as well as from many universities). The difference between the COA and EFC, often called

[19] Broadly, the increase in financial aid has been made available through five types of policies: student loan eliminations, student loan caps, parental contribution eliminations, tuition waivers, and Pell Grant matches.

[20] As described later, the term "no-loan program" is used to describe the package of recruitment, enrollment, and financial services provided by these campuses. See Lips (2011) for a discussion of various strategies for implementing no-loan programs.

the financial need gap, is covered in part by the university through scholarships, grants, and work-study programs (insofar as the student qualifies), and in part by the student through federal and/or private student loans.

Using the implementation of NLPs as a research design builds on the work by Rothstein and Rouse (2011) and Krishnan and Wang (2018). Rothstein and Rouse (2011) analyze the effect of the one-off implementation of a UNLP at Princeton University[21] and Krishnan and Wang (2018) study the effect of no-loan policies (both universal NLPs and those targeted to specific students) using data on students who appear on Crunchbase.com. I build on these papers by expanding both the number of students and the variables covered.[22]

When a university implements an NLP, it commits to covering the entire financial need gap through scholarships, grants, or work-study programs, eliminating the need for federal or private student loans. It is important to note that students are not required to take the full value of the loan reductions offered under the NLP. Some continue to take out loans in order to reduce work-study hours or parental contribution, or to permit more spending during college. Universities may implement NLPs either for a subset of students (e.g. low-income students) or for the entire student body. For example, as described by Rothstein and Rouse (2011), when Princeton University implemented an NLP for low-income students in 1998, loans were eliminated for new matriculants from families with incomes below $40,000. As a consequence, the average amount of student loans for low-income students fell from $15,000 (for students matriculating in 1997) to less than $500 (for students matriculating in 1998).[23] In 2001, Princeton extended its NLP to cover all students. Like Princeton, a number of other universities considered in this paper implemented their NLPs in two stages, first eliminating loans for students with low family incomes, then extending the policy to cover all students on financial aid. In the second stage, the policy applied to all students on campus at the time, regardless of cohort. Thus, a non-low-income student who was enrolled during the first phase would have been expected to take out loans for those years, but would have been covered by the NLP afterward.

During the period covered by my study, between 2001 and 2019, a total of 22 universities implemented UNLPs (i.e., NLPs covering all students, regardless of family income). Table

[21] I assume that the university studied in Rothstein and Rouse (2011) is Princeton: "Anon U is one of the most selective, expensive colleges in the country, and it admits only the most academically qualified students." Tellingly, "Anon U" implemented an NLP for low-income students in 1998 and an NLP for all students in 2001. To my knowledge, the only university that implemented these specific NLPs at these precise years are Princeton.

[22] For comparison, the dataset used in Krishnan and Wang (2018) includes the students who appear on Crunchbase.com, which is, on average, 0.54 students per university, or 0.017% of graduating students. (According to Table 1, Panel D, on p. 30 of Krishnan and Wang (2018), the average number of students per university in their sample is 2,144 + 837 + 168 = 3,149, and 0.54/3,149 = 0.017%.)

[23] See Rothstein and Rouse (2011), Figure 1, p. 152.

1 provides a list these 22 universities along with the year of implementation (and end year when applicable). Figure C4, Panel (a), shows the fraction of cohort-schools treated and the intensity of the treatment. For example, if a university implemented a UNLP in the fall of 2001, then the cohort of students graduating in the spring of 2002 would have been treated for 25% of the time they were in college, and the cohort graduating in the spring of 2003 would have been treated for 50% of the time they were in college.

Figure C2 in the online appendix shows the total number of NLPs implemented each year, as well as the number of other types of financial aid policies in place (i.e., student loan caps, parental contribution eliminations, tuition waivers, and Pell Grant matches).

4.2 Assumptions

In order to interpret the results of the event study specification as the causal treatment effects of UNLP on major choice, I rely on two key identifying assumptions: no anticipation of the treatment and parallel trends (Sun and Abraham (2021); Borusyak et al. (2021)). Under the no anticipatory effects assumption, I assume that units do not change their behavior in anticipation of the treatment.[24] The second identifying assumption is the parallel trends assumption, in which we assume that absent the reform, the difference in potential outcomes would be the same across all units and all periods conditional on the set of controls, unit and time fixed effects.[25]

The main threat to identification is that unobserved changes in the composition of students attending the UNLP school can explain both the timing of the UNLPs and changes in major choice. For example, with the implementation of the policy, the school may enroll students that are more interested in studying specific types of majors. In order to mitigate this issue, I focus on students that were enrolled at the time of the policy implementation. Specifically, I focus only on students that were sophomores, juniors, and seniors at the time the policy is implemented.

As highlighted by the recent econometric literature, the estimates may also be biased if there is heterogeneity in treatment effects between groups of units treated at different times. This bias can occur even if both the no-anticipatory effects and parallel trends assumptions hold (de Chaisemartin and D'Haultfoeuille, 2020; Sun and Abraham, 2021; Goodman-Bacon, 2021; Borusyak et al., 2021; Callaway and Sant'Anna, 2021). Given the potential for biased estimates, I employ the Sun and Abraham (2020) correction methodology, which adjusts for potential treatment heterogeneity across cohorts.

[24] Formally, in the notation of potential outcomes, this is equivalent to $E[Y_t - Y_t(0)|X] = 0$ for all t prior to the policy, conditional on covariates X.

[25] Formally, we assume that $E[Y_{it}(0) - Y_{it}0(0)|X]$ has to be the same across units i for all periods t, t0.

4.3 Determinants of Universal No Loan Policies

News accounts suggest that universities have been implementing UNLPs primarily in order to promote socioeconomic diversity within the student body, rather than to influence students' career outcomes (Rothstein and Rouse, 2011). The main goals are to ease the financial burden on students, boost enrollment of low-income students, and increase the pool of qualified candidates. University officials have stated that UNLPs are meant to help institutions fulfill their "commitment to providing financial aid to all students" (Linsenmeier et al., 2006).

Although changes in enrollment as a result of UNLPs are not the focus of my study, in the online appendix I provide descriptive evidence of how the proportion of low-income students (as measured by the number of Pell Grant recipients) evolves in the years before and after UNLP implementation. I do not find any discontinuous change in the number of Pell Grant recipients. However, I note that Rosinger et al. (2019) find increases in the enrollment rates of lower- and middle-class students following UNLP implementation. Their results highlight the selection bias that could result if my analysis included students who enrolled at a university only after it implemented a UNLP. I return to this issue below, in Section 4.5, where I discuss how I address it.

4.4 Benchmark Specifications

The empirical strategy is to study the effect of UNLPs on major choice, comparing individuals who were not exposed to an UNLP to those who were, at the time of policy implementation. I implemented this comparison with the following main specifications:

$$y_{i,s,t} = \sum_{j=-4}^{3} \mathbf{1}_{\text{cohort } j}\beta_j + \gamma_s + \gamma_t + X_i + \epsilon_{i,s,t}, \tag{6}$$

$$y_{i,s,t} = \alpha + \beta \times \text{Post}_t + \gamma_s + \gamma_t + X_i + \epsilon_{i,s,t}, \tag{7}$$

where (6) is the event study and (7) is the difference-in-differences model. The dependent variables in both models, $y_{i,s,t}$, is either a dummy variable indicating a specific major or occupation, or a continuous variable measuring an expected earnings characteristic, for individual i, in school s, in year t. The variable $\mathbf{1}_{\text{cohort } j}$ is a dummy variable indicating whether the student graduated in cohort j relative to policy implementation, where 0 is the year prior to implementation (and the omitted category). Finally, X_i represents demographic controls, γ_s is a vector of school fixed effects, and γ_t represents year fixed effects.

In regression (6), the vector β contains the main coefficients of interest. For each year, it

measures the average change in the outcome for an individual who was treated by the policy, relative to year 0, over and above the average change over the same period for an individual not treated by the policy. In the regression (7), β again contains the main coefficients of interest. All point estimates are reported with standard errors clustered at the school level.

4.5 Empirical Challenges and Robustness Specifications

In this subsection, I discuss three potential empirical challenges and concerns, and I describe how I address them.

One potential concern is that there could be a selection bias in the types of students who attend a given university before or after it implements a UNLP. For instance, a particular type of student might be more likely to apply to and enroll in universities that have a UNLP. I address this concern in two ways. First, I restrict the event window to include only graduation years up to $j = 3$. This means that I am only studying students who were already enrolled at the time of the policy implementation.

A second potential concern is that in the benchmark specification, the sample consists only of students from the 22 universities that implemented UNLPs, and the empirical model estimates the effects of UNLPs through variation in the implementation time. To test the robustness of the empirical results, I estimate a regression model which includes a control group of universities that never implemented a UNLP. As the control group I include all (non-UNLP-implementing) universities from Barron's list of the top 50 U.S. universities. (These universities include, for example, MIT, Duke, and UC Berkeley.) Specifically, I implement the following two regression models:

$$y_{i,s,t} = Treated_s \times \sum_{j=-4}^{3} \mathbf{1}_{\text{cohort } j}\beta_j + \gamma_s + \gamma_t + X_i + \epsilon_{i,s,t}, \tag{8}$$

$$y_{i,s,t} = \beta \times Treated_s \times \text{Post}_{s,t} + \gamma_s + \gamma_t + X_i + \epsilon_{i,s,t}, \tag{9}$$

where $Treated_s$ is a dummy variable taking the value 1 if the university is among the 22 that implemented UNLPs and 0 otherwise, and the other variables are as in the models above. The regression model (8) is a difference-in-differences event-study model, where the vector of coefficients of interest, β_j, captures the average difference in the outcome variable in year j relative to year 0 for treated schools. The regression model (4.6) is a standard collapsed difference-in-differences model, where β is the coefficient of interest.

A third potential concern is that there might be heterogeneity in the treatment effect across cohorts. A growing econometrics literature has shown that two-way fixed effects

(TWFE) models (such as (6) and (7)) provide a weighted average of each cohort-specific coefficient. This weighted estimate can sometimes be biased away from the standard interpretation of a difference-in-differences estimate, especially if the treatment effect differs across cohorts.[26] To address this concern, I also estimate the bias-correcting difference-in-differences event-study regression model recommended by Sun and Abraham (2020).

4.6 Instrumental Variable Approach

To quantify the causal effect of student on major choice, I also implement an instrumental variable (IV) approach. I use the implementation of a UNLP as an instrument for the amount of student debt. Using a standard two-stage least-squares estimator, the first stage is similar to Equation (4.6), with student debt as the dependent variable, with the one exception that I can only run the regression on the UNLP sample:

$$StudentDebt_{i,s,t} = \beta \times \text{Post}_{s,t} + \gamma_s + \gamma_t + X_i + \epsilon_{i,s,t}, \tag{10}$$

where $StudentDebt_{ist}$ is the amount of student debt at graduation for individual i at school s in cohort t. In the second stage, I regress the same outcome variables used in regression on the predicted value of student debt from Equation (10):

$$y_{i,s,t} = \eta \; \widehat{StudentDebt}_{i,s,t} + \gamma_s + \gamma_t + X_i + \epsilon_{i,s,t}. \tag{11}$$

The central identifying assumption for this strategy is that implementation of a UNLP is a valid instrument for student debt, which requires (1) a strong first stage, (2) that the UNLP is an independent instrument, (3) that the exclusion restriction holds, and (4) that the UNLP affects student debt monotonically. (In Section D in the Online Appendix , I discuss each of these four assumptions in detail.) Granting these assumptions, the coefficient η captures the local average treatment effect of student debt. That is, η captures the causal effect of student debt for students whose student debt was affected by the policy. For example, given appropriate scaling, the interpretation of η is *the effect of $1,000 of student debt on the implied initial wage from the choice of college major.*

[26] For the econometrics literature, see, for example, Borusyak and Jaravel (2017), de Chaisemartin and D'Haultfœuille (2020), Callaway and Sant'Anna (2020), and Sun and Abraham (2020). For an application in finance, see Baker et al. (2021).

4.7 Heterogeneity across Family Income Levels

Finally, I explore how the effects of student debt vary with family income. As noted in the introduction, most of the previous literature studies the effects of student debt for the average student. However, students' investments in human capital are often at least partially financed by their parents, and it is natural to assume that the parents' financial wealth can act as a credit constraint. Moreover, financial aid is a scarce resource. While estimates for the average student can serve as an argument to increase financial aid overall, such estimates do not address the more complex and pressing problem of how best to allocate financial aid. Instead, to design better allocation policies we need to ask: Who is most constrained by student debt? The answer is not obvious a priori. For example, students from the poorest families may already qualify for full tuition waivers, while students from the richest families may rely on their parents to fund their education. In this case, students from the middle of the income distribution would be likely to display the largest response to changes in student loan policies.

In order to decompose the effects of student loans according to family income, I have collected USPS records not only of students, but also of their parents. While I cannot directly observe family income, I can proxy for it by exploiting geographic information. Specifically, by matching the commencement program data to USPS records, I can see a history of addresses and approximate dates of residence for each student. Additionally, the USPS records allow me to see all individuals by name at a particular address. Using this information and other demographic data, I can allocate individuals into households. Then, from the ACS, I get median household income data at the ZIP Code level.

For this part of the analysis, I run the two main regression specifications, interacting the effect with proxies for family income. Specifically, I run the following two specifications:

$$y_{i,s,t} = \alpha + \beta_1 \text{Post}_t + \beta_2 \text{Family Income}_i + \beta_3 \text{Post}_t \times \text{Family Income}_i + \gamma_s + \gamma_t + X_i + \epsilon_{i,s,t}, \quad (12)$$

$$y_{i,s,t} = \sum_{j=-4}^{3} \mathbf{1}_{\text{cohort } j} \beta_j + \sum_{j=-4}^{3} \mathbf{1}_{\text{cohort } j} \mu_j \times \text{Family Income}_i + \gamma_s + \gamma_t + X_i + \epsilon_{i,s,t}, \quad (13)$$

where Family Income$_i$ is defined either as the natural logarithm of the median household income in the ZIP Code where the family lives, or as an indicator of whether the student student grew up in a ZIP Code where the median household income is above the nationwide median.

5 The Effect of Student Loans on Major Choice

In this section, I present the empirical results. First, I present the first-stage results on the effect of UNLPs on student debt take-up. Then I present the main result of the paper, on how UNLPs affect the choice of implied wage trajectories, and in a heterogeneity analysis, I detail how the effects vary with the family's income and credit scores. Next, to quantify the results, I present the IV results, and, to explore the economic mechanism driving the main result, I present the results on individual major and occupational choices. Finally, I provide additional evidence on how UNLPs affected the characteristics of major choice by affecting the difficulty of the chosen majors.

5.1 Effects on Student Debt

In this subsection, I present the results of UNLPs on student debt take-up using both publicly available data and the hand-collected commencement program data merged with credit bureau data.

First, as a proof of concept, I present a Difference-in-Differences (DID) test and an event study of the effect of the first stage using publicly available data from IPEDS. Table 3 reports the DID estimates, where the dependent variable is the fraction of undergraduates with student debt, and compares UNLPs to other financial aid programs, including income-specific NLPs, loan caps, parental contribution elimination, and tuition waivers. (For example, some colleges implement NLPs for families whose incomes fall below a specific level.) I find that UNLPs meaningfully decrease the percentage of students taking loans; specifically, they lower it by 12.6 percentage points.

Next, to understand the dynamic effects, Figure C4 reports the results from a standard event-study regression. The plot shows the coefficients on the year relative to implementation. We see a sharp drop in the fraction of students taking out loans that precisely coincides with the timing of the policy. To assess the robustness of this result, in Figure C4, I plot the regression coefficients and 95% confidence intervals from three different regressions of the percentage of students taking a student loan on year dummies relative to the implementation of a UNLP. A standard two-way fixed effect (TWFE) model is reported in blue, a TWFE model with non-implementing top-50 schools as a control group is reported in red, and a bias-corrected model allowing for treatment heterogeneity across cohorts (following Sun and Abraham (2020)) is reported in green. Across all three specifications, in the year when a UNLP is implemented, there is a significant drop in the share of students who take student loans.

Second, I present the results using the main sample with micro-data. Figure 5 shows the

coefficients from equation (6) where the dependent variable is the student loan balance in the year of graduation.[27] We see a sharp discontinues drop in the amount of student loans that students have that precisely coincides with the timing of the policy. As above, to assess the robustness of this result, I plot the regression results from both a standard TWFE, a TWFE with control groups, and a bias-corrected model allowing for treatment heterogeneity across cohorts. The results are similar across all three models.

Table 5 Column (1) reports the DID coefficient for the standard TWFE model. Following the implementation of the UNLP students, on average, have $4,900 less in student debt. As described in Section 4.5, the main sample covers the student who were already enrolled at the time of the policy implementation. Next, to study the effect of a full implementation, Table 5 Column (2) report the regression coefficient from a TWFE model where the post variable is continuous in treatment intensity. Specifically, I let the post variable take the values of 0.25, 0.50, and 0.75 for students who were seniors, juniors, and sophomores, respectively. I find a coefficient of $9,200. The interpretation of the regression coefficient is what the reduction in student debt would have been, if they had been treated for their entire time in college. Interestingly, this estimate is close in magnitude to the reduction in student loans of $11,000 that Rothstein and Rouse (2011) find.

Taken together, these results confirm that following the implementation of UNLPs, the use of student loans fell dramatically.

5.2 Implied Earnings Trajectories

In this subsection, I present the main results of the paper. Here, I regress the implied wage characteristics (imputed from the ACS data) on UNLP adoption.

Figure 6 shows the event-study plots. Panels (a), (b), (c), and (d) show the coefficients from the regression (6) with the outcome variable taken as the mean, standard deviation, slope, and initial wage, respectively. In Panels (a), (b), and (c) we see that following UNLP implementation, students choose majors that are associated with earnings that are both higher on average, have higher standard deviation, and have higher earnings growth. Prior to UNLP implementation, there is no pre-trend; afterwards, we see a discontinuous increase of between $800 and $1,700 in dollars in mean earnings. In Panel (d), we see that following UNLP implementation, students choose majors with initial earnings between $100 and $400 lower than those of the majors they chose before.[28]

[27] The credit bureau data offers a yearly snapshot on December 31. Therefor, the student loan balance variable captures the amount of student loan debt on December 31 on the year that the student graduated.

[28] In each of the four panels, I plot both the coefficients from a regression with only treated schools and those from another specification that includes non-treated schools.

To quantify the average effects on students, I turn to the collapsed difference-in-differences regression. Table 4 reports the regression coefficients. Column 1 show the coefficient for the mean, Column 2 for the slope, Column 3 for the initial earnings, and Columns 4 through 6 for the average earnings across three different decades (the 20s, 30s, and 40s, respectively). I find that after UNLP implementation, students choose majors associated with increases of $1,200 in mean earnings, an 0.65% increase in earnings slope, a decrease of $270 in initial earnings, a decrease of $147 in annual earnings in the 20s, and an increase in annual earnings of $830 and $1,500 in the 30s and 40s, respectively.

Table 5 Columns (3) through (8) report the IV estimates from the two-stage least squares model. I scale the coefficients by one thousand, such that they represent the effect of a $1,000 increase in (predicted) loan amount. We see that for each $1,000 increase in student loans, student choose majors that are associated with a decrease in mean life-time earnings of $190 with a lower earnings slope of 12%. When we look at implied earnings across the life-cycle, we see that a $1,000 increase in student loans leads students to choose majors associated with a $46 increase in initial annual earnings, and a $137 and $261 decline in annual earnings in their 30s and 40s, respectively.

These results highlight that higher student debt can simultaneously lead to both higher and lower earnings at different points in a student's career: Students with less debt are more likely to choose careers that have *lower initial wages* after college and *higher earnings growth* in the long term.

5.3 The Role of Family Income

In this Section, to provide a deeper perspective on how financial frictions affect human capital investments, I investigate heterogeneity in the effects of UNLPs across the income and credit score distributions.

First, to provide context, I plot the distributions of ZIP-level income and household-level credit scores. Figure C6, Panel (A), shows histograms of the distribution of ZIP-level income, with data for a random sample of the U.S. population in gray and data for students at universities with UNLPs in blue. There are two takeaways from this plot. The first is that families that send their children to universities with UNLPs are wealthier than the average family in the U.S. The second is that, nevertheless, students from lower-income families also attend these universities. Panel (B) shows the same histogram for UNLP families but compares them with the families of students who attended non-UNLP control schools. Here we see virtually no difference between the income distributions.

The same pattern holds for credit score distributions, as shown by the histograms in

Figure C7. Panel (A) shows the distribution of credit scores for the families of students at universities with UNLPs (in blue), relative to a random sample of the U.S. population (in gray). Panel (B) again shows the credit scores of UNLP families (in blue), but compares them to those of the families of students attending non-UNLP control schools (in gray).

Tables 6 and 7 summarize how the effects of a UNLP vary by family background. They report the regression coefficients from the collapsed difference-in-differences model interacted with the family characteristic (equation 12). In both tables, Column 1 reports the coefficients for the mean life-time wage by major choice, Column 2 for the slope, and Column 3 for the initial wage. Tables 6 reports the result from the regression where the family characteristic is an indicator variable taking the value 1 if the student's family lives in a ZIP Code where the median household income is below the median.

Prior to the policy implementation, students from families living in below-median income ZIP codes chose majors with an average lifetime earnings of more than $1,000 lower than students families that lives in above-median ZIP codes. Following the policy, both students from above- and below-median income ZIP codes chose higher earning majors. However, the effect of the policy is twice as large for students from low-income ZIP codes. In fact, the policy leads the gap in major choice between students from above- and below-median ZIP codes to fall by 44% ($463/$1,050=0.44). The same pattern is true when I study the effect of the policy of the implied earnings slope: prior to the policy, students from below-median ZIP codes chose majors with lower earnings slopes. Following the policy, students chose majors with steeper slopes, and the effects are roughly twice as large for students from below-median ZIP codes, which cuts the gap by 38%. The results are similar (but with reverse signs) when studying the the effect of the policy on initial wage. Prior to the policy, students from below-median ZIP codes chose majors with higher initial earnings. The policy leads students to chose majors with lower initial earnings, and the effects are largest among students from below-median ZIP codes, which cuts the gap between below- and above-median by 45%.

Table 7 report how the effects of UNLP vary by the parents credit score. The results are similar to those in Table 6. Prior to the policy, students who's parents have below-median credit scores chose majors with lower lifetime earnings, lower earnings slopes, and higher initial earnings. The policy leads students to chose majors with higher lifetime earnings, higher earnings slopes, and lower initial earnings. The effects are between 0.9 and 1.8 times larger for students from below-median credit score families, and the effects lower the gap between students from below- and above-median credit score families by between 38% and 43%.

Taken together, these results indicate that the financial constraints of the parents affect the choice of human capital investments of the children. In particular, students who grew

up in low-income families and families with low credit scores are more likely to be facing financial constraints.

5.4 Majors and Occupations

Next I explore the economic mechanism behind the main results, in two steps. First, I ask which majors students were choosing before and after the policy change. Specifically, I run the regression (7) with the outcome variable being a dummy variable that takes the value 1 if an individual graduated with a particular major.

Figure 7 shows a plot of the difference-in-differences coefficients. Each coefficient represents the change in the fraction of students choosing a specific major, controlling for student characteristics and for school and cohort fixed effects. (As before, the treated sample includes only students who were enrolled when the UNLP was introduced.)

We see that UNLPs induced more students to major in the physical sciences, biology and life sciences, history, and mathematics and statistics, while it reduced the number of students majoring in medical and health sciences and services (a category that includes nursing) and in business. Thus, the increases in the mean, standard deviation, and slope of earnings are primarily driven by the choice of STEM majors, history, and the social sciences.

Taken together, these results show that after UNLP implementation, students choose majors that are associated with, on average, a $1,000 increase in annual wage. As an illustration, this is equivalent to the increase that would result if all students majoring in psychology were to change their major to pre-law, or if 8% of all students majoring in medical and health sciences and services (which includes nursing) were to switch to biology and life sciences.[29]

As a second step, I ask whether UNLPs change the types of jobs that students get. Figure C5 reports the results from the model (7) for nine different job types. We see that following UNLP implementation, when student debt decreases, the probability that a student will become a teacher, nurse, or physician's assistant decreases. Interestingly, all of these jobs are characterized by stable earnings and lower expected earnings growth. In contrast, the probability that a student will become a doctor, lawyer, author, artist, or entrepreneur increases. These jobs are characterized by either low initial earnings after college (many prospective doctors and lawyers effectively have negative earnings while they are in graduate school and pay tuition), or high earnings risk (artists, authors, and entrepreneurs).

[29] The difference between the mean wages in pre-law and psychology is $-\$4,720 - (-\$5,787) = \$1,067$ (see Table D1). The difference between the mean wage for biology and that for medical and health sciences and services is $\$19,728 - \$6,523 = \$13,205$. Thus, if one in every 13 nursing students changes to biology, this corresponds to an average change of roughly $1,000 across the original population of nursing students.

5.5 Additional Results on Major Characteristics

To explore the economic mechanisms driving the main results, I analyze additional characteristics that may affect students' major choices, especially in the presence of financial frictions. Specifically, I collect data on the number of hours that students spend studying for each major, the difficulty level students report for each major, and the mean grade point average (GPA) by major.

First, from the results of the 2016 National Study of Student Engagement (administered by Indiana University), I obtain the average number of hours per week that students study, as reported by undergraduate students in each major. According to this survey, the majors requiring the most study time per week include chemical engineering (20 hours) and biomedical sciences (18 hours), while those requiring the least include social work (12 hours) and management (13 hours).

Second, I analyze whether UNLPs lead students to choose more difficult majors. For this, I use data provided by Vitaliy Novik, who has created a popular website compiling hundreds of thousands of student course evaluations. Each evaluation includes a self-reported level of difficulty, and Novik aggregates this data by major. The most difficult majors include biomedical sciences (3.62/5), chemical engineering (3.49/5), and molecular biology (3.42/5), while the least difficult include secondary education (2.38/5), elementary education (2.41/5), and human resources and personnel management (2.66/5).

Third, I analyze students' major choices in terms of GPA. I use data from Rask (2010), who provides statistics on mean GPA by major. The majors associated with the highest mean GPA include education (3.36/4) and languages (3.34/4), while those with the lowest mean GPA include chemistry (2.78/4), mathematics (2.9/4), and economics (2.95/4).

Table D6 reports the regression coefficients from difference-in-differences regressions where I regress each major choice characteristic on a dummy for UNLP times post. The coefficient is positive and significant for required hours of work and self-reported difficulty and negative for average GPA. These results suggest that by removing financial frictions, UNLPs induce students to choose majors that require more hours of work, are rated as more difficult, and are associated with lower grades. These findings are in line with recent findings from Luo and Mongey (2019) who show that financial assets affect the choice of non-wage amenities in job search.

6 Conclusion

In this paper, I study how student loans affect labor supply, human capital investments, and long-term labor market trajectories. Relatedly, I study how wealth transfers early in an individual's life (in the form of financial aid for college) affect these outcomes. I build a novel dataset by merging individual resume data with university records and postal service records. As the source of empirical variation, I exploit the staggered implementation of universal no-loan policies (UNLPs) at universities across the U.S.

This paper makes two main contributions. The first contribution relies on the insight that most human capital investments are intertemporal and require a trade-off between current and future income. For example, choosing to go to medical school will result in lower current income and higher expected future income. In the paper, I analyze how student loans affect this trade-off—in particular, the extent to which they induce students to choose career paths with lower earnings slopes. I find that following the implementation of a UNLP, students are much less likely to take on student debt; in addition, they choose different fields of study and end up with different career paths. Specifically, under a UNLP, students are more likely to choose majors and careers that require additional investments and that are associated with higher mean earnings and higher earnings slopes.

The paper's second contribution is its analysis of how student loans differentially affect students depending on their family backgrounds. Most of the previous literature studies the effects of student loans for the average student. However, students' investments in human capital are often at least partially financed by their parents, and it is natural to assume that the parents' financial wealth can act as a credit constraint. Moreover, financial aid is a scarce resource.

While estimates for the average student can serve as an argument to increase financial aid overall, such estimates do not address the more complex and pressing problem of how best to allocate financial aid. Instead, to design better allocation policies we need to ask: Who is most constrained by student debt? In this paper, I find that student debt has a stronger effect on the major choices and implied wage trajectories of students from low-income families and families with low credit scores.

References

Agarwal, S., Liu, C., and Souleles, N. S. (2007). The reaction of consumer spending and debt to tax rebates - evidence from consumer credit data. *Journal of Political Economy*, 115(6):986–1019.

Altonji, J. G. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics*, 11(1, Part 1):48–83.

Altonji, J. G., Blom, E., and Meghir, C. (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics*, 4(1):185–223.

Andrews, R. J., Imberman, S., and Lovenheim, M. (2017). Risky Business? The Effect of Majoring in Business on Earnings and Educational Attainment. Working Papers 2017-060, Human Capital and Economic Opportunity Working Group.

Andrews, R. J. and Stange, K. M. (2019). Price regulation, price discrimination, and equality of opportunity in higher education: Evidence from texas. *American Economic Journal: Economic Policy*, 11(4):31–65.

Arcidiacono, P. (2004a). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1):343–375. Higher education (Annals issue).

Arcidiacono, P. (2004b). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1):343–375. Higher education (Annals issue).

Arpita, P., Wiswall, M., and Zafar, B. (2020). College majors. In McCall, B. P., editor, *The Routledge Handbook of the Economics of Education*, chapter 1, pages 266–290. Taylor Francis Group, Oxford.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the “other 99 percent”. *Science*, 344(6186):843–851.

Baker, A., Larcker, D. F., and Wang, C. C. Y. (2021). How Much Should We Trust Staggered Difference-In-Differences Estimates? *Working Paper*.

Baker, S. R. (2018). Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data. *Journal of Political Economy*, 126(4):1504–1557.

Baker, S. R. and Yannelis, C. (2017). Income changes and consumption: Evidence from the 2013 federal government shutdown. *Review of Economic Dynamics*, 23:99–124.

Bartik, T. J. and Hershbein, B. J. (2018). Degrees of Poverty: The Relationship between Family Income Background and the Returns to Education. Upjohn Working Papers and Journal Articles 18-284, W.E. Upjohn Institute for Employment Research.

Bernstein, S., McQuade, T., and Townsend, R. R. (2018). Do household wealth shocks affect productivity? evidence from innovative workers during the great recession. Working Paper 24011, National Bureau of Economic Research.

Bettinger, E. (2004). How financial aid affects persistence. Working Paper 10242, National Bureau of Economic Research.

Bettinger, E., Gurantz, O., Kawano, L., Sacerdote, B., and Stevens, M. (2019). The long-run impacts of financial aid: Evidence from california's cal grant. *American Economic Journal: Economic Policy*, 11(1):64–94.

Black, S. E., Denning, J. T., Dettling, L. J., Goodman, S., and Turner, L. J. (2020). Taking it to the limit: Effects of increased student loan availability on attainment, earnings, and financial well-being. Working Paper 27658, National Bureau of Economic Research.

Boneva, T. and Rauh, C. (2017). Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns. *SSRN Electronic Journal*.

Bonin, H., Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2007). Cross-sectional earnings risk and occupational sorting: The role of risk attitudes. *Labour Economics*, 14(6):926–937. Education and Risk.

Borusyak, K. and Jaravel, X. (2017). Revisiting Event Study Designs. *Working Paper*.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation.

Callaway, B. and Sant'Anna, P. H. (2020). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, (Forthcoming).

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230. Themed Issue: Treatment Effect 1.

Carroll, C. D. (1997). Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis. *The Quarterly Journal of Economics*, 112(1):1–55.

Castleman, B. L. and Long, B. T. (2016). Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics*, 34(4):1023–1073.

Cervantes, A. P., Creusere, M. A., McMillion, R., McQueen, C., Short, M., Steiner, M., and Webster, J. (2005). Opening the doors to higher education: Perspectives on the higher education act 40 years later.

Chakrabarti, R., Fos, V., Liberman, A., and Yannelis, C. (2022). Tuition, Debt, and Human Capital. *The Review of Financial Studies*. hhac065.

Chapman, S. (2016). Student Loans and the Labor Market: Evidence from Merit Aid Programs. *Working Paper*.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2020). Income Segregation and Intergenerational Mobility Across Colleges in the United States*. *The Quarterly Journal of Economics*, 135(3):1567–1633.

Cronqvist, H., Makhija, A. K., and Yonker, S. E. (2012). Behavioral consistency in corporate finance: Ceo personal and corporate leverage. *Journal of Financial Economics*, 103(1):20 – 40.

Daniels, G. E. and Smythe, A. (2019). Student Debt and Labor Market Outcomes. *AEA Papers and Proceedings*, 109:171–175.

de Chaisemartin, C. and D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, 110(9):2964–2996.

Deaton, A. (1991). Saving and Liquidity Constraints. *Econometrica*, 59(5):1221–1248.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market*. *The Quarterly Journal of Economics*, 132(4):1593–1640.

Deming, D. J. and Noray, K. (2018). STEM careers and technological change. *Working Paper*.

Denning, J. T., Marx, B. M., and Turner, L. J. (2019a). Propelled: The effects of grants on graduation, earnings, and welfare. *American Economic Journal: Applied Economics*, 11(3):193–224.

Denning, J. T., Marx, B. M., and Turner, L. J. (2019b). Propelled: The effects of grants on graduation, earnings, and welfare. *American Economic Journal: Applied Economics*, 11(3):193–224.

Di Maggio, M., Kalda, A., and Yao, V. W. (2019). Second Chance: Life without Student Debt. *Working Paper*.

Di Maggio, M., Kermani, A., Keys, B. J., Piskorski, T., Ramcharan, R., Seru, A., and Yao, V. (2017). Interest rate pass-through: Mortgage rates, household consumption, and voluntary deleveraging. *American Economic Review*, 107(11):3350–3588.

Dillon, E. W. (2018). Risk and return trade-offs in lifetime earnings. *Journal of Labor Economics*, 36(4):981–1021.

Dynarski, S., Libassi, C., Michelmore, K., and Owen, S. (2018). Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students. Working Paper 25349, National Bureau of Economic Research.

Dynarski, S. M. (2003). Does aid matter? measuring the effect of student aid on college attendance and completion. *American Economic Review*, 93(1):279–288.

Egan, M., Matvos, G., and Seru, A. (2019). The market for financial adviser misconduct. *Journal of Political Economy*, 127(1):233–295.

Field, E. (2009). Educational debt burden and career choice: Evidence from a financial aid experiment at nyu law school. *American Economic Journal: Applied Economics*, 1(1):1–21.

Folch, M. and Mazzone, L. (2022). Go Big or Buy a Home: The Impact of Student Debt on Career and Housing Choices. *Working Paper*.

Friedman, M. (1957). The Permanent Income Hypothesis. In *A Theory of the Consumption Function*, volume I, pages 20–37. Princeton University Press.

Gallagher, E., Billings, S. B., and Ricketts, L. (2022). Human Capital Investment After the Storm. *The Review of Financial Studies (Accepted)*.

Gervais, M. and Ziebarth, N. L. (2017). Life after Debt: Post-Graduation Consequences of Federal Student Loans. *Working Paper*.

Gicheva, D. (2016). Student loans or marriage? A look at the highly educated. *Economics of Education Review*, 53:207–216.

Goodman, S., Isen, A., and Yannelis, C. (2021). A day late and a dollar short: Liquidity and household formation among student borrowers. *Journal of Financial Economics*, 142(3):1301–1323.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277. Themed Issue: Treatment Effect 1.

Hall, R. E. (1978). Stochastic Implications of the Lifecycle Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy*, 86(6):971–987.

Hamermesh, D. S. and Donald, S. G. (2008). The effect of college curriculum on earnings: An affinity identifier for non-ignorable non-response bias. *Journal of Econometrics*, 144:479–491.

Hearn, J. C. (2015). Outcomes-Based Funding in Historical and Comparative Context. *Lumina Issue Papers*.

Hemelt, S. W., Hershbein, B., Martin, S. M., and Stange, K. M. (2021). College majors and skills: Evidence from the universe of online job ads. *Working Paper*.

Hershbein, B. (2016). A college degree is worth less if you are raised poor. Brookings, https://www.brookings.edu/blog/social-mobility-memos/2016/02/19/a-college-degree-is-worth-less-if-you-are-raised-poor/.

Houle, J. N. and Berger, L. (2015). Is Student Loan Debt Discouraging Homeownership among Young Adults? *Social Service Review*, 4(89):589–621.

Jappelli, T. and Pistaferri, L. (2014). Fiscal policy and mpc heterogeneity. *American Economic Journal: Macroeconomics*, 6(4):107–36.

Jeffers, J. (2017). The impact of restricting labor mobility on corporate investment and entrepreneurship. *Working Paper*.

Johnson, D. S., Parker, J. A., and Souleles, N. S. (2006). Household expenditure and the income tax rebates of 2001. *American Economic Review*, 96(5):1589–1610.

Jørring, A. (2020). Financial Sophistication and Consumer Spending. *Working Paper*.

Krishnan, K. and Wang, P. (2018). The Impact of Student Debt on High Value Entrepreneurship and Venture Success: Evidence from No-Loans Financial Aid Policies. *Working Paper*.

Krishnan, K. and Wang, P. (2019). The Cost of Financing Education: Can Student Debt Hinder Entrepreneurship? *Management Science (Forthcoming)*.

Linsenmeier, D. M., Rosen, H. S., and Rouse, C. E. (2006). Financial Aid Packages and College Enrollment Decisions: An Econometric Case Study. *The Review of Economics and Statistics*, 88(1):126–145.

Lips, A. J. A. (2011). A Typology of Institutional Loan Replacement Grant Initiatives for Low- and Moderate-Income Students. *The Review of Higher Education*, 34(4).

Lucca, D., Seru, A., and Trebbi, F. (2014). The revolving door and worker flows in banking regulation. *Journal of Monetary Economics*, 65:17 – 32.

Luo, M. and Mongey, S. (2019). Assets and Job Choice: Student Debt, Wages and Amenities. *Working Paper*.

Mettler, S. (2014). *Degrees of Inequality - How the Politics of Higher Education Sabotaged the American Dream*. Basic Books.

Minicozzi, A. (2005). The short term effect of educational debt on job decisions. *Economics of Education Review*, 24(4):417–430.

Nielsen, H. S. and Vissing-jorgensen, A. (2006). The impact of labor income risk on educational choices: Estimates and implied risk aversion. working paper.

Osman, M. (2021). Mind-blowing linkedin statistics and facts (2021).

Parker, J. A. (1999). The Reaction of Household Consumption to Predictable Changes in Social Security Taxes. *American Economic Review*, 89(4):959–973.

Pool, V. K., Stoffman, N., and Yonker, S. E. (2015). The people in your neighborhood: Social interactions and mutual fund portfolios. *The Journal of Finance*, 70(6):2679–2732.

Rosinger, K. O., Belasco, A. S., and Hearn, J. C. (2019). A boost for the middle class: An evaluation of no-loan policies and elite private college enrollment. *The Journal of Higher Education*, 90(1):27–55.

Rothstein, J. and Rouse, C. E. (2011). Constrained after college: Student loans and early-career occupational choices. *Journal of Public Economics*, 95(1):149 – 163.

Ruggles, S., Genadek, K., Goeken, R., Grover, J., and Sobek, M. (2017). Integrated public use microdata series: Version 6.0 [data set]. *Minneapolis: University of Minnesota*.

Saks, R. E. and Shore, S. H. (2005). Risk and career choice. *The B.E. Journal of Economic Analysis & Policy*, 5.

Scott-Clayton, J. and Zafar, B. (2019). Financial aid, debt management, and socioeconomic outcomes: Post-college effects of merit-based aid. *Journal of Public Economics*, 170:68 – 82.

Souleles, N. S. (1999). The response of household consumption to income tax refunds. *American Economic Review*, 89(4):947–958.

Stange, K. M. (2012). An empirical investigation of the option value of college enrollment. *American Economic Journal: Applied Economics*, 4(1):49–84.

Stinebrickner, R. and Stinebrickner, T. (2008). The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *American Economic Review*, 98(5):2163–84.

Stinebrickner, R. and Stinebrickner, T. R. (2014). A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout. *Review of Economic Studies*, 81(1):426–472.

Sun, L. and Abraham, S. (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, (Forthcoming).

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199. Themed Issue: Treatment Effect 1.

Turner, S. E. and Bowen, W. G. (1999). Choice of major: The changing (unchanging) gender gap. *ILR Review*, 52(2):289–313.

Weidner, J. (2016). Does Student Debt Reduce Earnings? *Working Paper*.

Wiswall, M. and Zafar, B. (2014). Determinants of College Major Choice: Identification using an Information Experiment. *The Review of Economic Studies*, 82(2):791–824.

Wiswall, M. and Zafar, B. (2017). Preference for the Workplace, Investment in Human Capital, and Gender*. *The Quarterly Journal of Economics*, 133(1):457–507.

Xia, X. (2016). Forming wage expectations through learning: Evidence from college major choices. *Journal of Economic Behavior Organization*, 132:176–196.

Yannelis, C. and Tracey, G. (2022). Student loans and borrower outcomes. *Annual Review of Financial Economics*, 14(1):167–186.

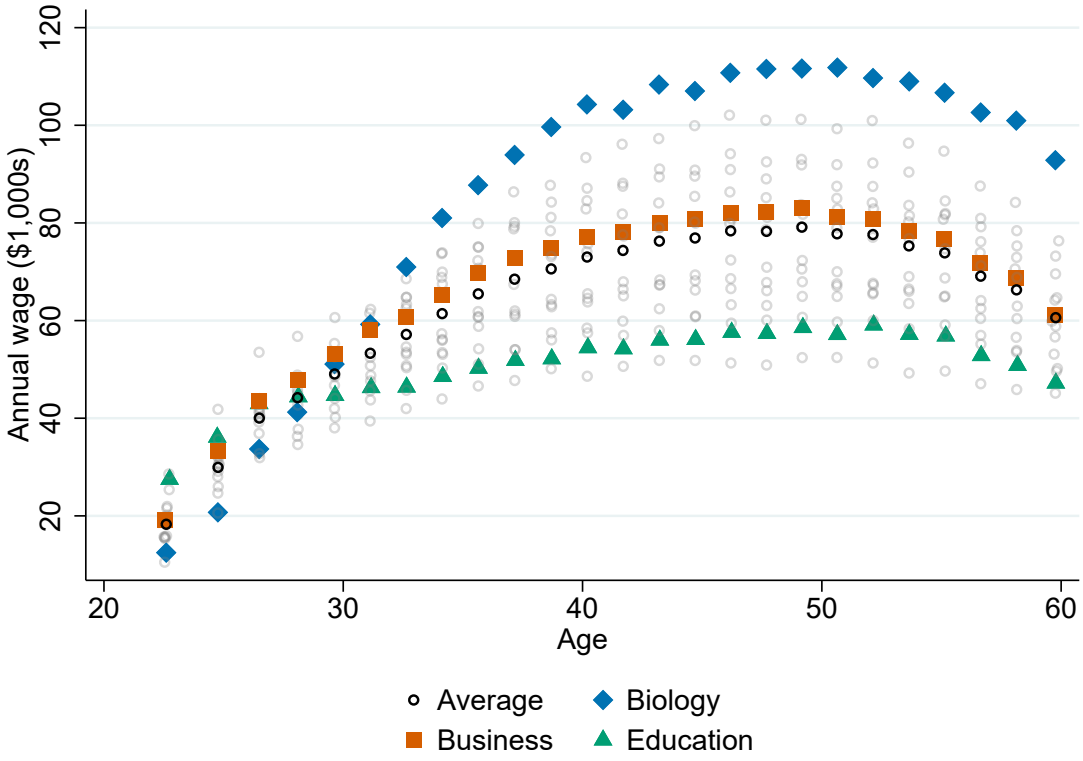
Zafar, B. (2011). How do college students form expectations? *Journal of Labor Economics*, 29(2):301–348.

Zafar, B. (2013). College major choice and the gender gap. *The Journal of Human Resources*, 48(3):545–595.

Zeldes, S. P. (1989). Consumption and Liquidity Constraints: An Empirical Investigation. *Journal of Political Economy*, 97(2):305–346.

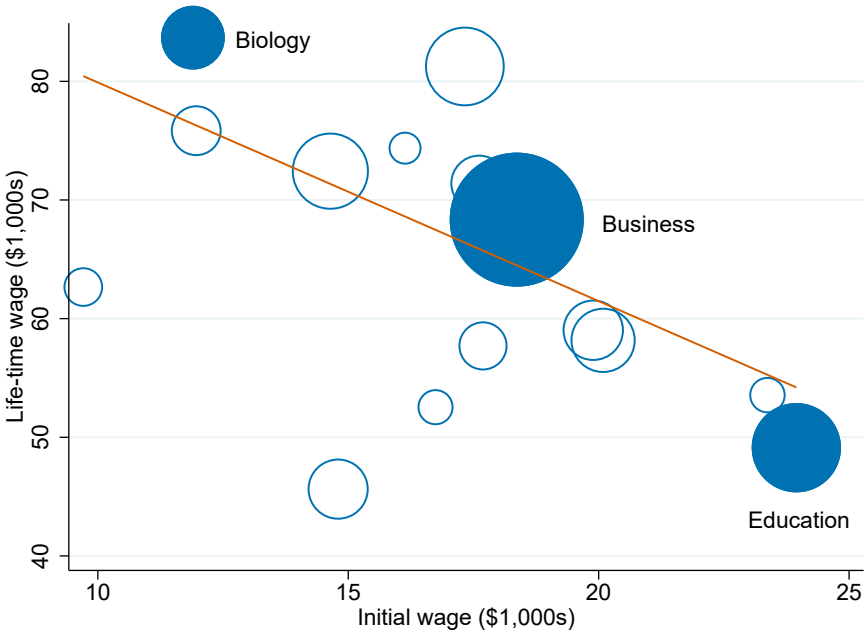
Figures

Figure 1: College Majors & Life-cycle Wages

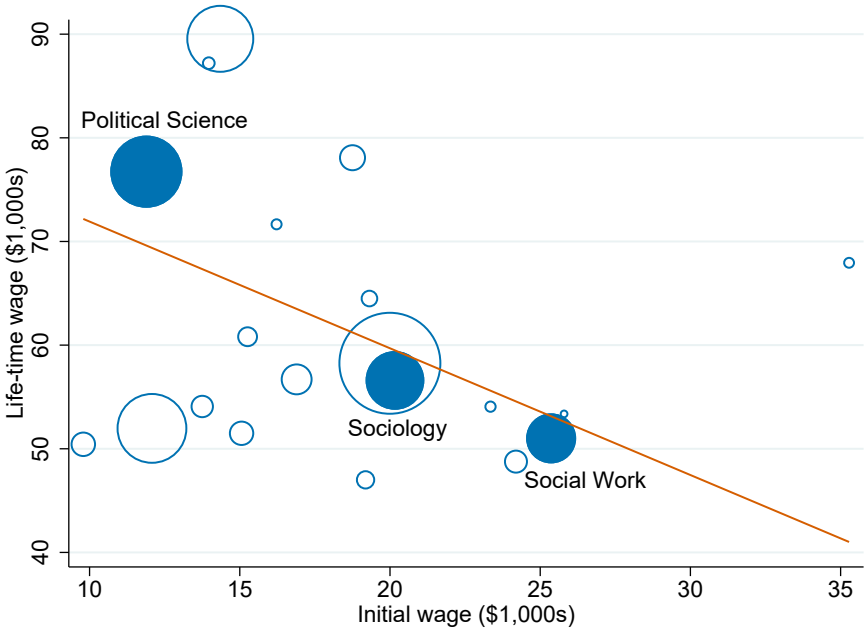


Note: This figure describes the trade-off across college majors between initial and life-time wage. It shows a binned scatter plot of residualized annual wages relative to age (controlling for race, ethnicity, gender, and year, and adding back the sample mean of each variable back to its residuals) across the 15 most frequent college majors (using the 2-digit CIP classification). The **black circles** report the average across all majors. The blue diamonds, red squares, green triangles report the residualized wages for students who majored in biology, business, and education, and the gray circles represent the remaining top-15 majors. The data source is the American Community Survey (ACS) and the National Center for Education Statistics (NCES).

Figure 2: The Intertemporal Trade-off in Major Choice



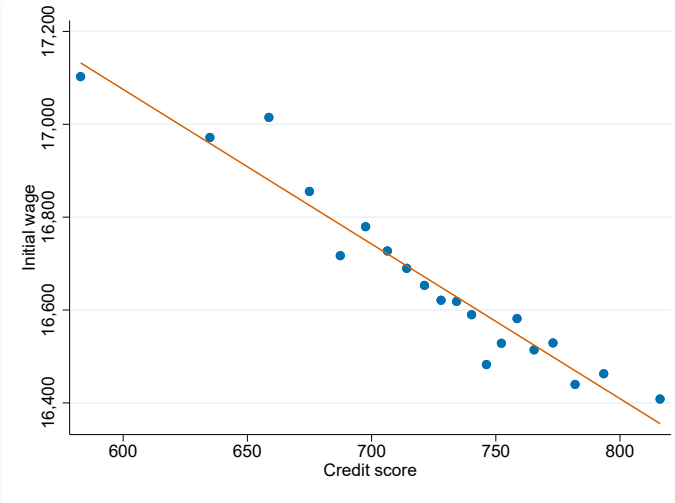
(A) Two-digit majors



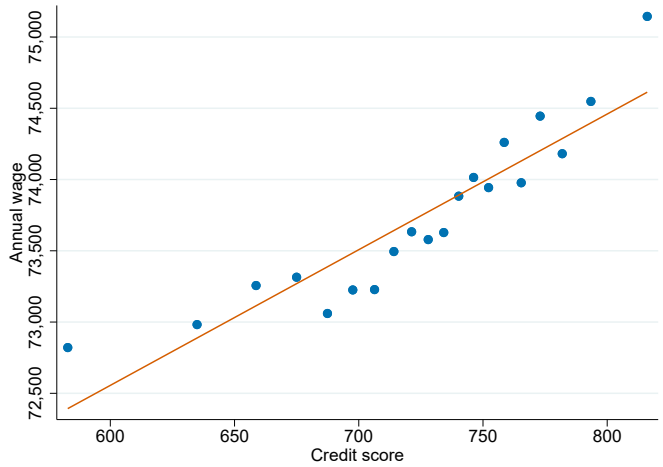
(B) Four-digit majors

Note: This figure describes the trade-off across college majors between initial and life-time wage. Panel (A) plots a binned scatter plot of residualized wages relative to age (controlling for race, ethnicity, gender, and year, and adding back the sample mean of each variable back to its residuals) across the 15 most frequent college majors (using the 2-digit CIP classification). The y-axis plots annual life-time wages and the x-axis plots the average initial wage calculated as the annual wage between the ages of 21 and 23. Panel (B) plots the same variables as Panel (A), but the sample includes only the 20 most frequent majors within "Social Science" and "Public Policy" using the 4-digit CIP classification. The data source is the American Community Survey (ACS) and the National Center for Education Statistics (NCES).

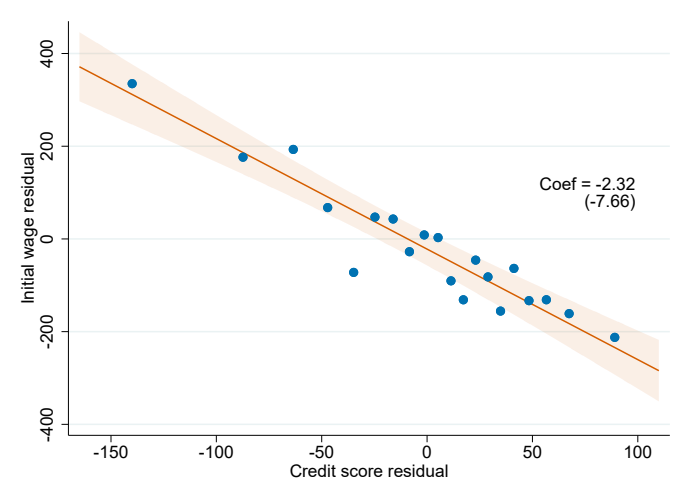
Figure 3: Major Choice and Parents' Credit Scores



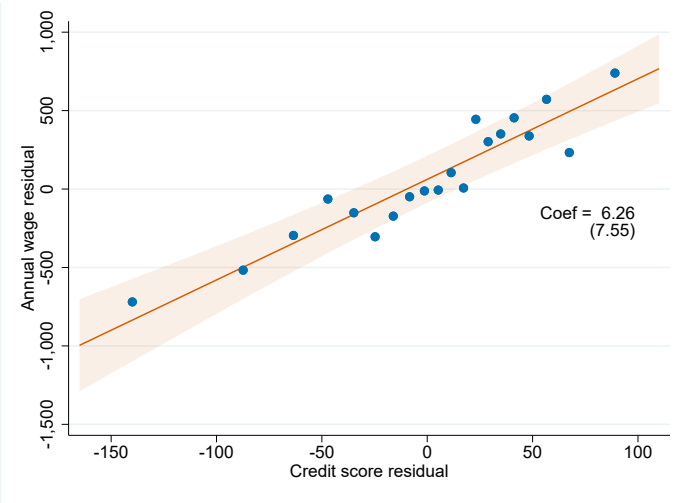
(A) Initial wage



(B) Average lifetime wage



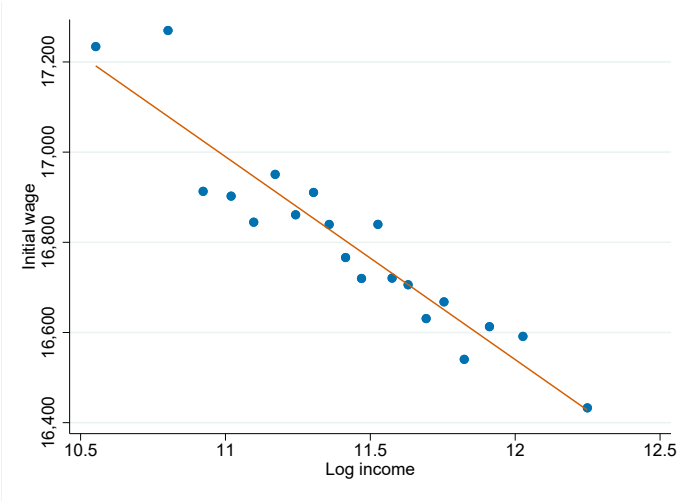
(C) Initial wage (residualized)



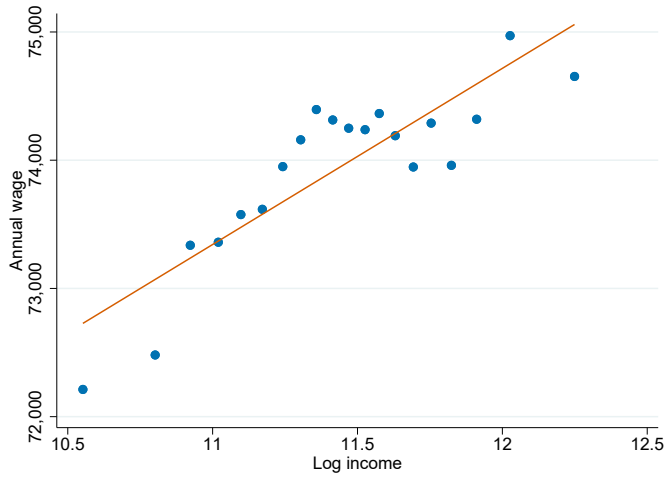
(D) Average lifetime wage (residualized)

Note: This figure describes the relationship between the parents' credit scores and the type of major that the student chose in college. Panels (A) and (B) plot binned scatter plots of the raw data, and both panels display the the parents' credit scores on the x-axis. In Panel (A) the y-axis plots the initial wage for the major that the student chose, and Panel (B) plots the average lifetime wage of the major that student chose on the y-axis. Panels (C) and (D) plot binned scatter plots of residualized wages and credit scores when controlling for race, ethnicity, gender, year, and school fixed effects. Panel (C) plot the initial wage relative to credit scores controlling for the fixed effects, and Panel (D) plot the average lifetime wage controlling for the fixed effects. In each panel, the dots represent 20 equal-sized bins based on the variable on the x-axis, and the solid line is a linear regression on the entire dataset. In Panels (C) and (D) the transparent bars represent the 95% confidence interval, and the coefficients are displayed (and the t statistic in parentheses with the standard errors are clustered at the school level). The data source is the American Community Survey (ACS) and the Commencement Program Database merged with Credit Bureau records.

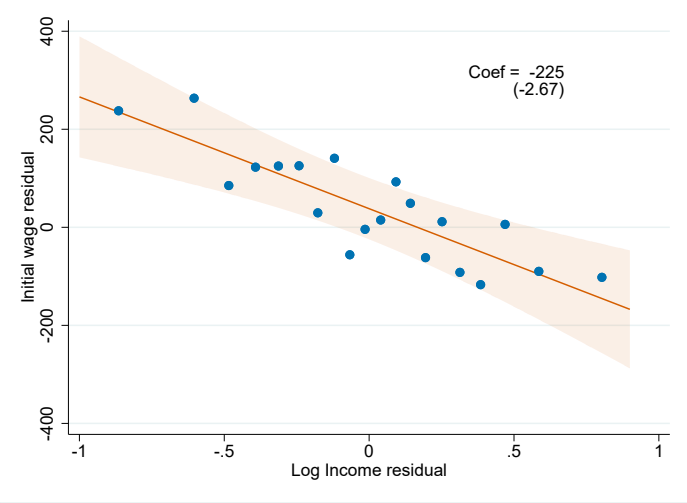
Figure 4: Major Choice and Parents' Income



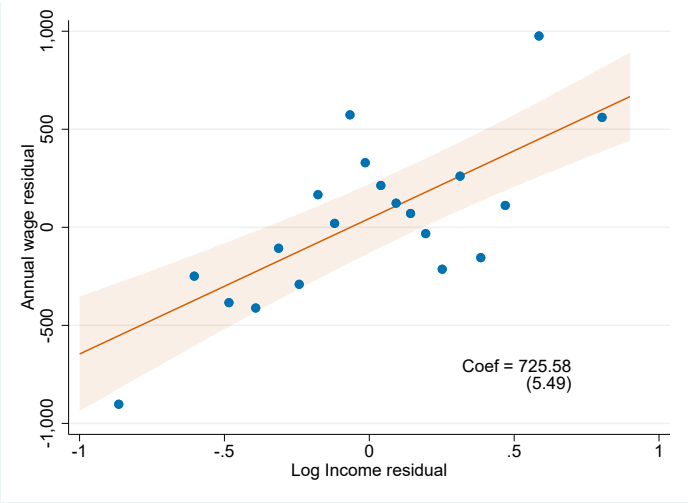
(A) Initial wage



(B) Average lifetime wage



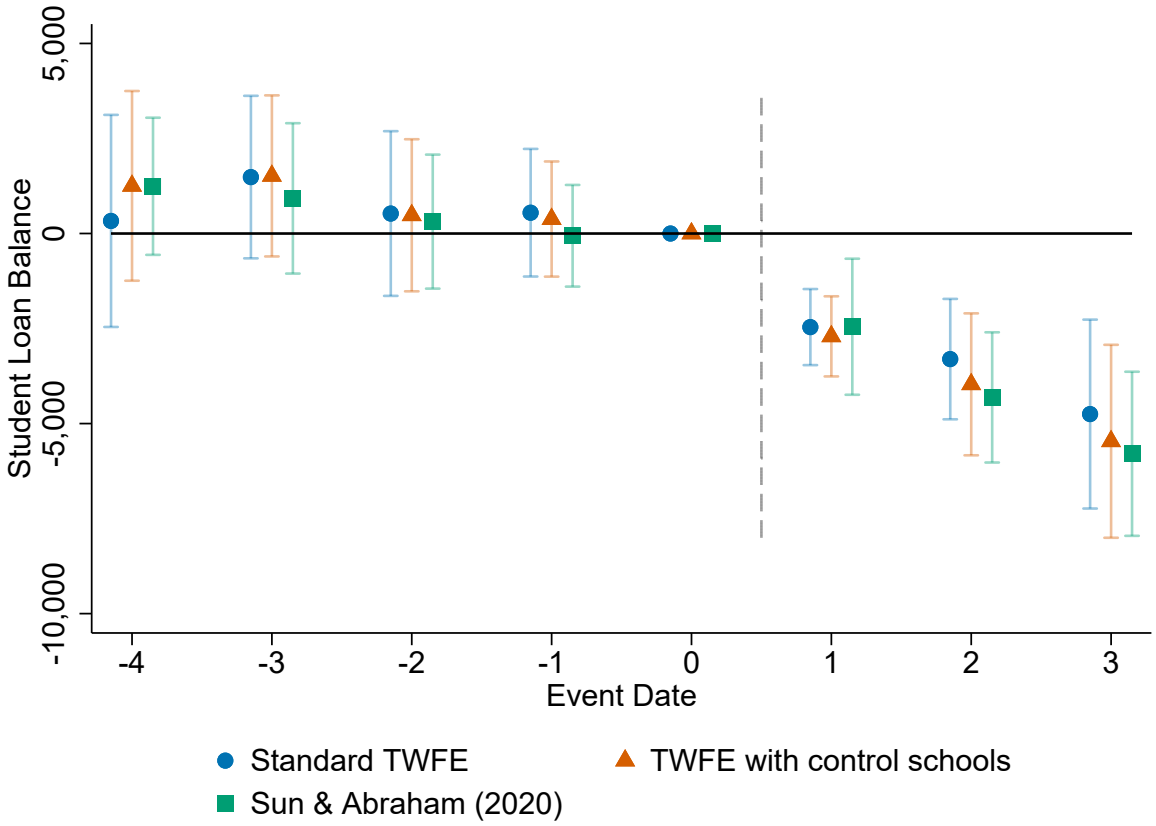
(C) Initial wage (residualized)



(D) Average lifetime wage (residualized)

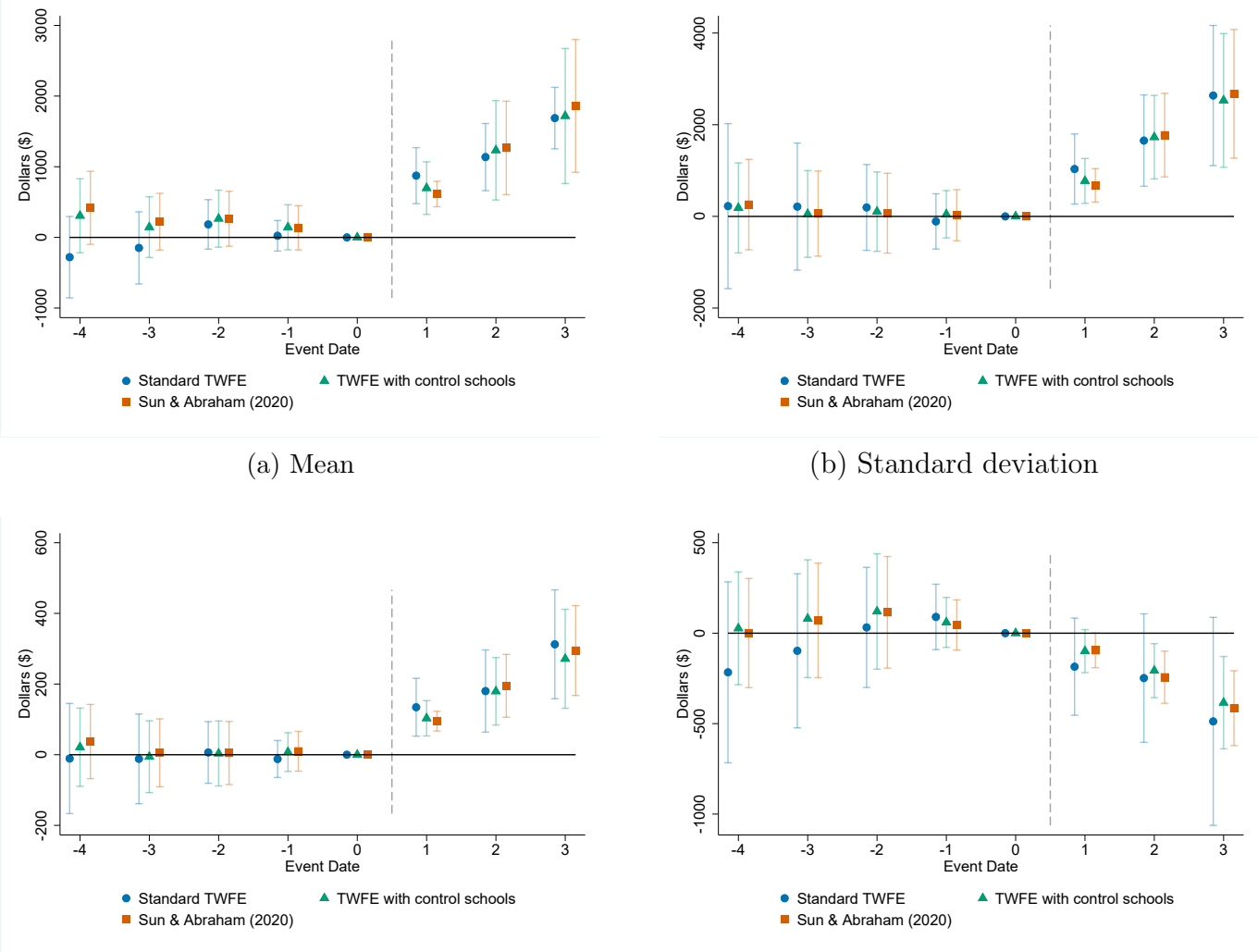
Note: This figure describes the relationship between the neighborhood that a student grew up in and the type of major that the student chose in college. Panels (A) and (B) plot binned scatter plots of the raw data, and both panels display the natural logarithm of the median income in the ZIP code where the parents' live on the x-axis. In Panel (A) the y-axis plots the initial wage for the major that the student chose, and Panel (B) plots the average lifetime wage of the major that student chose on the y-axis. Panels (C) and (D) plot binned scatter plots of residualized wages and income when controlling for race, ethnicity, gender, year, and school fixed effects. Panel (C) plot the initial wage relative to log income controlling for the fixed effects, and Panel (D) plot the average lifetime wage controlling for the fixed effects. In each panel, the dots represent 20 equal-sized bins based on the variable on the x-axis, and the solid line is a linear regression on the entire dataset. In Panels (C) and (D) the transparent bars represent the 95% confidence interval, and the coefficients are displayed (and the t statistic in parentheses with the standard errors are clustered at the school level). The data source is the American Community Survey (ACS) and the Commencement Program Database merged with Credit Bureau records.

Figure 5: Effect of UNLPs on amount of student debt



Note: This figure shows the regression coefficients from an event-study regression where the unit of observation is a student. The data source is hand-collected commencement programs merged with individual credit bureau records. The dependent variable is the student loan balance in the year of graduation. The main independent variable is a set of event-time dummies indicating when the student graduated relative to the implementation of the UNLP. A standard two-way fixed effect (TWFE) model without control schools is reported in blue and a TWFE model with non-implementing top-50 schools as a control group is reported in green, and a bias-corrected model allowing for treatment heterogenity across cohorts following Sun & Abraham (2020) is reported in red.

Figure 6: Effects of UNLPs on major wage characteristics



(a) Mean

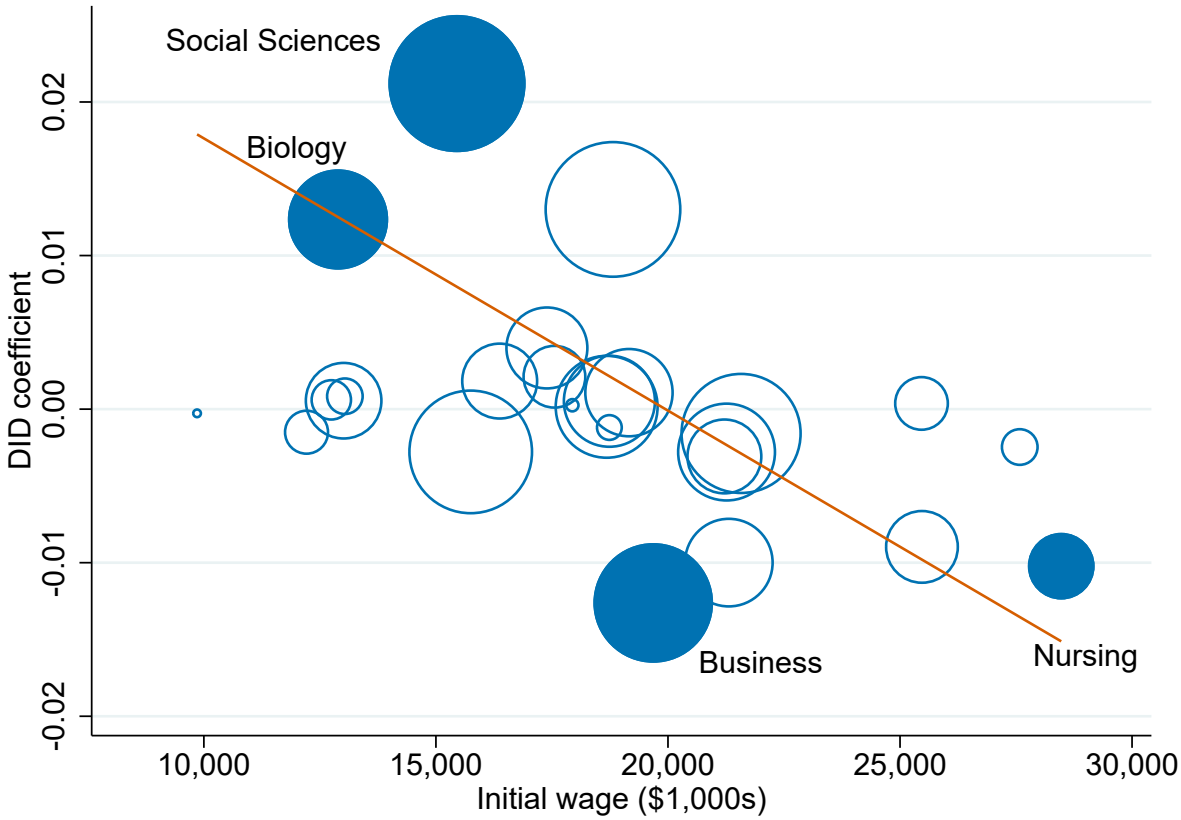
(b) Standard deviation

(c) Slope

(d) Initial wage

Note: This figure shows the regression coefficients from an event-study regression where the unit of observation is a student. The data source is hand-collected commencement programs. In Panel (a), the dependent variable is the mean across undergraduate majors of the residuals from the Mincer wage regression. In Panel (b), the dependent variable is the standard deviation of residuals. In Panel (c), the dependent variable is the slope. In Panel (d), the dependent variable is the initial wage. For all panels, the main independent variable is a set of event-time dummies indicating when the student graduated relative to the implementation of the UNLP. A standard two-way fixed effect (TWFE) model without control schools is reported in blue and a TWFE model with non-implementing top-50 schools as a control group is reported in green, and a bias-corrected model allowing for treatment heterogenity across cohorts following Sun & Abraham (2020) is reported in red.

Figure 7: Majors after NLP implementation



Note: This figure shows a scatter plot where the y-axis plots the regression coefficients from Equation (7), with the outcome variable being a dummy variable that takes the value 1 if an individual graduated with a particular major. The coefficient represents the change in the fraction of students who chose that major, controlling for student characteristics and for school and cohort fixed effects. The x-axis plots the residualized initial wage for the major. The solid line represents the best fitted line.

Tables

Table 1: Universal No-Loan Policies

	School	Implementation Year	Year Abandoned
1	Amherst College	2008	
2	Bowdoin College	2008	
3	Brown University	2018	
4	Claremont McKenna College	2008	2014
5	Colby College	2008	
6	College of the Ozarks	2013	
7	Columbia University	2008	
8	Dartmouth College	2008	2011
9	Davidson College	2007	
10	Grinnell College	2021	
11	Harvard University	2008	
12	Haverford College	2011	
13	John's Hopkins University	2018	
14	Northwestern University	2018	
15	Pomona College	2008	
16	Princeton University	2001	
17	Stanford University	2008	
18	Swarthmore College	2008	
19	University of Pennsylvania	2009	
20	Vanderbilt University	2009	
21	Williams College	2008	2011
22	Yale University	2008	

Note: This table provides the list of schools that implemented universal no-loan policies, their associated year of implementation, and, when relevant their ending date. The data is hand-collected.

Table 2: Summary statistics

Panel A: School-level characteristics

	i: UNLP vs. non-UNLP control schools				ii: UNLP vs. All U.S. schools			
	UNLP	non-UNLP	Diff.	Std. err.	UNLP	All U.S.	Diff.	Std. err.
Admissions rate	0.21	0.24	-0.03	0.02	0.21	0.73	-0.51	0.04***
SAT median (critical reading)	699	682	17.06	9.36	699	525	174.42	12.83***
SAT median (math)	702	701	0.92	11.38	702	526	176.40	13.72***
SAT median (writing)	701	683	17.15	12.98	701	495	206.02	22.27***
ACT median (cumulative)	31	31	-0.31	0.49	31	22	8.32	0.70***
ACT median (English)	31	32	-0.43	0.55	31	22	9.05	0.89***
ACT median (math)	31	31	-0.18	0.63	31	22	8.69	0.79***
ACT median (writing)	8	11	-2.94	2.73	8	9	-1.23	2.30
Enrollment of undergraduate students	4,279	6,649	-2,370	1,400	4,279	2,237	2,042	1,342
Share of students who are white	0.40	0.46	-0.06	0.05	0.40	0.44	-0.04	0.09
Share of students who are black	0.07	0.05	0.02	0.01*	0.07	0.18	-0.11	0.07
Share of students who are Hispanic	0.09	0.09	0.01	0.01	0.09	0.15	-0.06	0.07
Share of students who are Asian	0.09	0.13	-0.04	0.03	0.09	0.03	0.06	0.02*
Average net price	24,880	26,753	-1,873	2,091	24,880	18,335	6,545	2,815*
Average net price for $0-$30k fam. inc.	10,774	9,722	1,052	2,006	10,774	17,043	-6,270	2,615*
Average net price for $30k-$48k fam. inc.	10,982	10,370	612	1,910	10,982	17,946	-6,964	2,688**
Average net price for $48k-$75k fam. inc.	15,324	15,462	-138	1,921	15,324	19,925	-4,601	2,705
Average net price for $75k-$110k fam. inc.	22,843	23,898	-1,055	1,786	22,843	22,157	686	2,703
Average net price for $110k+ fam. inc.	42,289	41,893	396	1,497	42,289	23,803	18,486	3,202***
In-state tuition and fees	33,469	34,272	-802	2,941	33,469	11,626	21,843	1,735***
Out-of-state tuition and fees	33,862	37,656	-3,794	1,931	33,862	13,297	20,565	1,618***
Net tuition revenue per full-time eq. student	17,769	21,632	-3,863	2,798	17,769	16,207	1,563	27,586
Average faculty salary	10,156	11,018	-861	443	10,156	5,224	4,932	407***
Proportion of faculty that is full-time	0.81	0.81	-0.00	0.03	0.81	0.59	0.21	0.06***
Pct. of undergraduates with a Pell Grant	0.21	0.15	0.06	0.04	0.21	0.50	-0.29	0.07***
Completion rate	0.90	0.89	0.00	0.02	0.90	0.45	0.44	0.04***

Note: This table reports summary statistics and covariate balance t-test for comparing universities that implemented UNLPs with the top-50 control group. For each variable, I remove a year fixed effect and add back the sample mean. The data source is the Integrated Postsecondary Education Data System (IPEDS) and the College Scorecard. Significance level: *p<0.1; **p<0.05; ***p<0.01.

Table 2 (continued): Summary statistics

Panel B: Demographic summary statistics of regression sample vs full sample

	Regression sample (N=757,052)		Full sample (N=895,292)	
	Mean	Std. dev.	Mean	Std. dev.
Gender				
Female (pct.)	46.31	49.86	45.96	49.84
Male (pct.)	43.66	49.60	43.92	49.63
Race				
Native American (pct.)	0.22	4.68	0.22	4.68
Asian (pct.)	3.87	19.28	3.84	19.22
Black (pct.)	4.60	20.94	4.73	21.22
Hispanic (pct.)	2.39	15.27	2.50	15.62
White (pct.)	69.73	45.94	69.44	46.07

Panel C: Comparing UNLP vs non-UNLP schools

	Regression sample (N=757,052)				Full sample (N=895,292)			
	UNLP (N=232,199)	non-UNLP (N=524,853)	Diff.	Std. err.	UNLP (N=370,439)	non-UNLP (N=524,853)	Diff.	Std. err.
Gender								
Female (pct.)	44.33	47.18	-2.85	0.12***	44.24	47.18	-2.95	0.11***
Male (pct.)	45.33	42.91	2.42	0.12***	45.35	42.91	2.44	0.11***
Race								
Native American (pct.)	0.19	0.23	-0.04	0.01***	0.20	0.23	-0.03	0.01**
Asian (pct.)	3.89	3.85	0.04	0.05	3.82	3.85	-0.03	0.04
Black (pct.)	5.15	4.35	0.79	0.05***	5.25	4.35	0.90	0.05***
Hispanic (pct.)	2.90	2.16	0.73	0.04***	2.98	2.16	0.82	0.03***
White (pct.)	68.66	70.21	-1.54	0.11***	68.36	70.21	-1.85	0.10***

Note: This table reports demographic summary statistics for the students. Panel B compares gender and race for the regression sample and the full sample. The full sample includes all students. The regression sample restricts the full sample by excluding students who enrolled after the policy was implemented. Panel C compares students who attended UNLP schools and those who attended non-UNLP top-50 schools across the regression sample and the full sample. Significance level: *p<0.1; **p<0.05; ***p<0.01.

Table 3: Proof of Concept with Publicly Available Data

	Dependent variable: Pct. of Undergraduates with Student Debt				
	No-Loan Programs		Other Programs		
	Universal	Targeted	Loan Cap	No Par. Contr.	No Tuition
	(1)	(2)	(3)	(4)	(5)
Financial Aid Program	-12.61^{***}	0.36	2.58^{***}	-1.69	0.10
	(1.34)	(1.07)	(0.77)	(1.13)	(2.63)
Year FE	✓	✓	✓	✓	✓
School FE	✓	✓	✓	✓	✓
Observations	324	972	641	360	378
R^2	0.756	0.636	0.668	0.853	0.755
Adjusted R^2	0.726	0.611	0.647	0.841	0.731

Note: This table reports regression coefficients from regressing the percentage of undergraduates with student debt at school s in year t on an indicator taking the value 1 if school s has implemented a given financial aid policy in year t, controlling for school and year fixed effects. Column 1 reports the coefficient when the policy is a UNLP applying to all students. Column 2 reports the coefficient when the policy is an NLP restricted to low-income students. Columns 3–5 report the results for loan caps, parental contribution waivers, and tuition waivers, respectively. The data source is the Integrated Postsecondary Education Data System (IPEDS). Standard errors (in parentheses) are clustered at the school-year level, and the p-values are as follows: *p<0.1; **p<0.05; ***p<0.01.

Table 4: The Effect of UNLP on Major Choice (DID)

	Mean life-time wage	Slope	Initial wage	Year bins		
	(1)	(2)	(3)	(4) 20s	(5) 30s	(6) 40s
UNLP $\times Post$	1,161	0.65	-266.0	-147.6	834.2	1497.1
	(2.38)	(3.07)	(-2.14)	(-0.93)	(2.43)	(2.65)
Control schools	✓	✓	✓	✓	✓	✓
School FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Demographics	✓	✓	✓	✓	✓	✓
N	720,885	720,885	720,885	720,885	720,885	720,885

Note: This table reports regression coefficients from a difference-in-differences specification. Column 1 give the coefficients for the mean life-time wage, Column 3 the coefficient for the slope, Column 3 the coefficients for the initial wage, and Columns 4, 5, and 6 the coefficients for the wage in different age decades. The t-statistics are reported in parentheses, and the standard errors are clustered at the school level. The data sources are hand-collected commencement programs and the ACS.

Table 5: The Effect of UNLP on Major Choice (IV)

	OLS		IV					
	Loan amount		Mean life-time wage	Slope	Initial wage	Year bins		
						20s	30s	40s
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	-4,878							
	(-6.47)							
Intensity		-9,272						
		(-5.96)						
Loan amount			-190.46	-0.12	46.35	22.69	-136.80	-260.85
			(-2.99)	(-2.38)	(2.41)	(1.01)	(-2.31)	(-3.07)
School FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Demographics	✓	✓	✓	✓	✓	✓	✓	✓
N	227,429	227,429	227,429	227,429	227,429	227,429	227,429	227,429
F-Stat	118.04							

Note: This table reports regression coefficients from the IV regression model. Column 1 give the coefficient from the first stage regression which regresses student debt loan amount on a post dummy. Column 2 gives the OLS coefficient regressing student debt loan amount on the treatment intensity. Columns 3 through 8 report the second-stage coefficients (using the predicted loan amount from Column 1) on mean life-time wage, wage slope, initial wage, and average annual wage in age decades. In Columns 3-8, the coefficients are scaled to represent the effect of a $1,000 increase in predicted loan amount. The t-statistics are reported in parentheses, and the standard errors are clustered at the school level. The data sources are hand-collected commencement programs, the Equifax credit bureau, and the ACS.

Table 6: Treatment Effects by Family Characteristics I: Neighborhood

	Mean life-time wage	Slope	Initial wage
	(1)	(2)	(3)
Post	447	0.32	-67.8
	(1.25)	(2.43)	(-1.21)
Below-Median Family Income (ZIP)	-1,050	-0.91	236.8
	(-2.60)	(-4.13)	(2.50)
Post × Below-Median Family Income (ZIP)	463	0.35	-106.6
	(2.24)	(2.56)	(-2.14)
School FE	✓	✓	✓
Year FE	✓	✓	✓
Demographics	✓	✓	✓
N	227,429	227,429	227,429

Note: This table reports regression coefficients from a difference-in-differences specification. Post is a dummy indicating whether a student has been treated. Below-Median Family Income (ZIP) is an indicator taking the value 1 if the student's family lives in a ZIP Code where the median household income is below the median. Column 1 reports the coefficients for the mean life-time wage by major choice, Column 2 for the slope, and Column 3 for the initial wage. The t-statistics are reported in parentheses, and the standard errors are clustered at the school level. The data sources are hand-collected commencement programs, the credit bureau reports, and the ACS.

Table 7: Treatment Effects by Family Characteristics II: Credit Score

	Mean life-time wage	Slope	Initial wage
	(1)	(2)	(3)
Post	425	0.36	-53.6
	(1.00)	(2.83)	(-1.12)
Below-Median Vantage Score	-973	-0.89	234.6
	(-2.72)	(-3.75)	(2.94)
Post × Below-Median Vantage Score	408	0.38	-93.0
	(2.04)	(2.75)	(-2.66)
School FE	✓	✓	✓
Year FE	✓	✓	✓
Demographics	✓	✓	✓
N	227,429	227,429	227,429

Note: This table reports regression coefficients from a difference-in-differences specification. Post is a dummy indicating whether a student has been treated. Below-Median Vantage Score is an indicator taking the value 1 if the average Vantage Score of the student's parents is below the median. Column 1 reports the coefficients for the mean life-time wage by major choice, Column 2 for the slope, and Column 3 for the initial wage. The t-statistics are reported in parentheses, and the standard errors are clustered at the school level. The data sources are hand-collected commencement programs, the credit bureau reports, and the ACS.

Online Appendix

Contents

A Theoretical Appendix

A.I Conceptual Framework

In this section, I describe in three steps the theoretical framework that guides my empirical tests of the effects of student debt on career trajectories. First, I set up a standard buffer-stock consumption–savings model for the problem that a student faces when choosing a college major. I then derive the closed-form solution for the model. Finally, guided by the model, I outline the implications for the empirical tests.

A.I.1 Model

The model is a standard buffer-stock model, where a student i chooses her job j and her savings a_1 to maximize utility from consumption subject to a per-period budget constraint and a borrowing constraint. There are two periods: $t \in \{0, 1\}$. The period $t = 0$ corresponds to the first five years of the student's career after college, and $t = 1$ corresponds to the rest of her career. The student chooses between two jobs $j \in \{0, 1\}$, where $j = 0$ represents a low-skilled job that any college graduate could fill without further investment in human capital, and $j = 1$ represents a high-skilled job that requires additional human capital investments. For tractability, I assume that utility is logarithmic, that there is no discounting, that there is student debt repayment in the first period only, and that the borrowing constraint is zero. The problem faced by the student is then

$$\max_{j,a_1} \ \ln(c_0) + \ln(c_1) \qquad \text{s.t.} \qquad a_1 = (y_0 - c_0 - d^i - j\theta^i) \quad \text{and} \quad a_1 \geq 0, \tag{14}$$

where c_t denotes consumption, and a_1 denotes liquid assets that earn a per-period interest of r and are subject to a borrowing constraint of 0. The term d^i denotes the student debt payments; the superscript i indicates that d^i varies exogenously across students.

For simplicity, I assume that there is no uncertainty, and that the two jobs have the same job amenities. This means that if the student chooses the low-skilled job, $j = 0$, then she pays no investment cost in the first period, and her wage in the second period is w_0. If she chooses the high-skilled job, $j = 1$, then she pays an investment cost of θ^i in the first period, and her wage in the second period is w_1.

For tractability, I assume that $\theta^i > 0$, $w_1 > w_0$, and $(1+r)y_0 < w_0$. This ensures that the low-skilled job has a higher wage in the first period, that the high-skilled job has a higher wage in the second period, and that the credit constraint is binding. Finally, I assume that $y_0 > d^i$ to ensure positive consumption.

We can now solve the model in closed form. The solution is characterized by the following

threshold condition: The student will choose to invest θ^i to get the high-skilled job, $j = 1$, if and only if

$$\theta^i < \left(\frac{w_1 - w_0}{w_1}\right)(y_0 - d^i). \tag{15}$$

The inequality (15) states that the optimal choice of j^* depends on the relative wage difference $w_1 - w_0$, the initial resources y_0, the investment cost θ^i, and the amount of student debt d^i. Furthermore, we see that the probability of choosing $j = 1$ (i.e., investing in human capital and getting the high-skilled job) is increasing with respect to relative wage difference and initial resources, and decreasing with respect to investment cost and the amount of student debt:

$$j^* = j\left(\underbrace{w_1 - w_0}_{+}, \underbrace{y_0}_{+}, \underbrace{\theta^i}_{-}, \underbrace{d^i}_{-}\right). \tag{16}$$

A.I.2 Accounting for Wage Uncertainty

This model is similar to the one above, except there is time-discounting and uncertainty over the wages in the first period. As before, there are two periods: $t \in \{0, 1\}$. The period $t = 0$ corresponds to the first five years of the student's career after college, and $t = 1$ corresponds to the rest of her career. The student chooses between two jobs $j \in \{0, 1\}$, where $j = 0$ represents a low-skilled job that any college graduate could fill without further investment in human capital, and $j = 1$ represents a high-skilled job that requires additional human capital investments.

Relative to the benchmark model, I now take into account a discount rate β, and the wages are uncertain. Specifically, the low-skilled job has a wage of w_{0H} with probability π_0 and a wage of w_{0L} with probability $1 - \pi_0$, where $w_{0H} > w_{0L}$. Similarly, the high-skilled job has a wage of w_{1H} with probability π_1 and a wage of w_{1L} with probability $1 - \pi_1$, where $w_{1H} > w_{1L}$.

The maximization problem is then

$$\max_{j, a_1} \ln(c_0) + \beta\mathbb{E}[\ln(c_1)] \qquad \text{s.t.} \qquad a_1 = (y_0 - c_0 - d^i - j\theta^i) \quad \text{and} \quad a_1 \geq 0. \tag{17}$$

The budget constraint can be described as the consumption choices given each choice and

each state of the world:

$$
\begin{aligned}
& & & c_0 = y_0 - D^i - j\theta^i - a_1; \\
& \text{if } j = 0, & \text{w. prob. } \pi_0 : & \quad c_1 = w_{0H} + a_1(1+r); \\
& & \text{w. prob. } 1-\pi_0 : & \quad c_1 = w_{0L} + a_1(1+r); \\
& \text{if } j = 1, & \text{w. prob. } \pi_1 : & \quad c_1 = w_{1H} + a_1(1+r); \\
& & \text{w. prob. } 1-\pi_1 : & \quad c_1 = w_{1L} + a_1(1+r).
\end{aligned}
$$

We can now solve the model in closed form. The solution is characterized by the following threshold condition: The student will choose to invest θ^i to get the high-skilled job, $j = 1$, if and only if

$$\theta^i \leq (y_0 - d^i)\frac{w_{1,H}^{\beta\pi_1} w_{1,L}^{\beta(1-\pi_1)} - w_{0,H}^{\beta\pi_0} w_{0,L}^{\beta(1-\pi_0)})}{w_{1,H}^{\beta\pi_1} w_{1,L}^{\beta\pi_1}}. \tag{18}$$

The inequality (18) states that the optimal choice of j^* depends on the initial resources y_0, the investment cost θ^i, the amount of student debt d^i, and the relative wage difference scaled by their probabilities. Furthermore, we see that the probability of choosing $j = 1$ (i.e., investing in human capital and getting the high-skilled job) is increasing with respect to initial resources, and decreasing with respect to investment cost and the amount of student debt:

$$j^* = j\left(\underbrace{y_0}_{+}, \underbrace{\theta^i}_{-}, \underbrace{d^i}_{-}\right). \tag{19}$$

A.I.3 Empirical Hypotheses and Discussion

From the benchmark model, I can present my three empirical hypotheses. The first two hypotheses test the average effects of student debt and family characteristics on job choices, whereas the third hypothesis tests the heterogeneous effects of student debt across family characteristics.

For the first hypothesis, we draw on the empirical relationship that high-skilled jobs have (a) higher average lifetime earnings, (b) higher variability in earnings, and (c) higher earnings growth.

We then notice that from Equations (18) and (19), higher student debt leads to a more restrictive condition for the investment choice; in other words, the derivative of j with respect to d^i is negative: $\frac{\partial j}{\partial d^i} < 0$. Hypothesis 1 tests this relationship in the data:

Hypothesis 1 (H1.A): *Higher student debt decreases the probability of choosing a job with higher average earnings.*

Hypothesis 1 (H1.B): *Higher student debt decreases the probability of choosing a job with more variable earnings.*

Hypothesis 1 (H1.C): *Higher student debt decreases the probability of choosing a job with a steep earnings path.*

For the next hypothesis, notice that from Equations (15) and (16), a higher investment cost leads to a more restrictive threshold; that is, the derivative of j with respect to θ^i is negative: $\frac{\partial j}{\partial \theta^i} < 0$. Hypothesis 2 tests this relationship in the data:

Hypothesis 2 (H2): *Students' average job choices (e.g. whether they choose jobs with flat or steep earnings paths) differ across family characteristics such as their parents' incomes, assets, credit scores, and debt-to-income ratios.*

For the third and final hypothesis, I ask: Does the effect of student debt on job choices differ based on family characteristics? In terms of the model, I am estimating the cross-derivative, $\frac{\partial^2 j}{\partial d^i \partial \theta^i}$.

Hypothesis 3 (H3): *The effect of student debt on job choices varies with the demographic and financial characteristics of the student and her parents.*

Testing H3 is the main contribution of this paper. Although no consensus exists yet, there is already a nascent literature studying the effects of student debt on labor choices (H1) and the variation of labor choices across family characteristics (H2). In this project I go further, studying not just the effects of student debt for the average student, but also how these effects vary with family characteristics.

Importantly, note that H3 is two-sided. In other words, I have not taken a stand on whether the cross-derivative between student debt and family characteristics is positive, negative, or even monotonic. It is theoretically unclear whether the costs (and benefits) of investing in human capital are higher for high-income or low-income families. Recall that the choice of job j occurs simultaneously with the consumption choice c. One extra dollar

of student debt can mean either less consumption or less investment in human capital:

$$\underbrace{\frac{\partial c}{\partial d^i}}_{\text{Marginal propensity to consume (MPC)}} \qquad \underbrace{\frac{\partial j}{\partial d^i}}_{\text{Marginal propensity to invest (MPI)}}$$

There is a longstanding literature on household finance examining the marginal propensity to consume (MPC).[30] A key result in this literature is that the MPC covaries positively with credit constraints (Johnson et al., 2006; Jappelli and Pistaferri, 2014; Baker, 2018). For example, low-income households with few liquid assets or high credit card debt display higher MPCs. On the other hand, it is not obvious that the same is true for the marginal propensity to invest (MPI). For example, if human capital is a luxury good, then we might expect that a student from a wealthy family is more likely to invest in human capital when she receives more financial aid, and that a student from a low-income family might have a high MPC and choose to consume the extra financial aid rather than invest it. That is, it is theoretically ambiguous whether the MPI is increasing or decreasing with respect to household income:

$$\frac{\partial \text{MPC}}{\partial \text{household income}} < 0; \qquad \frac{\partial \text{MPI}}{\partial \text{household income}} \overset{?}{\lesseqgtr} 0.$$

In sum, the model of Subsection A.I.2 provides a theoretical framework that guides my empirical tests. The model highlights that both student debt and exogenous investment costs may lead students to choose careers with higher initial earnings and lower earnings growth. Furthermore, and critically for this project, the model highlights that the effects of student debt depend on the costs (and benefits) of investing in human capital. And these costs might vary substantially across the student population.

[30] Going back, at least, to Friedman (1957) and Hall (1978), economists have studied how changes in income affect consumer spending. Seminal theoretical work by Deaton (1991) and Carroll (1997) emphasizes the role of borrowing constraints, and a large empirical literature has identified certain spending responses to predictable changes in income (Zeldes, 1989; Parker, 1999; Souleles, 1999; Agarwal et al., 2007; Di Maggio et al., 2017; Baker and Yannelis, 2017). See Jørring (2020) for a recent review.

B Data Collection

B.I Description of College Commencement Programs, Yearbooks, and Degree Conferral Rosters

I collect and OCR university commencement programs, yearbooks, and degree conferral rosters for the 50 schools in my sample. The commencement program contains five sources of individual-level information for cohorts graduating between 1996 and 2022: (1) full name which includes first name, middle name, and last name, (2) realized major choices, (3) realized minor choices, (4) cities, states, and countries of origin, (5) awards that include phi beta kappa and cum laude awards. If the commencement programs did not contain some of the information above, I supplement by using archived yearbooks that contain overlapping information. This is also used as a method of verifying graduation dates and majors for students in the sample.

B.II Description of LinkedIn, Doximity, and CV Data

I collected publically available CVs from LinkedIn, Doximity, and professional websites.

LinkedIn is a social media platform used primarily for professional networking. It allows job seekers to post their CVs and employers to post jobs. The platform is widely used with over 740 million members worldwide in 2021. In the United States, there are over 170 million users (Osman, 2021). It is popular among professionals with 50% of the users holding a college degree (Osman, 2021).

Doximity is the leading professional network in medicine. Doximity members include over 2 million healthcare professionals within the United States: 80% of Physicians, 50% of Nurse Practitioners and Physician Assistants, and 90% of graduating Medical Students.

Users create online public profiles that contain CV information. This contains information on all previous work experience, including job title, employer, location, job descriptions, start and end dates. Individuals can also post education and training, skills, and a personal photo. In addition, individuals can connect with other users on the platform in an online social network. I collected this data twice, first in June of 2018 and second in June of 2022. As a result, the CV information is current up to 2022.

The data is cleaned by parsing the information on the CV and reshaping the data such that a quarterly panel is created based on the start and end dates of employment. I expand the data to include observations for when someone is in nonemployment, i.e. periods where there is a gap on the CV between the start and end dates of two consecutive positions. I then collapse the dataset to the yearly level. For each year, individuals are assigned to the

position in which they have spent the most time during that year. If there are ties, the position with the longer tenure takes precedence.

B.III United States Postal Service Records

I collect a history of postal addresses for each student in my sample. I use two prominent data aggregators called Intelius and BeenVerified in order to do so.

Intelius, Inc. is a public records business that provides information services, including people and property search, background checks and reverse phone lookup. Users also have the ability to perform reverse address lookups to find people using Intelius' services and an address.

BeenVerified is a background check company that provides background checks and people search services. BeenVerified uses traditional background methods in addition to Web 2.0 and social networking to return results to a requesting user.Users enter the name and/or email address of the person they are requesting information on and are given information from public records and other privately licensed databases of public record information.

The search process is as follows. Each student's full name and approximate age, based on cohort, is fed through a search algorithm that returns one or many results. Each search result is then manually verified by a research assistant using a unique sequence of locations coming from hometown, university locations, and locations from LinkedIn. This process yields a unique match.

These data providers also use an internal algorithm to assess possible relatives. This is a based on address overlaps and last names. This allows me to construct a household of possible relatives. I cross-reference both BeenVerified and Intelius to verify a household.

C Data Matching

C.I Mapping Onet Codes to Occupations in CVs

This Section describes the data matching process between ONET codes and occupations in the resume data.

C.I.1 Pre-processing

Given job title strings extracted from parsed LinkedIn webpages data, the text strings are preprocessed (1) with hard conditional rules; (2) with a constructed dictionary.

- *Hard processing.* First, the hard conditional rules rid the text strings off punctuation and delimiters such as " - " and "/", as these tokens pose no informational increment for the matching process; furthermore, for some scrapped strings in non-English languages, such as French, translation is imposed on these strings.

- *Dictionary-based processing.* Secondly, the dictionary is composed of two parts, and updated dynamically. The first part of the dictionary outlines junk tokens that conveys no information in text strings *e.g.* "senior", "sr.", etc. The second part of the dictionary characterizes tokens that tend to be ambiguous *e.g.* "advocate", "associate", etc. For the junk tokens, they are removed from the text string, and for the text strings containing indeterminate tokens, they are skipped and stored as residual samples for next step processing.

- *Dictionary update.* For each match sample that's generated *and* labeled to be a good or bad match, the sample is stored in the dictionary-generating dataset (The labeling process is detailed in later sections). Consequently, for each pre-processing call, a updated dictionary is generated by scanning all matches labeled a bad match, and identify tokens with aforementioned undesirable characteristics. This step is done by tokenizing all text strings labeled bad, computing token frequencies, and adding tokens with highest frequencies in the bad matches to the updated dictionary.

C.I.2 Matching Algorithm

The matching process uses the modified ONET API. After pre-processing, the processed job title text strings are batched and uploaded to ONET database for the best match as determined by ONET Web Service search. However, as the returned best search are sometimes inadequate, an checking process is devised to determine whether the returned search is an adequate match for the true information as conveyed by the LinkedIn text string.

C.I.3 Robot Checker

To train a robot checker for purpose stated above, a sample of approximately 8700 samples are selected for manual labeling with match adequacy determined by human eye to serve as training set.

- *Feature generation.* To generate features for model training, we first pull alternative job titles of the best-match ONET title from the API *e.g.* job function "Lawyers" is also known as 'Attorney', 'Attorney at Law', 'Attorney General', etc.

  $$\text{“Lawyers”} \xrightarrow{\text{is also known as}} [\text{“Attorney”, “Attorney at Law”, “Attorney General”}]$$

  according to ONET API. The intuition behind this step is that, should the LinkedIn title be a right match, then the title should be able to at some level fuzzy-match with some of these alternative titles as well. Using different fuzzy-matching functions, such as regular fuzzy ratio (*i.e.* Levenshtein distance), partial fuzzy ratio, token set ratio, token sort ratio, etc., the input argument of (LinkdeIn title, alternative title list) produces a probability vector scaled by 100 *e.g.*

  $$\texttt{fuzzy\_func(``attorney'', titleList)} \xrightarrow{\textit{outputs}} [88, 61, 58, 67, 40, 45, 17, 29]$$

  where 0 is the lowest possible fuzzy score and 100 is the highest. Finally, on different score vectors generated by the fuzzy function, we apply statistical aggregate functions, such as standard deviation, mean and maximum. This process results in a total number of (Num of func)$\times$(Num of Stats) different features

  Note to author: Othre features such as ONET code first two digits to be added

- *Model: Tree-based classifier.* For the purpose of identifying whether the match by ONET is adequate or not, we employ a tree-based classifier, such as Random Forest or XGBoost library (simple decision tree is ignored as random forest ensemble tends to vastly outperform decision tree). The intuition behind this decision is easy to understand: The magnitude (high or low) of the generated statistical aggregates from different fuzzy functions must convey some information regarding the adequacy of the match, even if the aggregate cannot be easily interpreted by a human. Hence, we use decision trees to interpret these indicators, and let the trees determine pertinent threshold values for the model.

C.II Mapping CIP codes to Majors

After collecting commencement programs with student majors, we had to standardize them. We did this by assigning Classification of Instruction Program codes (CIP) to each major from the National Center of Education Statistics in a master list. The master list was used in a mapping process to match majors and CIP codes, creating a standardized list of majors per institution. The mapping process would have holes when schools had niche or interdisciplinary majors. In this circumstance, we would manually cross-reference the major's description to the National Center for Education Statistics to best identify a CIP code. For example, Vanderbilt had a major called "INTERDISCIPLINARY: CORPORATE LEADERSHIP AND FINANCE." Looking further into Vanderbilt's course offerings, we closely aligned each interdisciplinary major to something that fell under the CIP code(s). Sometimes to best capture the interdisciplinary major, we would include two CIP codes, like business management and finance for the example above.

C.III Matching between Commencement Programs and CV Profile

Starting from the commencement programs for classes 2004 to 2022, I collect publicly available CVs from Linkedin, Doximity, and personal pages. I matched students to their online profiles based on the following variables:

- Full name
- Undergraduate school name
- Year of graduation

I require names, undergraduate school and class year to match perfectly to be considered a match. For women, I require only the first names to match and we conduct an online verification for those that may have changed their last names due to marriage (e.g., a wedding registry webpage). This verification process also includes alternative names listed in the USPS data.

C.IV Linking between Credit Bureau data and Hand-collected data

This project was reviewed by Boston College's IRB and determined to fall under a Category 4 Exemption. This project falls under a research exemption for purposes of FERPA.

The linking procedure between the hand-collected data that I provided and Experian Credit Bureau followed the below procedure:

- First, an independent third party created a random, study-specific identifier. They then provided Experian with personally identifiable information (PII) necessary for linking and the random identifier.
- Experian queried their data and created an anonymized, individual-level dataset that contained the random identifier and credit bureau records.
- The anonymized dataset was 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.

C.V Assigning Gender and Ethnicity

I obtain gender and ethnicity through algorithmic assignments.

For gender, I use "genderize" which is an R function. This function predicts the gender of a first name given a year or range of years in which the person was born. The genderize functions works by the "ssa" method looks up names based from the U.S. Social Security Administration baby name data. The "ipums" method looks up names from the U.S. Census data in the Integrated Public Use Microdata Series. The "napp" method uses census microdata from Canada, Great Britain, Denmark, Iceland, Norway, and Sweden from 1801 to 1910 created by the North Atlantic Population Project. The "kantrowitz" method uses the Kantrowitz corpus of male and female names.

For ethnicity, I use "predictrace," which is an R function. The goal of predictrace is to predict the race of a surname or first name and the gender of a first name. This package uses U.S. Census data which says how many people of each race has a certain surname. For first name data, this package uses data from Tzioumis (2018). From this we can predict which race is mostly likely to have that surname or first name. The possible races are American Indian, Asian, Black, Hispanic, White, or two or more races.

D Discussion of IV assumptions

In this section, I discuss the validity of the instrumental variables (IV) strategy. Specifically, I discuss and test the four central assumptions that are required to interpret the IV estimates as a causal effect of student debt.

D.I Relevance

First, I discuss the strength of the first stage. As discussed in Section 5.1, I can use both publicly available data aggregated at the school-cohort level as well as my main sample of student-level micro data to test the relevance assumption.

First, I test the assumption using publicly available data from IPEDS. Table 3 reports the DID estimates, where the dependent variable is the fraction of undergraduates with student debt, and compares UNLPs to other financial aid programs, including income-specific NLPs, loan caps, parental contribution elimination, and tuition waivers. (For example, some colleges implement NLPs for families whose incomes fall below a specific level.) I find that UNLPs meaningfully decrease the percentage of students taking loans; specifically, they lower it by 12.6 percentage points.

Next, to understand the dynamic effects, Figure C4 reports the results from a standard event-study regression. The plot shows the coefficients on the year relative to implementation. We see a sharp drop in the fraction of students taking out loans that precisely coincides with the timing of the policy. To assess the robustness of this result, in Figure C4, I plot the regression coefficients and 95% confidence intervals from three different regressions of the percentage of students taking a student loan on year dummies relative to the implementation of a UNLP. A standard two-way fixed effect (TWFE) model is reported in blue, a TWFE model with non-implementing top-50 schools as a control group is reported in red, and a bias-corrected model allowing for treatment heterogeneity across cohorts (following Sun and Abraham (2020)) is reported in green. Across all three specifications, in the year when a UNLP is implemented, there is a significant drop in the share of students who take student loans.

Second, I test the assumption using the main sample with micro-data. Figure 5 shows the coefficients from equation (6) where the dependent variable is the student loan balance in the year of graduation.[31] We see a sharp discontinues drop in the amount of student loans that students have that precisely coincides with the timing of the policy. As above, to assess the robustness of this result, I plot the regression results from both a standard TWFE, a

[31] The credit bureau data offers a yearly snapshot on December 31. Therefor, the student loan balance variable captures the amount of student loan debt on December 31 on the year that the student graduated.

TWFE with control groups, and a bias-corrected model allowing for treatment heterogeneity across cohorts. The results are similar across all three models. Table 5 Column (1) reports the DID coefficient for the standard TWFE model. Following the implementation of the UNLP students, on average, have $4,900 less in student debt.

Taken together, these results confirm that following the implementation of UNLPs, the use of student loans fell dramatically. In other words, there is a strong first stage.

D.II Independence

It is important to consider whether the implementation of a UNLP is an *independent* instrument for student debt: Does the first stage represent a causal estimate of the effect of the policy? There are two central arguments in favor. First, the staggered implementation of the policies makes it unlikely that contemporaneous macroeconomic shocks are driving the effects; additionally, with the inclusion of year fixed effects, I study the variation across schools only within each year. Second, as I include school fixed effects, I study the variation across cohorts only within the same school. Thus, any correlation between "selective" schools and low student debt is eliminated.

Additionally, it is worth noting that the assumption for an independent instrument are similar to those discussed in Sections 4.2 (the assumptions required to interpret the reduced form evidence as causal evidence.) Specifically, I rely on two key identifying assumptions: no anticipation of the treatment and parallel trends (Sun and Abraham (2021); Borusyak et al. (2021)). Under the no anticipatory effects assumption, I assume that units do not change their behavior in anticipation of the treatment.[32] The second identifying assumption is the parallel trends assumption, in which we assume that absent the reform, the difference in potential outcomes would be the same across all units and all periods conditional on the set of controls, unit and time fixed effects.[33]

The main threat to identification is that unobserved changes in the composition of students attending the UNLP school can explain both the timing of the UNLPs and changes in major choice. For example, with the implementation of the policy, the school may enroll students that are more interested in studying specific types of majors. In order to mitigate this issue, I focus on students that were enrolled at the time of the policy implementation. Specifically, I focus only on students that were sophomores, juniors, and seniors at the time the policy is implemented.

As highlighted by the recent econometric literature, the estimates may also be biased if

[32] Formally, in the notation of potential outcomes, this is equivalent to $E[Y_t - Y_t(0)|X] = 0$ for all t prior to the policy, conditional on covariates X.

[33] Formally, we assume that $E[Y_{it}(0) - Y_{it}0(0)|X]$ has to be the same across units i for all periods t, t0.

there is heterogeneity in treatment effects between groups of units treated at different times. This bias can occur even if both the no-anticipatory effects and parallel trends assumptions hold (de Chaisemartin and D'Haultfoeuille, 2020; Sun and Abraham, 2021; Goodman-Bacon, 2021; Borusyak et al., 2021; Callaway and Sant'Anna, 2021). Given the potential for biased estimates, I employ the Sun and Abraham (2020) correction methodology, which adjusts for potential treatment heterogeneity across cohorts.

Importantly for the independence assumption for the two-stage least squares, I do not find any difference between the TWFE event-study results and the bias-corrected results that correct for potential treatment heterogeneity across cohorts.

D.III Exclusion

The exclusion restriction posits that UNLPs affect outcomes only through their effects on student debt. This would not be the case if, for example, the student body changed as it became more advantageous to apply to a school after the implementation of an UNLP. I address this issue in two ways. First, I study only cohorts who were already enrolled by the time of the UNLP implementation. That is, I study only the three cohorts following the implementation of the UNLP. Second, when analyzing student characteristics, I do not find any discontinuous change in student characteristics around the time of implementation of the UNLPs.

Figures (C12) and (C13) show that there is no discontinuous change in student characteristics around the event date. Figure (C12) show average SAT scores for students for each year relative to the implementation of an NLP. There is a slight upward trend for both the 25th and 75th percentiles in verbal and math, but importantly, no discontinuous change around the event date. Figure (C13) show the average fraction of students who are White, Black, Hispanic, and Asian respectively. We see slight trends over time. For example, the fraction of Hispanic students is trending upwards. Importantly, we see no discontinuous change around the event date.

As mentioned above, it is important to note that UNLPs are not implemented at random, and other policies that affect labor market outcomes could have been implemented simultaneously. For example, universities could have implemented other non-financial aid policies (e.g. hired more guidance counselors) in the same year they implemented an UNLP. To address this concern, one needs to exercise careful due diligence and, to the extent possible, collect data on these other types of policies. I have begun the process of collecting data both on other types of financial aid policies and on non-financial aid policies. For example, I have collected data on the implementation of work-study programs, but I have not found

any significant differences that correlate with UNLP implementation.

Additionally, I have explored the fact that a few schools rolled back their UNLPs, causing increases in student debt. In line with the exclusion restriction, I find a symmetrical effect of student debt on the outcome of interest regardless of whether the change in student debt is caused by implementation or roll-back of an UNLP.

D.IV Monotonicity

The monotonicity assumption posits that there are no *defiers* of the treatment. That is, no students take on more student debt as a result of the UNLP. For example, the monotonicity assumption would not hold if grants for low-income students were crowded out with the implementation of universal loan elimination programs, which would leave low-income students worse off and presumably with more student debt. Given the results on the extensive margin reported in table 3, this would be a particularly important assumption to test, if I was using loan caps as an instrument for student debt (instead of UNLPs).

I test the monotonicity assumption using the credit bureau data in the following way. First, I tabulate students by family income quartile.[34] For each quartile I calculate (a) the average amount of student debt, (b) the fraction of students that have any student debt, and (c) the fraction of students that have more than $10,000 of student debt. Next, I find that neither (a), (b), or (c) weakly decrease after the policy implementation within each quartile.

Finally, I refer to Krishnan and Wang (2018), who have also studied the monotonicity assumption. They show that there is no crowding out of other financial aid with the introduction of NLPs. They find that when schools implemented NLPs and increased financial aid grants, they did not simultaneously decrease other forms of financial aid.

[34] I have done this both using the imputed value of household income from W2 data provided by Experian, and by using ZIP codes as a proxy for income.

E Comparison with Other Financial Aid Programs

In this Section, I compare UNLPs with other financial aid programs. Specifically, I compare UNLPs with no-loan programs that were targeted to specific income groups (targeted NLPs or TNLPs).

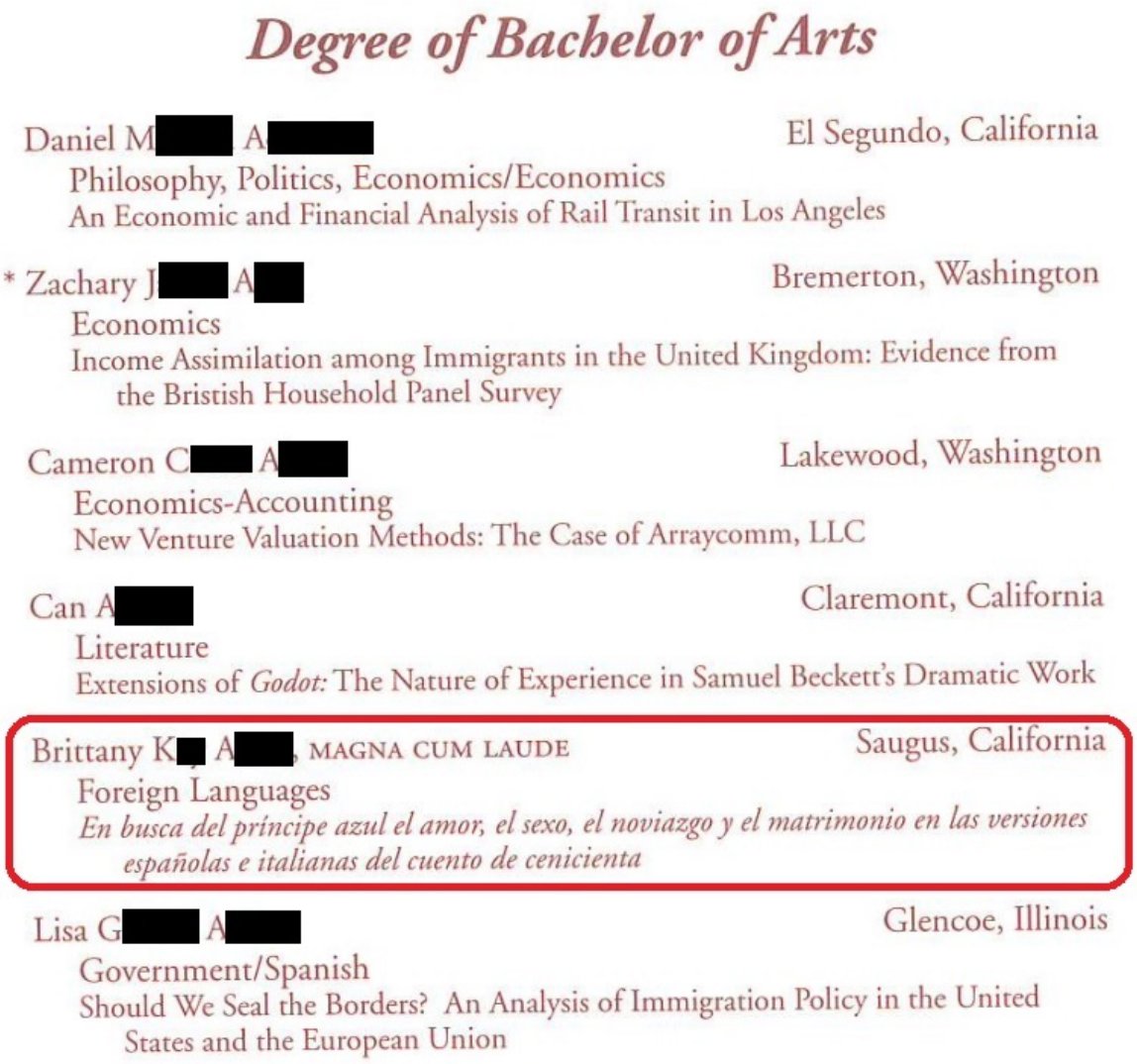
Since 1998, at least 85 universities have implemented new financial aid policies. Figure C2 displays the implementation of financial aid policies over time. Panel A displays the implementation of all five financial aid policies (including NLPs, student loan caps, parental contribution eliminations, tuition waivers, and Pell Grant matches), and Panel B displays the implementation of just NLPs. The black bars denote policies targeted specifically at low-income students, and the grey bars denote policies available for all students. The years refer to starting year of the policy. For example, 20 universities implemented a NLP for low-income students matriculating in 2008 (and graduating in 2012). Notably, three universities reversed their policies: Dartmouth College and Williams College in 2011 and Claremont McKenna College in 2014.

Table D7 and Table D8 report summary statistics. Table D7 reports summary statistics for all institutions that implemented any financial aid policy. The first column report the mean statistic across all institutions, and following columns report the mean statistics for institutions that implemented an NLP, a Loan Cap Policy, No Parental Contribution, Pell Grant Match, and Tuition Waiver respectively. On average, institutions that implemented NLPs are slightly smaller, slightly more selective, and have a higher tuition cost. Table D8 reports the summary summary statistics separately for institutions that implemented an NLP for all students and those that implemented an NLP for low-income students only. Schools that implemented for all students tend to be smaller, have lower admission rates, higher tuition, and have slightly higher SAT scores for their admitted students.

The distribution across universities also vary. Figure C8 display the distribution of out-of-state tuition. Panel A display tuition in 2006 and Panel B tuition in 2015. We see that annual tuition for universities that implemented NLPs for some students (grey bars) vary from less than $20,000 per year to almost $50,000 per year (in 2016). The tuition for the universities that implemented NLPs for all students is narrower (blue bars): between $40,000 and $48,000. Figure C9 display the distribution of endowments in 2006 (Panel A) and in 2015 (Panel B). The universities that have implemented The majority of universities have endowments of between $1 and $10 billion, while a few universities have endowments above $20 billion. Figure C10 display the admission rates, and Figure C11 display the average SAT scores for the 25th and 75th percentiles of admitted students.

F Appendix Figures

Figure C1: Commencement Program Example



Degree of Bachelor of Arts

Daniel M[redacted] A[redacted] — El Segundo, California
Philosophy, Politics, Economics/Economics
An Economic and Financial Analysis of Rail Transit in Los Angeles

* Zachary J[redacted] A[redacted] — Bremerton, Washington
Economics
Income Assimilation among Immigrants in the United Kingdom: Evidence from the Bristish Household Panel Survey

Cameron C[redacted] A[redacted] — Lakewood, Washington
Economics-Accounting
New Venture Valuation Methods: The Case of Arraycomm, LLC

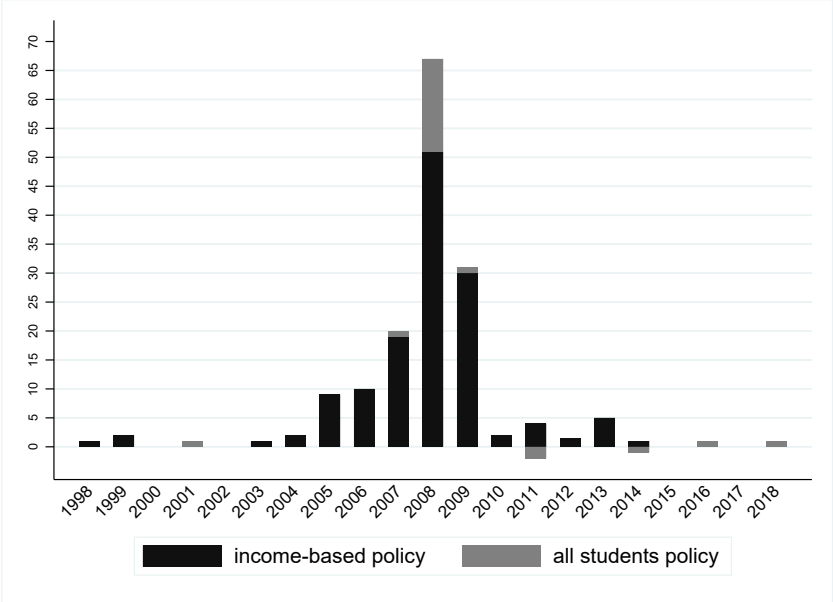
Can A[redacted] — Claremont, California
Literature
Extensions of *Godot:* The Nature of Experience in Samuel Beckett's Dramatic Work

Brittany K[redacted] A[redacted], MAGNA CUM LAUDE — Saugus, California
Foreign Languages
En busca del príncipe azul el amor, el sexo, el noviazgo y el matrimonio en las versiones españolas e italianas del cuento de cenicienta

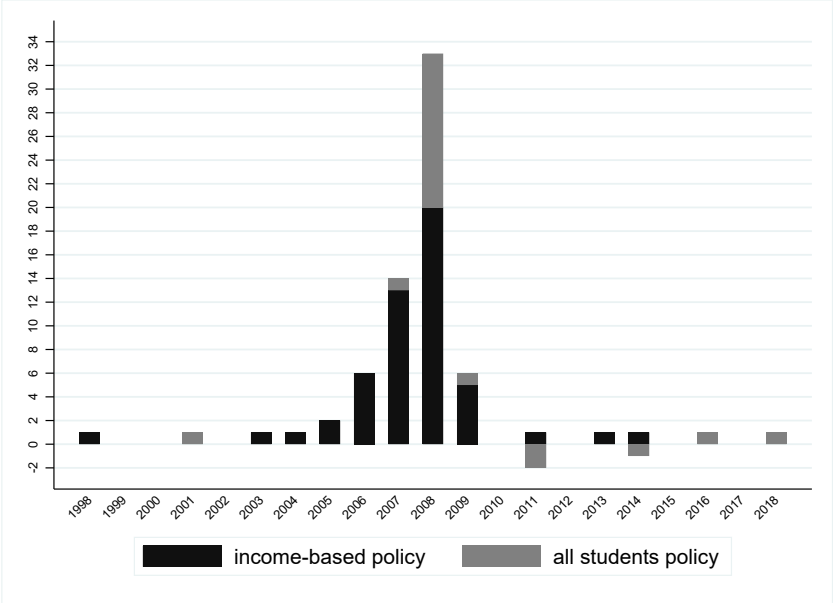
Lisa G[redacted] A[redacted] — Glencoe, Illinois
Government/Spanish
Should We Seal the Borders? An Analysis of Immigration Policy in the United States and the European Union

Note: This picture is an example of commencement program from the 2008 cohort of Claremont Mckenna College. The picture includes full names, major choices, hometowns and states, awards as indicated by *, and thesis titles. This digitization was procured from the the Claremont Schools Special Archives and Dean's Office.

Figure C2: Number of financial aid policies implemented, 1998–2018



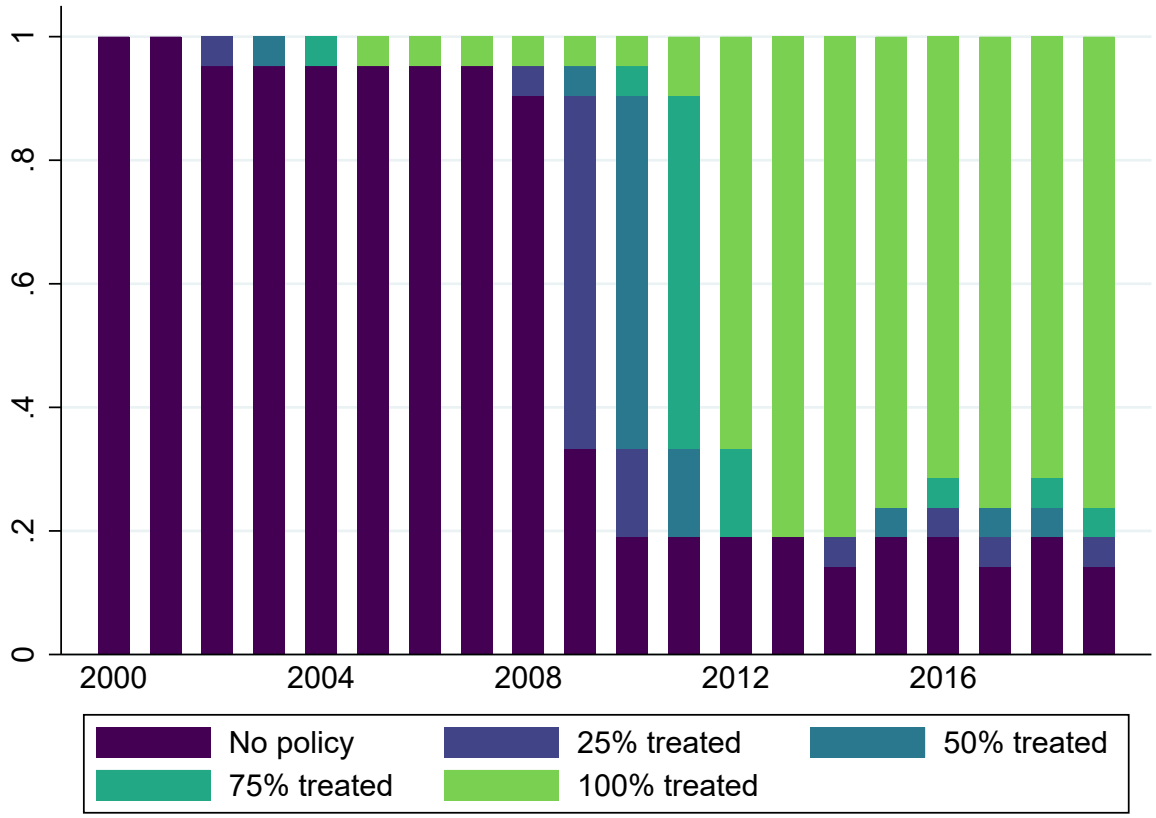
(A) Implementation of *all* financial aid policies



(B) Implementation of loan elimination policies (no-loan policies)

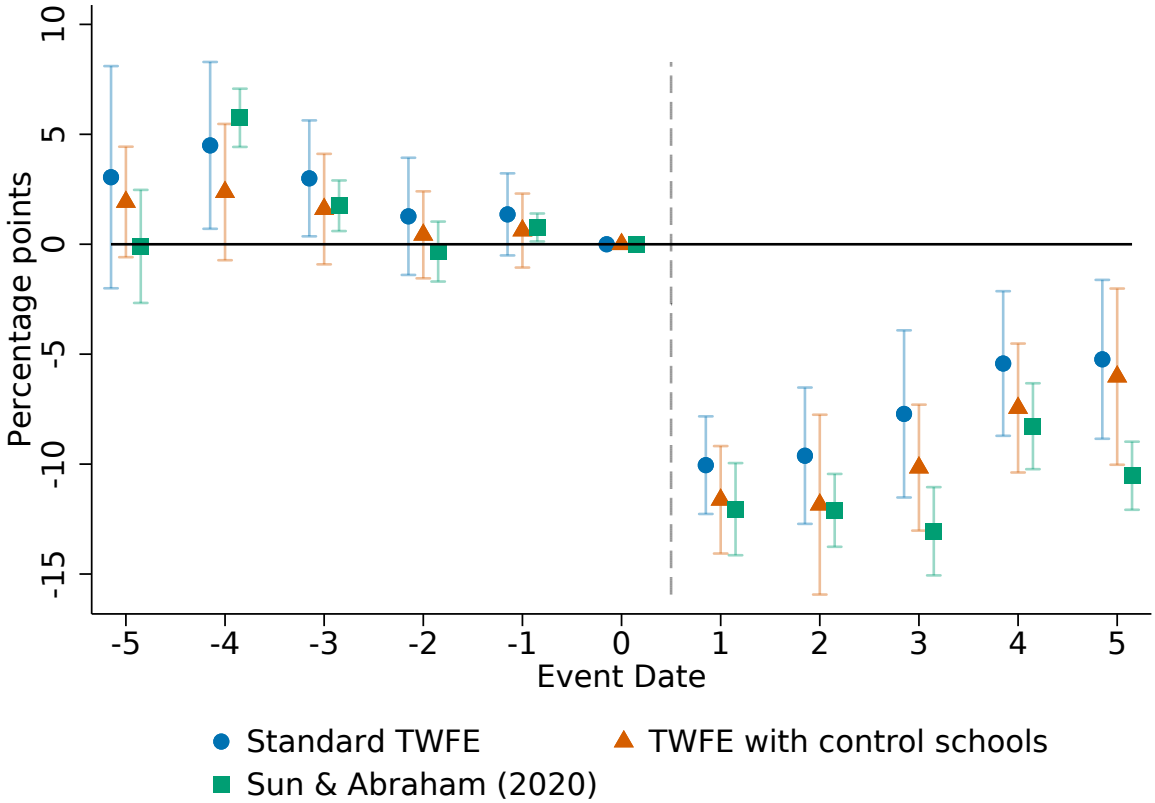
Note: This figure shows the number of financial aid policies that have been implemented for particular household income bands and for all students over time. Panel (a) shows the total number of universities implementing loan eliminations, loan caps, parental contribution waivers, Pell Grant matches, and tuition waivers. Panel (b) shows the total number of no-loan policies implemented over time. The black bars denote policies targeted specifically at low-income students, and the gray bars denote policies available for all students. The data source is the Integrated Postsecondary Education Data System (IPEDS).

Figure C3: UNLP implementation timeline



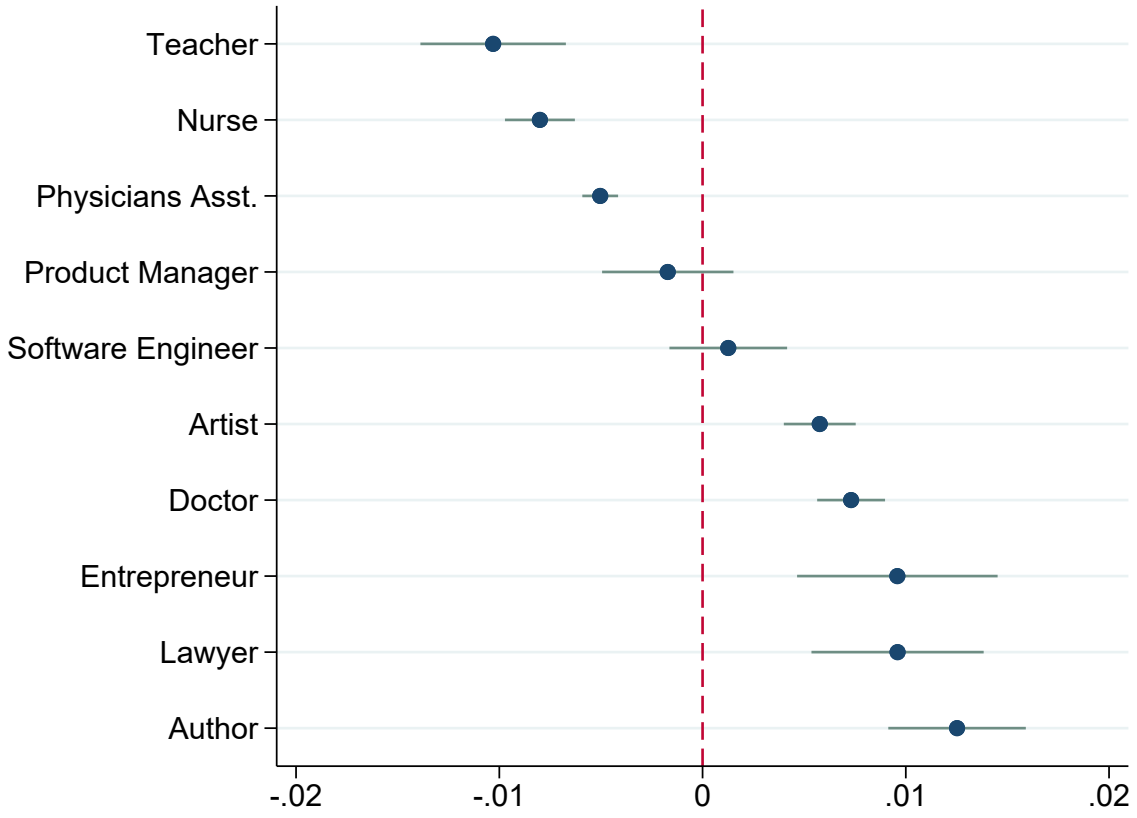
Note: This figure describes the implementation timeline of the universal no-loan policies (UNLPs). The figure shows the fraction of cohort-schools treated and the intensity of the treatment. For example, if a university implemented a UNLP in the fall of 2001, then the cohort of students graduating in the spring of 2002 would have been treated for 25% of the time they were in college, and the cohort graduating in the spring of 2003 would have been treated for 50% of the time they were in college. The data source for both panels is the Integrated Postsecondary Education Data System (IPEDS).

Figure C4: Effect of UNLPs using publicly available data



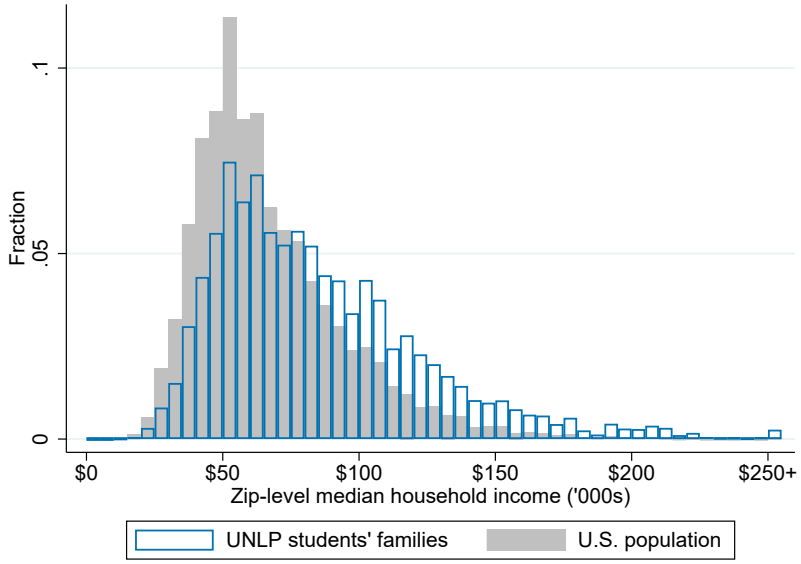
Note: This figure describes the effect of universal no-loan policies (UNLPs) using publicly available data. The figure shows the regression coefficients and 95% confidence intervals from two regressions of the percentage of students taking a student loan on year dummies relative to the implementation of a UNLP. A standard two-way fixed effects (TWFE) model is reported in blue, a TWFE model with non-implementing top-50 schools as a control group is reported in red, and a bias-corrected model allowing for treatment heterogeneity across cohorts (following Sun and Abraham (2020)) is reported in green. The data source for both panels is the Integrated Postsecondary Education Data System (IPEDS).

Figure C5: Occupations after NLP implementation

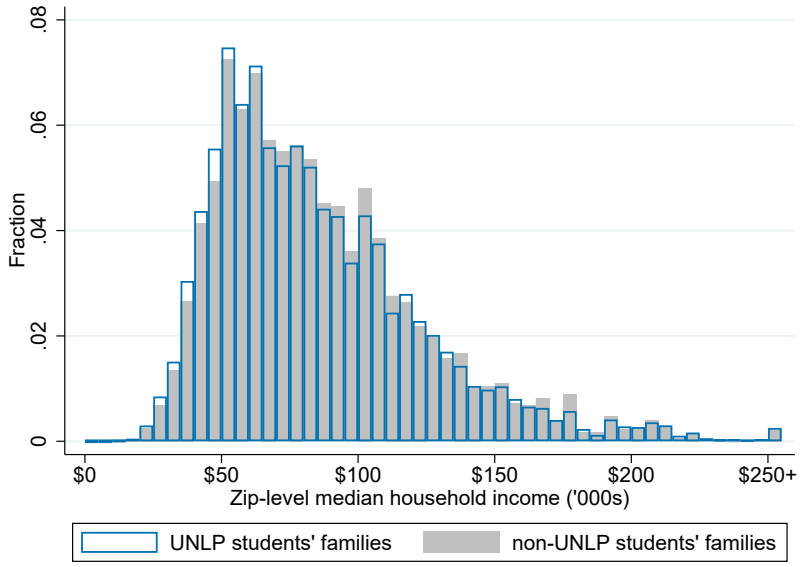


Note: This figure shows a plot of the regression coefficients from Equation (7), with the outcome variable being a dummy variable that takes the value 1 if an individual has a particular occupation seven years after graduation. The coefficient represents the change in the fraction of students with that occupation, controlling for student characteristics and for school and cohort fixed effects. The treated sample includes only students who were enrolled at the time the NLP was implemented.

Figure C6: ZIP-level Household Income



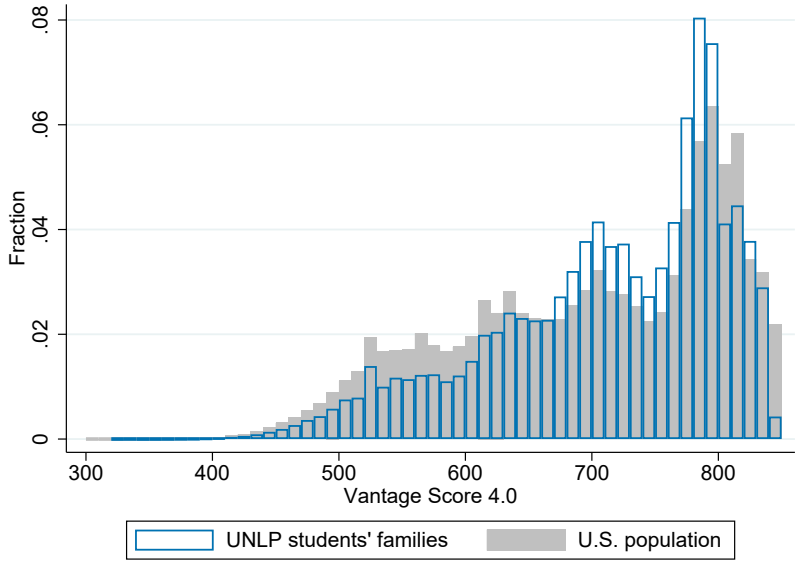
(A) UNLP families vs US population



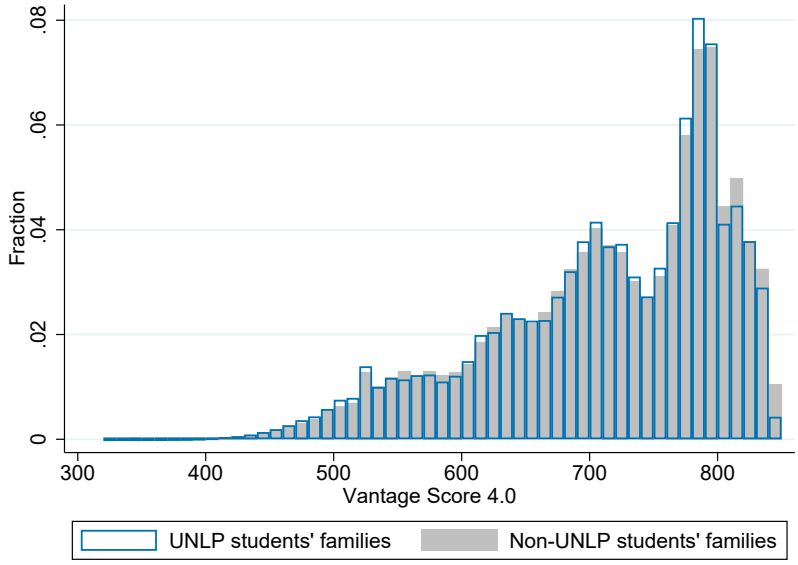
(B) UNLP families vs non-UNLP families

Note: The figure plots histograms of the distribution of median household income by ZIP codes. Panel (A) plots the distribution of ZIP code incomes for the UNLP students' families (in blue) relative to a 1% random sample of the U.S. population (in gray). Panel (B) plots the distribution of ZIP code incomes for the UNLP students' families (in blue) relative to a the families of students from non-UNLP control schools (in gray).

Figure C7: Vantage Score Distributions



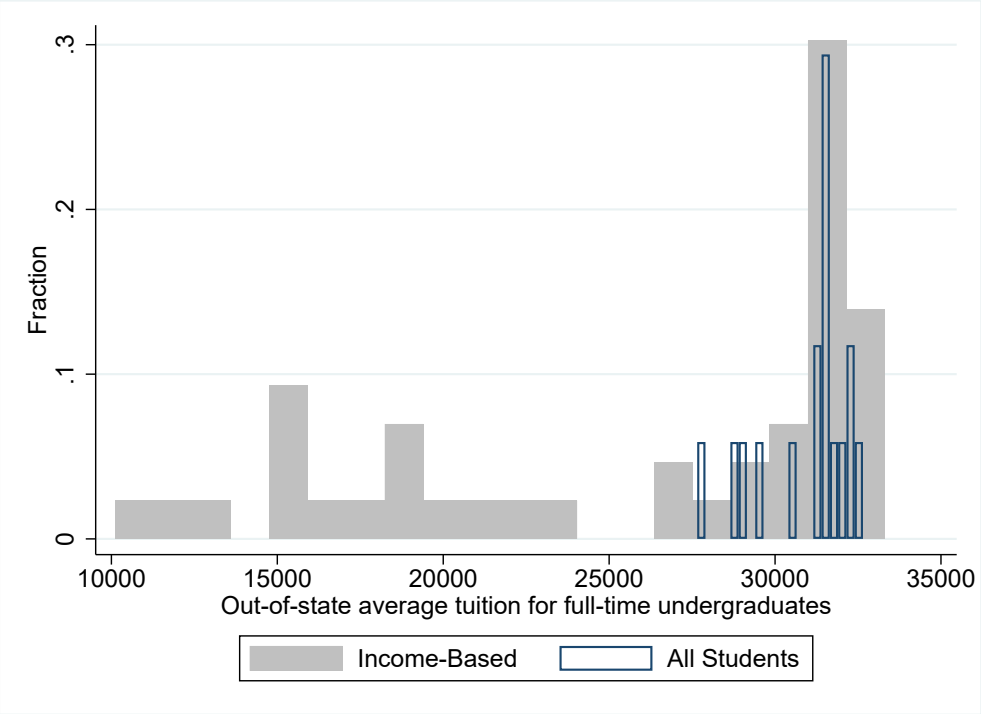
(A) UNLP families vs US population



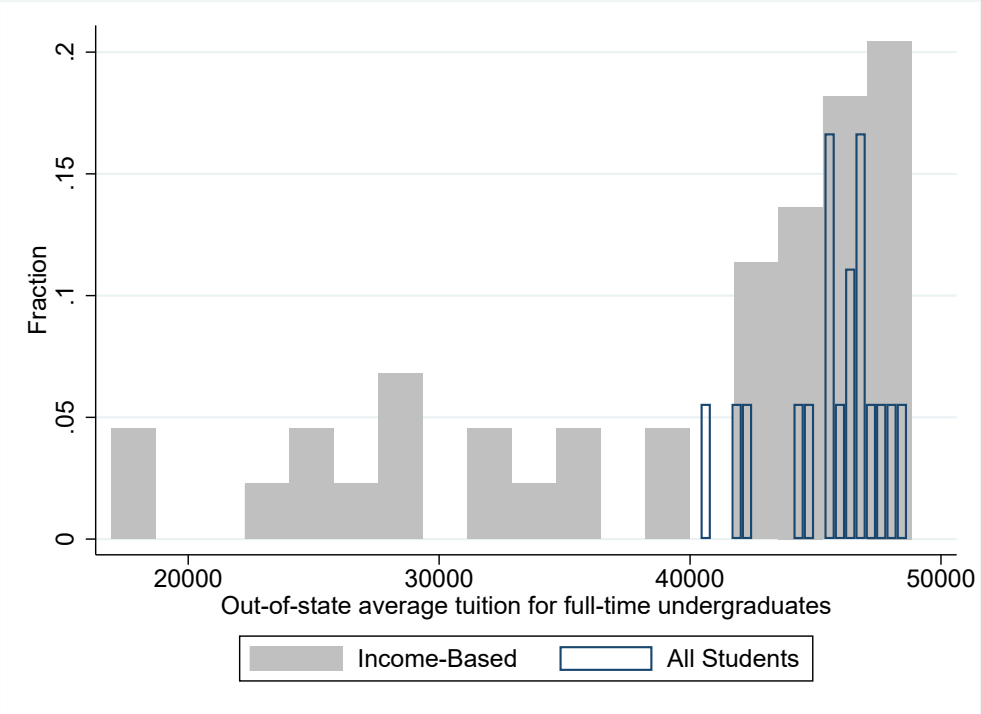
(B) UNLP families vs non-UNLP families

Note: The figure plots histograms of the distribution of Vantage Scores in the sample. Panel (A) plots the distribution of Vantage Scores for the UNLP students' families (in blue) relative to a 1% random sample of the U.S. population (in gray). Panel (B) plots the distribution of Vantage Scores for the UNLP students' families (in blue) relative to a the families of students from non-UNLP control schools (in gray).

Figure C8: Out-of-State Tuition in 2006 and 2015



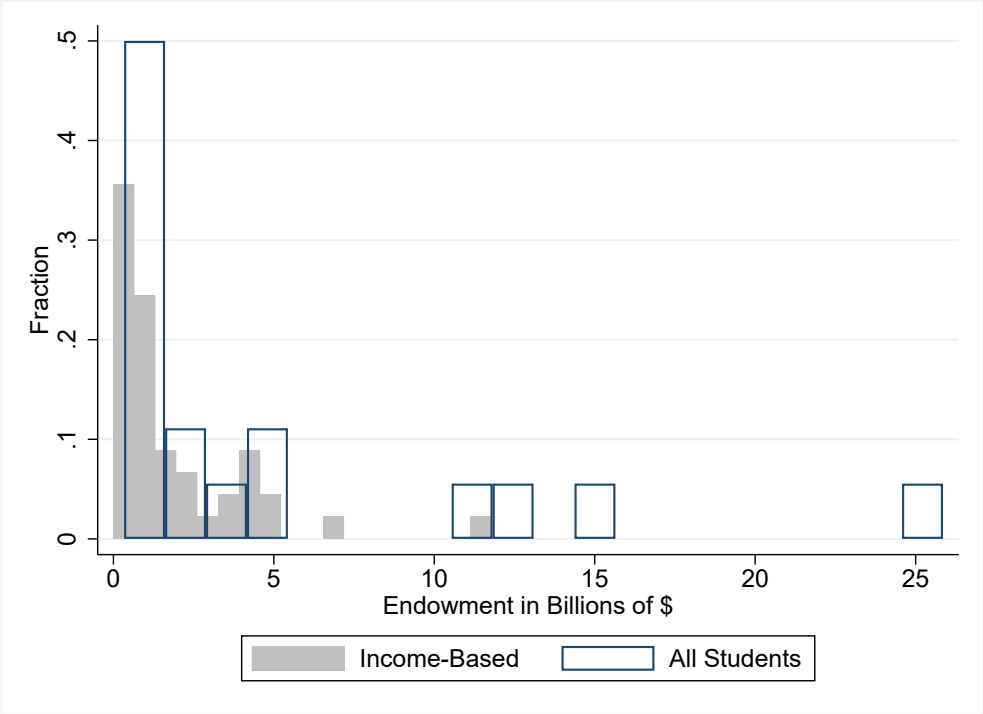
(A) Average Out-of-State Tuition in 2006



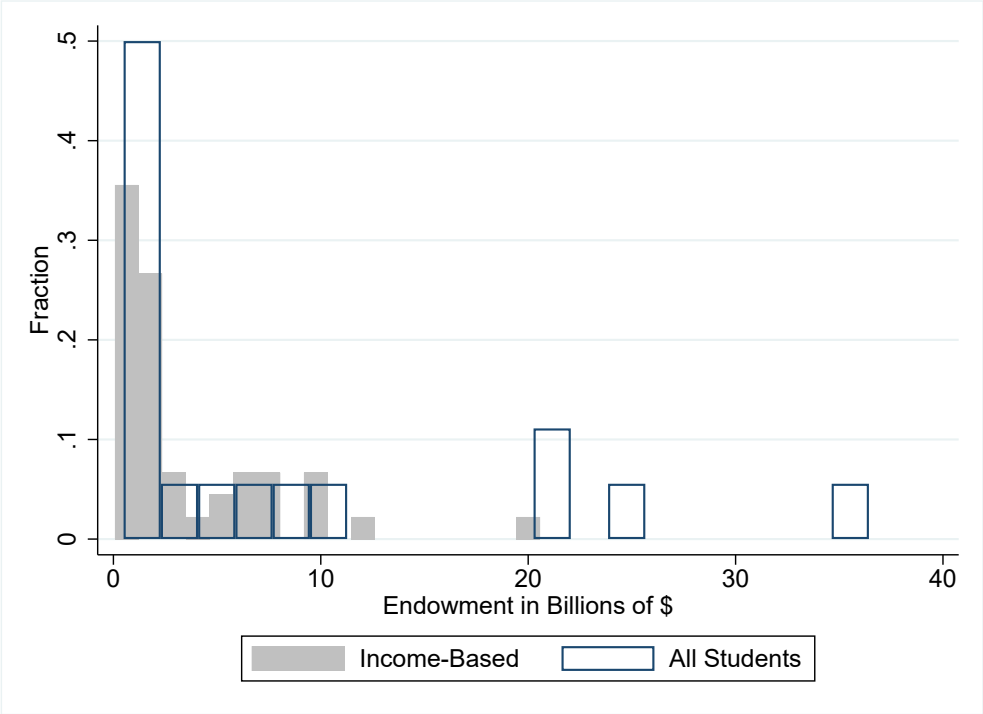
(B) Average Out-of-State Tuition in 2015

Note: This figure plots the average out of state tuition in 2006 and 2015 for schools that implemented income-based financial aid policies and schools that implemented loan elimination programs for all students. The data source is the Integrated Postsecondary Education Data System (IPEDS).

Figure C9: Value of Endowment in Billions in 2006 and 2015



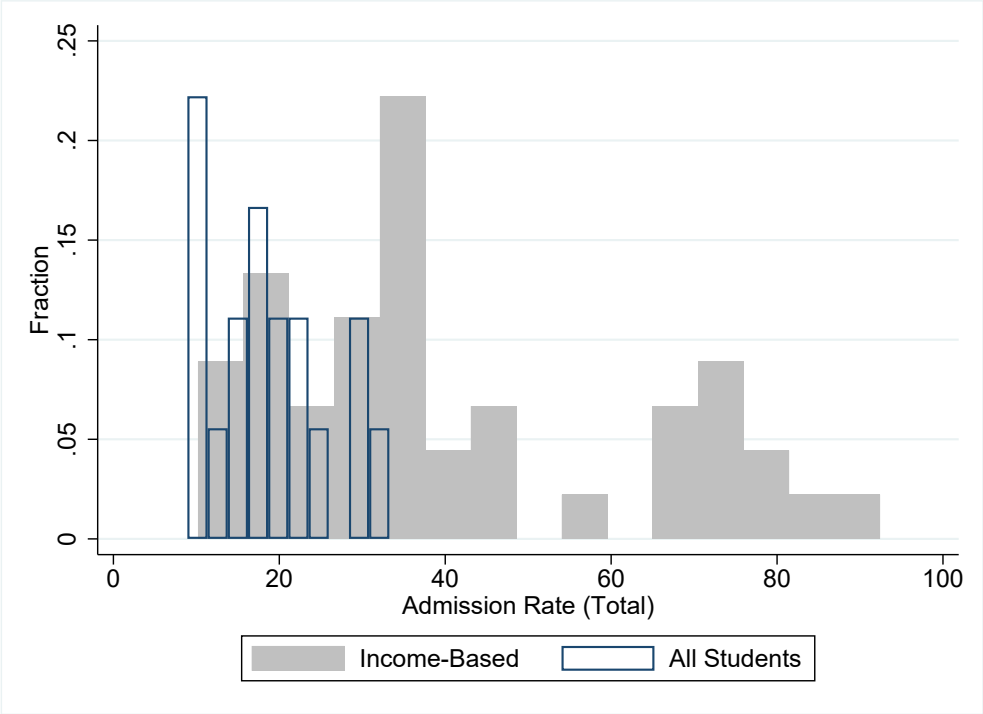
(A) Value of Endowments in 2006



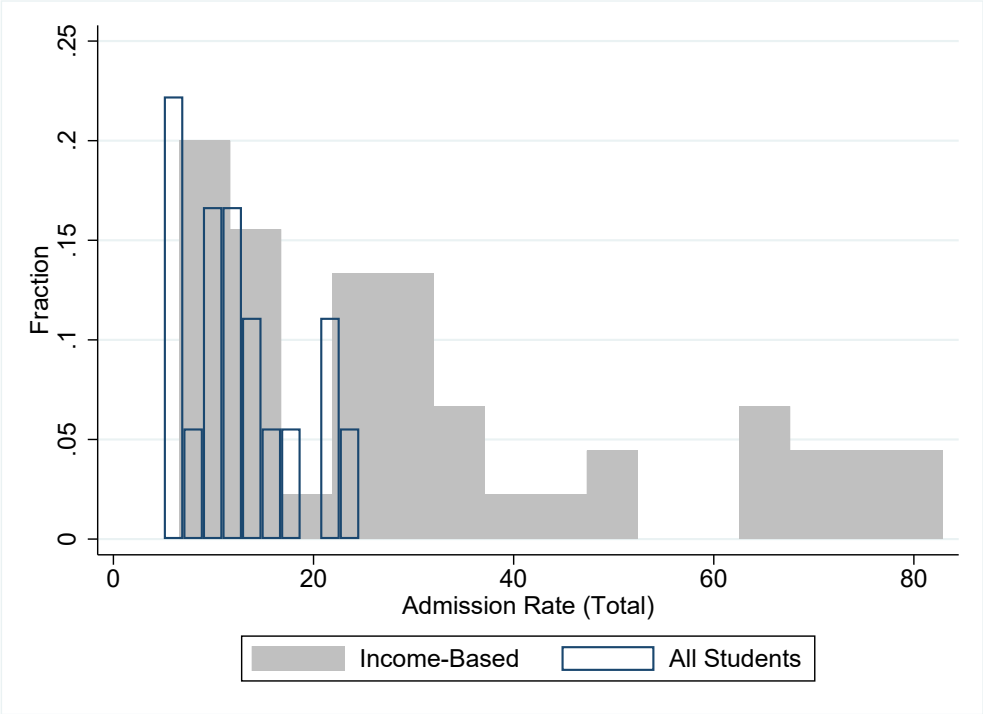
(B) Average Out-of-State Tuition in 2015

Note: This figure plots institution endowments in 2006 and 2015 for schools that implemented income-based financial aid policies and schools that implemented loan elimination programs for all students. The data source is the Integrated Postsecondary Education Data System (IPEDS).

Figure C10: Admission Rate in 2006 and 2015



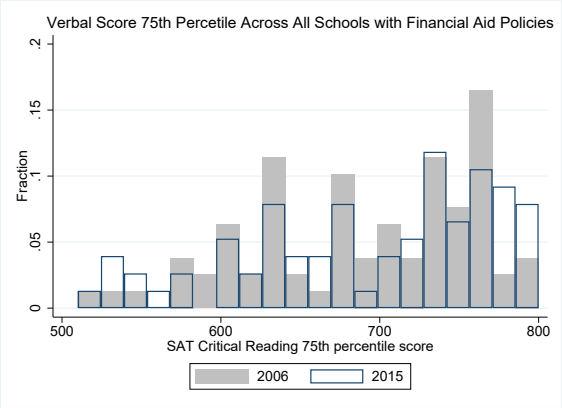
(A) Admission Rate in 2006



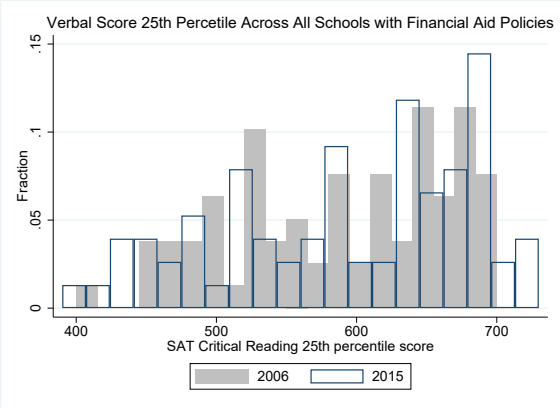
(B) Admission Rate in 2015

Note: This figure plots the average admission rate in 2006 and 2015 for schools that implemented income-based financial aid policies and schools that implemented loan elimination programs for all students. The data source is the Integrated Postsecondary Education Data System (IPEDS).

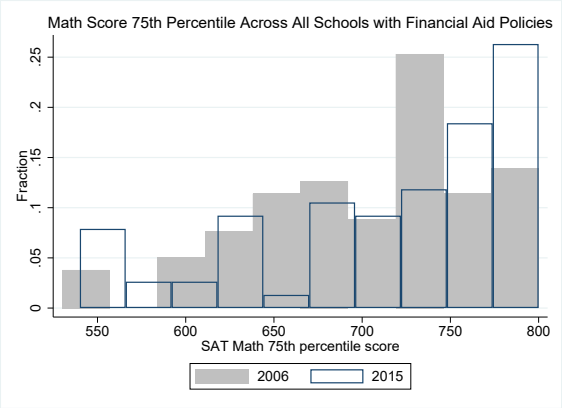
Figure C11: SAT in 2006 and 2015



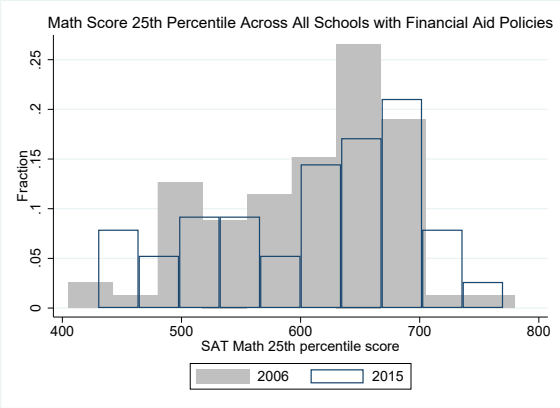
(A) Average of the 75th Percentile in Verbal



(B) Average of the 25th Percentile in Verbal



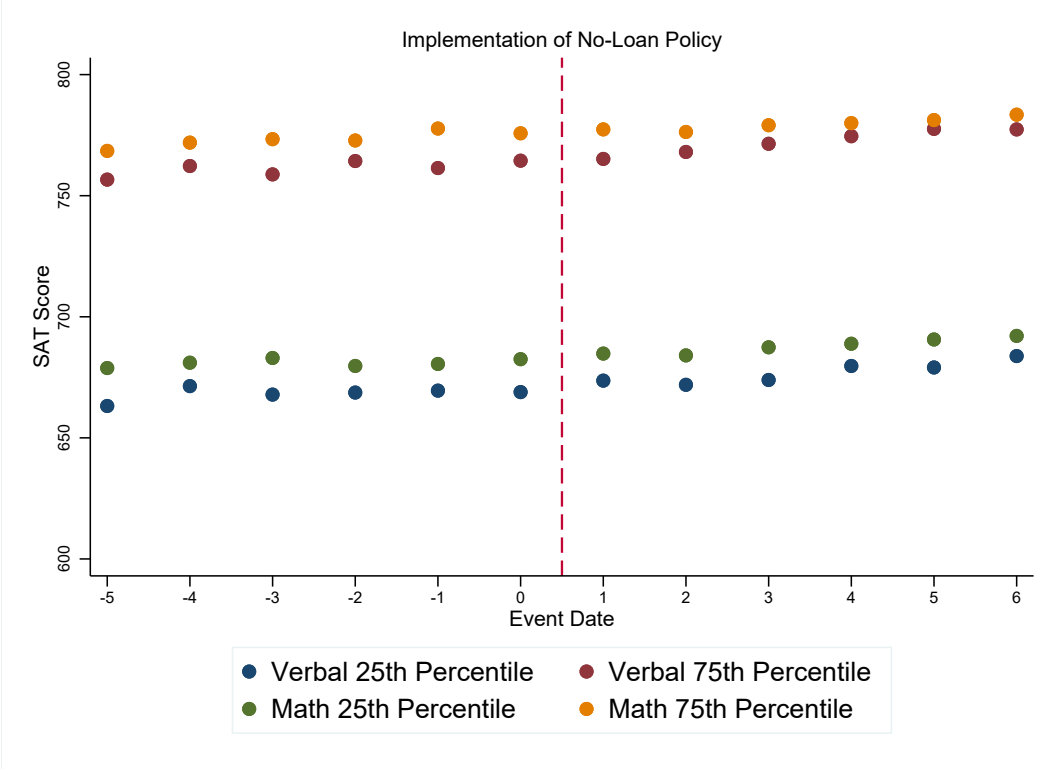
(C) Average of the 75th Percentile in Math



(D) Average of the 25th percentile in Math

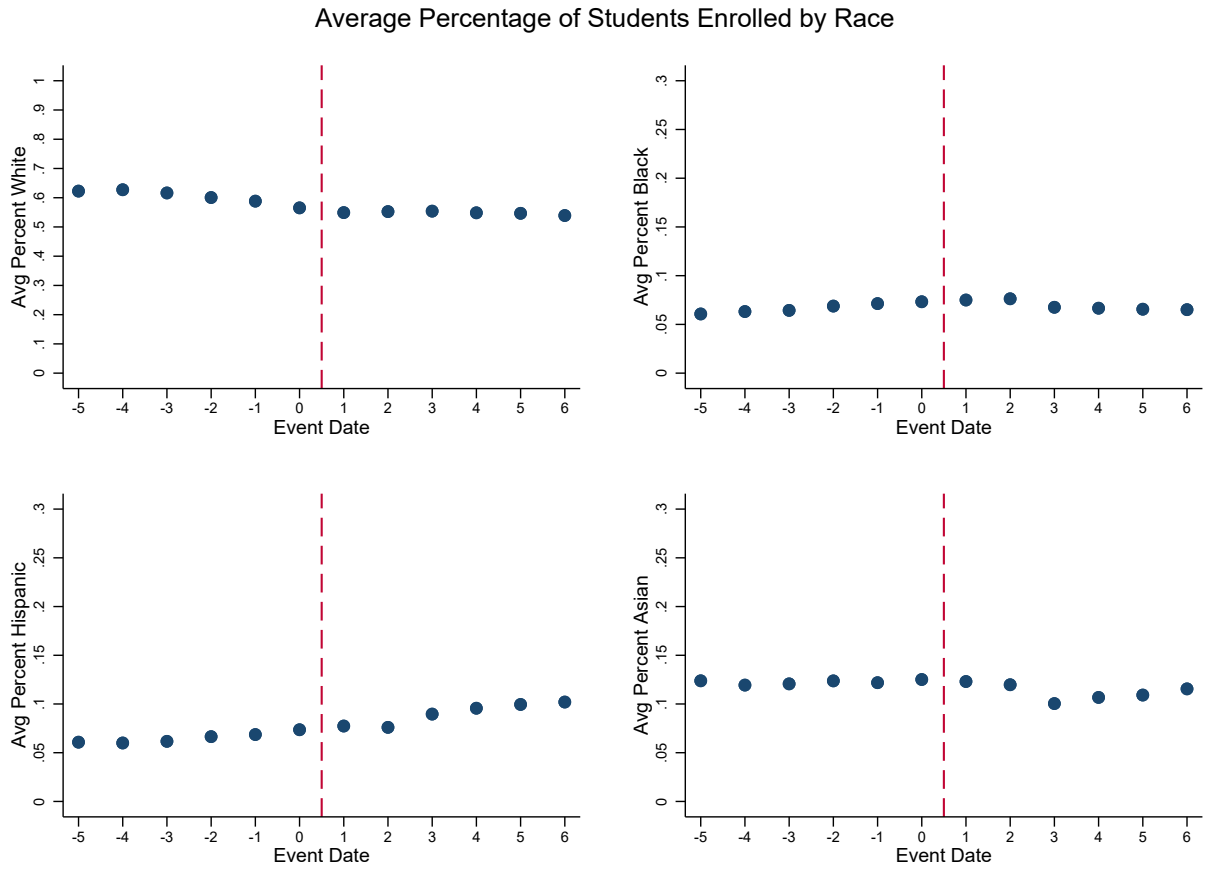
Note: This figure plots the average of the 25th and 75th percentiles in both verbal and math. The data source is the Integrated Postsecondary Education Data System (IPEDS).

Figure C12: Average SAT



Note: This figure plots the averages of the 25th and 75th percentiles of SAT scores relative to year that an institution implemented an NLP for all students. The data source is Integrated Postsecondary Education Data System (IPEDS).

Figure C13: Race event



Note: This figure plots the average of the percent of White, Black, Asian, and Hispanic students across institutions that implemented an NLP for all students relative to the event year. The data source is the Integrated Postsecondary Education Data System (IPEDS).

G Appendix Tables

Table D1: Mean and standard deviation of residualized wages

	Mean	StdDev	Coef_var	Slope_lvl	Slope_log
Biology and Life Sciences	19,728	97,176	4.93	7,899	26.38
Engineering	17,298	83,694	4.84	6,773	20.13
Physical Sciences	11,886	89,019	7.49	6,331	20.32
Mathematics and Statistics	10,414	85,148	8.18	6,158	16.98
Social Sciences	8,474	90,187	10.64	5,940	12.90
Computer and Information Sciences	7,443	71,722	9.64	5,615	18.99
Medical and Health Sciences and Services	6,523	57,669	8.84	3,668	12.69
Business	4,390	79,458	18.10	4,797	9.70
Transportation Sciences and Technologies	2,630	75,224	28.60	5,467	16.07
Area, Ethnic, and Civilization Studies	2,102	75,643	35.99	4,514	9.41
History	-1,294	84,321	-65.15	5,539	13.47
Interdisc and Multi-Disc Studies	-2,118	64,046	-30.24	4,282	10.58
Law	-4,720	64,945	-13.76	3,847	15.62
Communications	-4,934	63,667	-12.90	3,608	3.44
Engineering Technologies	-4,988	64,569	-12.95	4,743	19.70
Psychology	-5,787	60,417	-10.44	3,443	7.98
Linguistics and Foreign Languages	-6,127	66,547	-10.86	3,725	7.12
English Lang, Lit, and Comp	-6,255	69,124	-11.05	3,810	7.22
Architecture	-8,198	65,959	-8.05	3,664	9.38
Public Affairs, Policy, and Social Work	-10,402	49,919	-4.80	2,477	6.92
Philosophy and Religious Studies	-10,464	80,478	-7.69	4,991	11.91
Liberal Arts and Humanities	-11,420	62,077	-5.44	3,489	9.30
Criminal Justice and Fire Protection	-11,986	48,567	-4.05	3,805	14.63
Physical Fitness and Leisure	-12,074	50,626	-4.19	4,260	14.98
Environment and Natural Resources	-12,662	56,124	-4.43	3,699	9.59
Family and Consumer Sciences	-12,820	44,407	-3.46	2,104	0.70
Agriculture	-14,551	61,743	-4.24	3,512	8.29
Education Administration and Teaching	-14,834	40,040	-2.70	2,225	9.01
Fine Arts	-18,325	53,725	-2.93	2,704	3.20
Theology and Religious Vocations	-34,594	43,975	-1.27	2,341	6.50

Note: This table reports the average residual wage (in Column 1) and the standard deviation in residuals (in Column 2) from a Mincer regression of annual wage on age, age-squared, race, ethnicity, sex, and survey-year fixed effects. Column 3 reports the predicted slope on age from a Mincer regression of annual wage where each major interacts with age and age-squared; the slope is evaluated at age 22. The data source is the ACS.

Table D2: Summary statistics—demographics and degrees conferred

	Mean	SD	p25	p50	p90	N
Field of Studies						
Agriculture	52752.7	62031.3	12000.0	40000.0	110000.0	54203
Environment and Natural Resources	52310.4	56901.0	17000.0	43000.0	100000.0	31210
Architecture	58543.2	66841.5	11500.0	48000.0	120000.0	28823
Area, Ethnic, and Civilization Studies	55798.2	79317.4	7200.0	35450.0	125000.0	13566
Communications	52978.7	65712.2	11200.0	40000.0	110000.0	200275
Communication Technologies	45639.5	49623.3	12000.0	36000.0	100000.0	7543
Computer and Information Sciences	74841.0	72376.7	28800.0	65000.0	144000.0	146201
Education Administration and Teaching	36929.8	39811.5	9000.0	35000.0	72000.0	340408
Engineering	86362.3	83226.7	34500.0	75000.0	160000.0	288313
Engineering Technologies	69351.1	63715.4	30000.0	61000.0	130000.0	33260
Linguistics and Foreign Languages	49058.3	71409.6	3700.0	32000.0	107000.0	37430
Law	52340.8	66750.8	10000.0	40000.0	110000.0	8507
English Lang, Lit, and Comp	52766.1	75012.9	6000.0	35000.0	116000.0	117561
Liberal Arts and Humanities	47770.3	64319.6	6000.0	34300.0	100000.0	59898
Biology and Life Sciences	84785.3	110323.4	14000.0	50000.0	201000.0	224085
Mathematics and Statistics	75138.41	90229.41	17200.00	54000.00	154000.00	47773
Interdisc and Multi-Disc Studies	51037.43	69180.54	9000.00	36000.00	109000.00	36964
Physical Fitness and Leisure	44069.73	53027.25	11000.00	35000.00	90000.00	48301
Philosophy and Religious Studies	60164.48	87041.38	7700.00	37000.00	135000.00	26146
Theology and Religious Vocations	36285.29	43973.20	7000.00	29400.00	75000.00	23393
Physical Sciences	81801.82	97693.12	20000.00	55000.00	180000.00	114863
Psychology	48013.33	66043.93	6700.00	32650.00	101000.00	188336
Criminal Justice and Fire Protection	50798.33	49823.82	20000.00	42000.00	100000.00	87316
Public Affairs, Policy, and Social Work	39903.65	51863.90	9000.00	32000.00	80000.00	46623
Social Sciences	70153.45	93916.89	12000.00	45000.00	150000.00	302590
58 incwage	39727.69	35453.24	3000.00	35000.00	90000.00	65
Fine Arts	39411.20	53559.16	3100.00	28000.00	90000.00	191799
Medical and Health Sciences and Services	58920.95	60680.20	20500.00	50000.00	110000.00	312050
Business	66628.25	79578.79	20000.00	50000.00	135000.00	927466
History	66734.67	91210.77	12000.00	42000.00	150000.00	78114
Demographic Characteristics						
Age	40.54	10.91	31.00	40.00	56.00	4111126

Note: The above table reports mean values of annual wages, standard deviations, percentiles, and count values across individuals who are of working age (between 21 and 60) and have at least a bachelor's degree. Column 1 gives mean wages in dollars, Column 2 gives the standard deviation in dollars, and Columns 3–5 give percentiles of wages in dollars. The data source is the American Community Survey (ACS).

Table D3: Regression coefficients for wage imputation

	(1) Wage	(2) Log(wage)
Age	8656.8	0.28
	(94.53)	(106.51)
Age^2	-91.4	-0.0035
	(-125.29)	(-124.15)
Non-Hispanic White	13148.2	0.46
	(53.12)	(42.54)
African American	-755.2	0.46
	(-1.26)	(26.60)
Hispanic	788.1	0.052
	(3.11)	(4.96)
Asian	12662.4	-0.19
	(15.79)	(-9.87)
Female	-32790.5	-1.17
	(-52.62)	(-94.26)
Survey Year FE	✓	✓
N	5,628,805	5,628,805
R^2	0.11	0.04

Note: This table reports regression coefficients from the Mincer regression of annual wage and log(wage) on age, age squared, race, ethnicity, gender, and survey-year fixed effects. The data source is the American Community Survey (ACS).

Table D4: The intertemporal trade-off across college majors

	Life-time wage	Year bins			Life-time wage	Year bin
	(1)	(2)	(3)	(4)	(5)	(6)
		30s	40s	50s		30-34
Initial wage (ACS)	-1.04	-0.74	-1.72	-1.61	-0.81	
	(-2.22)	(-1.76)	(-2.69)	(-2.57)	(-3.33)	
Initial wage (Experian)						-0.76
						(-2.15)
Sample	ACS	ACS	ACS	ACS	ACS	Experian
CIP Classification	2 digits	2 digits	2 digits	2 digits	4 digits	2 digits
Residualization fixed effects	Year	Year	Year	Year	Year	Year & School
N	36	36	36	36	121	24
R^2	0.13	0.08	0.18	0.16	0.09	0.17

Note: This table reports regression coefficients from regressing the average residualized wage characteristics across college majors. Columns (1) and (5) regress the average annual life-time wage regressed on the initial wage. Columns (2)-(4) and Column (6) regress the average annual wage across a specific decade on the initial wage. Columns (1)-(4) and (6) use the 2-digit CIP classification of majors, while Column (5) use the 4-digit classification. Columns (1)-(5) use initial wage calculated from the ACS, while Column (6) use initial wage caclculated from the Experian data.

Table D5: Major Choice and Parents' Financial Characteristics

	Initial wage				Mean life-time wage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vantage Score	-3.61	-2.32			8.57	6.27		
	(-5.07)	(-7.66)			(4.40)	(7.49)		
Log(income)			-434.7	-225.4			1348.5	725.6
			(-2.14)	(-2.67)			(2.29)	(5.49)
School FE		✓		✓		✓		✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Demographics	✓	✓	✓	✓	✓	✓	✓	✓
N	227,429	227,429	227,429	227,429	227,429	227,429	227,429	227,429

Note: This table reports regression coefficients from difference-in-differences regressions regressing implied wages from the choice of college major on the parents' financial characteristics. Columns 1 through 4 report the coefficient where the dependent variable is the initial wage for the major, and Columns 5 through 8 report the coefficients where the dependent variable is the mean life-time wage for the major.

Table D6: UNLP and the difficulty of college majors

	Hours	Average GPA	Difficulty
	(1)	(2)	(3)
UNLP $\times Post$	0.10	-0.012	0.025
	(2.07)	(-3.59)	(1.93)
Control schools	✓	✓	✓
School FE	✓	✓	✓
Year FE	✓	✓	✓
Demographics	✓	✓	✓
Data Source	NSSE (2016)	Novik (2022)	Rask (2010)
N	697,441	555,190	714,521
R^2	0.11	0.11	0.07

Note: This table reports regression coefficients from difference-in-differences regressions regressing characteristics of college major choice on UNLP status. Column 1 reports the effect on average hours worked, Column 2 on average GPA, and Column 3 on self-reported difficulty.

Table D7: Summary Statistics For All Schools With Any Financial Aid Policies

	All	No Loan	Loan Cap	Parent Contrib.	Pell Grant	Tuition
Total Number of Undergraduates	10745.86	9708.73	12732.83	8911.40	10280.40	13978.57
	(10327.99)	(10778.45)	(8890.03)	(5767.39)	(11859.30)	(11307.02)
Admission Rate	34.41	31.10	37.36	12.65	58.16	52.66
	(24.46)	(21.56)	(19.88)	(6.10)	(25.59)	(30.89)
Out of State Tuition for Undergraduates	28218.79	30164.19	27267.33	32647.75	16357.10	22306.47
	(10692.34)	(9738.34)	(10299.76)	(8064.78)	(10287.07)	(11554.65)
Percentage of Undergraduates Awarded Loans	35.50	32.50	39.95	27.06	55.48	40.05
	(14.33)	(12.09)	(8.40)	(12.20)	(15.70)	(19.47)
Percentage of Undergradutes Awarded Institutional Grants	49.25	48.01	52.60	48.45	64.02	44.62
	(15.08)	(12.48)	(12.64)	(6.91)	(15.57)	(24.60)
Percentage of White Undergraduates	0.55	0.58	0.47	0.43	0.71	0.53
	(0.18)	(0.14)	(0.18)	(0.07)	(0.18)	(0.24)
Percentage of Black Undergraduates	0.06	0.07	0.05	0.07	0.04	0.07
	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)	(0.07)
Percentage of Hispanic Undergraduates	0.11	0.08	0.12	0.10	0.04	0.20
	(0.12)	(0.04)	(0.09)	(0.03)	(0.04)	(0.26)
Percentage of Asian Undergraduates	0.14	0.12	0.21	0.19	0.08	0.08
	(0.10)	(0.07)	(0.13)	(0.07)	(0.08)	(0.09)
SAT Verbal 25th Percentile	610.79	628.38	585.04	666.88	534.54	547.68
	(78.57)	(64.60)	(68.91)	(35.75)	(88.89)	(103.23)
SAT Verbal 75th Percentile	710.13	724.51	688.70	764.29	653.59	650.15
	(70.06)	(56.01)	(60.80)	(27.09)	(82.95)	(98.15)
SAT Math 25th Percentile	630.93	643.78	613.80	687.38	565.51	568.36
	(77.34)	(62.65)	(63.72)	(29.65)	(106.99)	(108.58)
SAT Math 75th Percentile	727.82	737.14	719.65	781.67	674.69	667.16
	(66.85)	(53.31)	(54.08)	(16.03)	(86.60)	(99.35)
Observations	2226	1197	412	190	104	304

Note: The above table report mean values across universities that implement no loan programs, loan caps, parental contribution programs, pell grant match programs, and tuition waivers between 2001-2015. The data source is the Integrated Postsecondary Education Data System (IPEDS).

Table D8: Summary Statistics For No Loan Policy Schools

	All Students	Low Income	Difference
Total Number of Undergraduates	4480.16	11847.69	-7367.54**
	(3367.39)	(11973.23)	(0.00)
Admission Rate	16.25	37.04	-20.79**
	(7.60)	(22.46)	(0.00)
Out of State Tuition for Undergraduates	33904.68	28665.84	5238.84**
	(7176.66)	(10219.15)	(0.00)
Percentage of Undergraduates Awarded Loans	26.21	35.01	-8.81**
	(11.06)	(11.56)	(0.00)
Percentage of Undergraduates Awarded Institutional Grants	48.25	47.91	0.34
	(6.81)	(14.12)	(0.68)
Percentage of White Undergraduates	0.53	0.61	-0.08**
	(0.11)	(0.15)	(0.00)
Percentage of Black Undergraduates	0.07	0.07	0.00*
	(0.02)	(0.03)	(0.07)
Percentage of Hispanic Undergraduates	0.09	0.08	0.02**
	(0.03)	(0.04)	(0.00)
Percentage of Asian Undergraduates	0.13	0.11	0.02**
	(0.05)	(0.08)	(0.00)
SAT Verbal 25th Percentile	666.48	612.78	53.70**
	(24.25)	(69.32)	(0.00)
SAT Verbal 75th Percentile	760.41	709.81	50.60**
	(23.05)	(58.83)	(0.00)
SAT Math 25th Percentile	673.87	631.46	42.41**
	(24.02)	(69.10)	(0.00)
SAT Math 75th Percentile	765.30	725.61	39.68**
	(25.39)	(57.32)	(0.00)
Observations	342	855	1197

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The above table report mean values across universities that implemented a NLP between 2001-2015. Column 1 are mean values across universities that implemented an NLP for all students and Column 2 are mean values across universities that implemented an NLP for students below a certain household income level. Column 3 reports the difference in means. The data source is the Integrated Postsecondary Education Data System (IPEDS).