Essays on Student Debt

Jung Hwan Kim

A Thesis

in

The Department

of

Economics

Presented in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy (Economics) at

Concordia University

Montréal, Québec, Canada

October 2023

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Jung Hwan Kim**

Entitled: **Essays on Student Debt**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Economics)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

______________________ Chair
Dr. Jean-Philippe Warren

______________________ External Examiner
Dr. Ying Feng

______________________ Examiner
Dr. Jorgen Hansen

______________________ Examiner
Dr. Tatyana Koreshkova

______________________ Examiner
Dr. Panos Margaris

______________________ Supervisor
Dr. Heejeong Kim

Approved by ______________________
Jorgen Hansen, Chair
Department of Economics

__________ 2023 ______________________
Pascale Sicotte, Dean
Faculty of Arts and Science

Abstract

Essays on Student Debt

Jung Hwan Kim, Ph.D.

Concordia University, 2023

The thesis studies student debt in the United States. The first chapter examines the quantitative effects of rising college costs, wage inequality, and delinquencies on growing student debt balances in the U.S. We build an incomplete markets overlapping-generation (OLG) model with choices for a college education, student loans, and delinquency. We solve transitional dynamics with the estimated time-varying changes in college costs and wage inequality, in addition to a stronger preference for college education, that affect the repayment decision of borrowers. We find that these sources increase aggregate student debt balances by $480 billion between 1979 and 2015. Rising college costs increases borrowing by recent college students. The declining average ability of college students and increasing volatility of wage shocks lead to a higher delinquency rate among borrowers over time. Importantly, we find that when borrowers are not delinquent on their payments, the aggregate student debt only increases by 50% of the increase in the benchmark economy, despite all the time-varying sources. This suggests that although the rising college costs largely affect the borrowing behavior of college students, the increasing delinquency rate over time significantly contributes to the rapid growth of U.S. student debt. The second chapter empirically studies how holding student debt affects initial labor market outcomes for college graduates in the United States. In particular, I examine the implications of student debt on college graduates' initial unemployment duration, underemployment rate, and initial labor income. By conducting hazards models analyses using the NLSY97 and O*NET dataset, I first find that college students graduating with student debt, on average, have a shorter unemployment duration before their first job compared to those graduating without student debt. Furthermore, using the average, excluding a self-amount, institution-level grant amount as an instrument, I find that graduating with a higher level of student debt increases the probability of underemployment for college graduates and decreases the initial labor income at their first jobs. The results highlight the impact of student

debt on the trade-off between search time and the job quality.

Contribution of Authors

This thesis contains co-authored material. Chapter 1 is co-authored with Heejeong Kim. All authors have equally contributed and are equally responsible for the work.

Acknowledgments

The six years of Ph.D. journey at Concordia University have been challenging, and I truly acknowledge that I would not have achieved so far without the support and encouragement from numerous people. First and foremost, I thank God for everything that He has allowed me to go through the Ph.D. journey for the last six years.

Words cannot express my sincere gratitude to my supervisor Heejeong Kim for her invaluable guidance and patience. She has been the source of intellectual and emotional support that encouraged me to overcome challenges whenever I felt lost and discouraged. I thank for her unwavering guidance and countless hours devoted to expanding my research ability and providing valuable feedback.

I also thank the examiners for my dissertation–Jorgen Hansen, Tatyana Koreshkova, Panos Margaris, and Ying Feng–for their time and feedback. I am also grateful to Christian Sigouin for his support throughout the Ph.D. program, and I would like to acknowledge the support and resources from the Department of Economics at Concordia University.

My sincere thanks also go to my fellow graduate students, especially my friend Saba Ranjbar, for feedback and moral support. I am also thankful to my friends and church members for their support.

Lastly, I would like to express my deepest appreciation and love to my family who has been endlessly supportive and praying for me.

Contents

List of Figures

List of Tables

Chapter 1

Sources of Rising Student Debt in the U.S.: College Costs, Wage Inequality, and Delinquency

1.1 Introduction

The total outstanding student debt has risen sharply in the U.S. from $50 billion in 1985 to almost $1 trillion in 2015. This 20-fold increase in just 30 years increased the outstanding student debt to approximately 7% of the GDP in 2015.[1] Now, in the U.S., the total outstanding student debt balance exceeds credit card debt and auto loans, making it the second-largest form of household debt after mortgages (Brown, Haughwout, Lee, Scally, and Klaauw (2014)).[2]

Despite growing concern about the rising student debt in the U.S., there is little understanding of the causes of the rapid increase. To fill this gap in the literature, this paper quantitatively studies the contribution of rising college costs, wage inequality, and delinquencies to the rapid growth of U.S. student debt.[3] We

[1] As shown in Figure D1 in Appendix A.4, the rise in student debt is accompanied by increases in the total number of borrowers and the average amount of student debt per borrower. For example, the total number of borrowers has doubled from 22 million in 2004 to 44 million in 2015, and the average student debt per borrower has increased from $15,106 to $21,677 over the same period.

[2] Student debt increased even during the Great Recession, while other consumer debt, such as mortgages, credit cards, auto loans, and home equity lines of credit, decreased.

[3] Notably, rising wage inequality also affects rising net tuition. Cai and Heathcote (2022) show that rising income inequality is the key driver of the increase in U.S. college net tuition between 1990 and 2016. This is because rich households become more willing to pay higher tuition for better education, experience, and networks, providing top-quality universities the incentive to raise their tuition and fees.

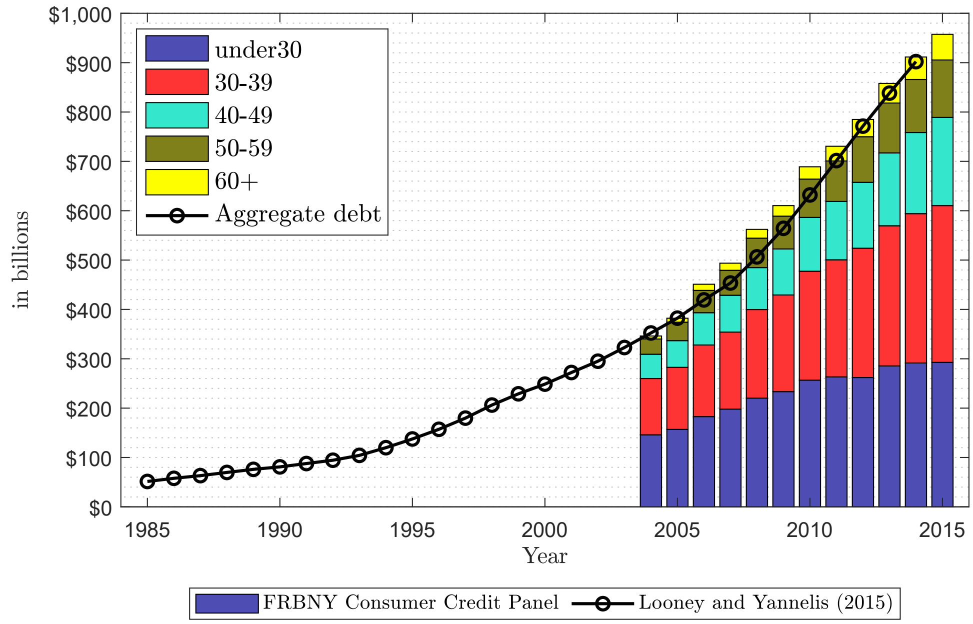


Figure 1.1: Total outstanding student debt balance

Note: Dynamics of the outstanding student debt in the U.S. between 1985 and 2015. The bar graphs show the aggregate amounts as well as the age composition of the debt (FRBNY Consumer Credit Panel). The black circled line indicates the aggregate total outstanding debt from the data (Looney and Yannelis (2015)). This amount is expressed in 2004 U.S. billion dollars.

focus on these sources as the total student debt balance reflects both changes in the borrowing behavior of college students and the repayment behavior of borrowers.[4] Importantly, to the best of our knowledge, this is the first paper in the literature that investigates the dynamics of student debt and simultaneously examines the implications of rising college costs, wage inequality, and the delinquent behavior of borrowers for the growing student debt in the U.S.

We build an incomplete-markets overlapping-generations (OLG) model in which individuals choose whether to pursue a college degree, how much they want to borrow from government student loans, and whether to be delinquent on the payments for student debt after graduation. Individuals born with different abilities and parental transfers decide to attend college or not. College education is costly: individuals pay the education cost, face the psychic cost of education, and lose out on four years of earnings. College students can finance their education through three different sources: parental transfers, labor income, and government student loans. Individuals enter the labor market with one of the two skill levels and receive

[4]Given that the federal government has increased its investment in higher education since the 1960s, rising student debt can be a natural consequence of the growing number of college-educated alongside rising wage inequality and education costs. On the other hand, it can reflect the financial distress borrowers face when paying off their existing student debt; therefore they choose to roll over their debt and increase the total balance in the economy.

skill-specific hourly wages and labor market experience premia. Once in the labor market, individuals with student debt have to pay it off following a fixed payment schedule. However, for each period, they can choose to be delinquent on their scheduled payment, with the utility cost of delinquency.

We solve the model's transitional dynamics with the following estimated time-varying sources: college wage premia, variances of persistent and transitory wage shocks, and college costs. The first two (wage inequality) are estimated using the 1968–2015 Panel Study of Income Dynamics (PSID) data, while the last is estimated using the National Longitudinal Survey of Youth (NSLY) 97 alongside data on tuition, fees, and costs of room and board (TFRB). We find that the college wage premia have almost doubled between 1968 and 2015, and the variance of wage shocks increased sharply during the same period. The average annual college education cost has increased from $\$4,723$ in 1979 to $\$12,000$ in 2015.[5] In addition, we lower the psychic cost of education over time, so the model can match the observed increase in college completion rates in the presence of changes in college costs and wage inequality.[6]

The calibrated economy is consistent with the observed distribution of college completion rates, education costs, and student debt amounts over ability terciles and parental transfer quartiles in the data. Matching the micro-level data to this level of rigor is unprecedented in the literature, but it is essential to study the secular increase in student debt balances over time. College completion rates increase in both ability and parental transfers. The net education cost is greater for individuals with high ability and large parental transfers, reflecting that they are more likely to attend a top-quality college with higher tuition. Furthermore, the amount borrowed by a college student is increasing with ability. This is because as ability is persistent through labor income, high-ability individuals wish to borrow more from future income to smooth out their consumption. Given that there are no other financial assets students can use to borrow from the future, they use student loans as a means.

We find that rising college costs, wage inequality, and delinquencies, in addition to a stronger preference for college education, increase student debt balances from $\$45$ billion in 1979 to $\$525$ billion in 2015. This explains 50% of the observed increase in the U.S. These sources explain most of the dynamics of student debt until 2007, suggesting the importance of other factors, such as the Great Recession or the 2006 student loan reform for the increasing student debt since 2007. Rising college costs largely determine the borrowing behavior of college students, leading a higher fraction of college students to borrow from government student

[5]These amounts are in 2004 U.S. dollars.

[6]The decreasing psychic cost of education is consistent with empirical findings in Moschini, Raveendranathan, and Xu (2023) that college students become overly optimistic about their graduation probability.

loans and a substantial amount when they do. At the same time, a more volatile wage shock process over time increases the delinquency rate of recent college graduates. This is because a more volatile wage shock process implies a higher probability of a persistent negative wage shock, preventing individuals from paying off their student debt as planned. In addition, a stronger preference for college education lowers the average ability of college graduates, which further raises the delinquency rate for student debt in the economy.

Crucially, and most importantly, we find that in a counter-factual economy without a delinquency choice, but with all the time-varying sources, aggregate student debt only increases by 50% of the increase in the benchmark economy. This is first because the no delinquency option discourages college students from borrowing as they do not have the option to defer their payments when they experience financial hardship. Second, without delinquencies, no student debt is rolled over. In sum, although rising college costs largely affect the borrowing behavior of college students, we find that delinquency significantly contributes to the rising student debt in the U.S.

The remainder of the paper is organized as follows. Section 1.2 discusses the related literature. Section 1.4 presents the model economy. Section 1.5 discusses the estimation and calibration. Section 1.6 presents the quantitative results, and Section 1.7 concludes.

1.2 Related Literature

This paper contributes to the literature that explores the implications of different student loan policies for educational attainment, borrowing, and repayment behavior in a quantitative macroeconomic framework. F. A. Ionescu (2008) studies the 2006 reform that eliminated a lock-in interest rate option for federal student loans in a model in which individuals face earnings and interest rate risks. F. Ionescu (2009) further quantifies the effects of flexible repayment options and relaxed eligibility for student loans on college enrollment, borrowing, and default rates. Our model is built on the recent paper by Abbott, Gallipoli, Meghir, and Violante (2019), which examines the effects of government grants and loans on college degree attainment, welfare, and the aggregate economy. However, we extend their model with a delinquency choice for student debt and focus on the dynamics of student debt rather than on educational policies.

Central to our analysis is the interaction among college attainment, parental transfers, and borrowing. A large body of literature discusses the role of credit constraints and family income in education decisions, including Belley and Lochner (2007), Hai and Heckman (2017), Keane and Wolpin (n.d.), Chetty, Friedman, Saez, Turner, and Yagan (2017), and Carneiro and Heckman (2002). For example, using the NLSY79, Keane and Wolpin (n.d.) and Carneiro and Heckman (2002) find that family income has a small effect on college attendance during the early 1980s. Consistent with this, our calibrated economy implies that few individuals are credit constrained for their college education, especially in 1979.

Another important component of student loans, especially in recent periods, is the private lending market. F. Ionescu and Simpson (2016) and Lochner and Monge-Naranjo (2011) explore the interaction between the private lending market and government student loans. Despite the recent importance, our work is agnostic about the role of private loans in the dynamics of student debt. This is because private loans account for only a relatively small fraction ($6 - 7\%$) of the total outstanding student debt.[7] In addition, the private lending market operates in a starkly different way from government loans, which is hard to introduce into the model. For example, in the private market, the terms, eligibility, and interest rates of loans depend on borrowers' default risk.

The interaction between student loans and labor market outcomes is essential for determining college graduates' repayment behavior with regard to student debt. Papers that examine the implications of student debt for job search and labor market outcomes include Ji (2021b), Luo and Mongey (2019a), Weidner (2016), and Rothstein and Rouse (2011b). Interestingly, Ji (2021b) finds that indebted agents spend less time on job search and end up with lower-paid jobs, while Luo and Mongey (2019a) and Rothstein and Rouse (2011b) find that college graduates with high student debt are likely to accept jobs with higher wages but lower job satisfaction. Although our model allows for an endogenous labor supply that depends on age, education, and student debt balances, given the complexity of the current model, we do not explicitly account for how student debt affects heterogeneous labor market outcomes among college-educated individuals.

The relationship between student loans and other forms of household debt has important policy implications. Mezza, Ringo, Sherlund, and Sommer (2020) use nationally representative administrative data to study the effect of student loan debt on subsequent homeownership and find that an increase in $\$1,000$ student loan debt lowers the homeownership rate by 1.8 percentage points for public 4-year college-goers. They find that the damage from credit scores from student debt delinquency is likely the channel through which student debt affects homeownership, and they suggest that student loan forgiveness or reduction in tuition growth may encourage borrowers to become homeowners. Using a general equilibrium overlapping generations model with student loans and credit card debt, Irwin and Kim (2023) quantitatively assess

[7] See Figures D2 and D3 in Appendix A.4.

the impact of allowing student debt to be liquidated in consumer bankruptcy on aggregate productivity via education decision, consumption smoothing behavior, and government budget constraint.[8]

1.3 Federal student loans program in the United States

The federal student loans program (FSLP) in the U.S. began in 1958 and substantially expanded under the Higher Education Act of 1965 to provide financial assistance for students in postsecondary and higher education. The program initially started as the Guaranteed Student Loan Program, also known as the Federal Family Education Loan Program (FFELP). Under this program, the banks or private institutions provided government-subsidized and guaranteed loans to students. In 1980, the Parent PLUS loan program was introduced for parents of dependent undergraduate students to help pay for college. In 1986, the Federal Loan Consolidation Program was first created. Under the consolidation program, students are allowed to combine all of their federal loans into one loan to simplify payments.

In 1992, the Higher Education Amendments introduced the Federal Direct Loan Program (FDLP) to improve the delivery of the student loans to college students. The difference between FDLP and FFELP is that the entitlements for FDLP accrue to individual borrowers while the entitlements for FFELP accrue to lenders and guaranty agencies. Also, with the Higher Education Amendments of 1992, the government began to provide different types of student loans under the FDLP. For instance, the government initiated Unsubsidized Stafford loans to all students regardless of their financial needs. Students under unsubsidized loans, unlike subsidized loans, have to cover the interest rates while in school. In 2005, the government allowed the graduate students to take student loans through the amendment to the Federal Parent PLUS loans. In 2010, the FFELP was officially terminated, and all students were required to take the FDLP.

The interest rates for student loans are set by the Congress. The rates are identical for unsubsidized and subsidized loans but are different between undergraduate and graduate borrowers. For undergraduate students, the interest rates were fixed, ranging from 6% in 1960s to 10% between 1998 and 1992. In 1992, the Congress enacted the variable rates, based on the interest rates on short-term US Treasury securities plus 3.1% markup rate, capped at 9%. From July 1st 2006, the interest rate was fixed at 6.8% for subsidized and unsubsidized loans. For subsidized loan interest rates, there was an incremental cut on interest rate from 2008 to 2012, from 6.8% to 3.4%. From July 1st 2013, all annual interest rate, both subsidized and

[8] While credit card debt can be discharged in a personal bankruptcy claim, student debt cannot be discharged.

unsubsidized loans, were linked to 10-year US treasury rate plus 2.05% markup rate, capped at 8.25%.

After leaving school, individuals with student debt begin the repayment of student loans after a six-month grace period . The standard repayment plan for student loans is a 10-year plan with fixed amount of repayment and can be extended up to 25 years. In addition, income-based and income-contingent repayment plans are available for borrowers with low incomes, but the take-up of income-driven repayment plans has been historically low (Looney & Yannelis, 2015). The eligibility for these programs depends on the type of loan, time of entry into borrowing and into repayment, and debt-to-income ratios.

Students are considered to be delinquent if they miss a payment. When students are delinquent for more than 270 days, their loans are considered to be in default. Unlike other types of consumer loans, student loans are almost impossible to discharge in bankruptcy. When in default, borrower's loans are assigned to a collection agency, and borrowers have to pay for late fees, collection costs, and accruing interest. The government also collects defaulted loans through wage garnishment up to 15% of the borrower's wages and through withholding tax refunds and other transfers such as Social Security or disability benefits.[9]

1.4 Model

1.4.1 Overview

The life-cycle of individuals consists of three stages: college education, work, and retirement. At age $j = 1$, an individual with different ability $\theta \sim \Theta(\theta)$ and parental transfers $a_0 \sim A(\theta, \xi_a)$ makes a college education decision. For parental transfers, we assume the following log-linear relationship to capture its positive correlation with ability in the data:

$$\log a_0 = \psi_0 + \psi_1 \log \theta + \psi_2 \xi_a, \quad \xi_a \sim N(0, 1), \tag{1}$$

where ξ_a is a stochastic component that follows a standard normal distribution.[10] Parental transfers are paid only once at the initial age for individuals who do not attend college, while those who enroll in college receive a fixed amount a_0 every period until graduation.[11]

[9] The Department of Education provides further information on types of repayment plans and penalties associated with default.

[10] Without a random component, there is a perfect correlation between ability and parental transfers. As a result, in contrast to the data, we do not include individuals with low (high) ability and high (low) parental transfers in the model. A random component is introduced to overcome this issue and allow more heterogeneity across individuals.

[11] Thus, college students receive a larger amount of parental transfers in total than those who do not attend college.

College education takes four years and lasts until age $J_c = 4$.[12] In each period, an enrolled college student has to pay net tuition and fees and devotes a fixed fraction of time $\bar{t}$ to studying. A college education can be financed through three different sources: parental transfers a_0, labor supply n, and government student loans b. There is no idiosyncratic risk during college, and thus no endogenous dropout. However, college graduates randomly lose their skills with a probability of ϕ_d before they enter the labor market. This is to capture the relatively high college drop-out rate in the U.S.

After college, an individual enters the labor market with one of the two distinct skill levels $e \in \{l, h\}$. A fraction $1 - \phi_d$ of college graduates become skilled workers ($e = h$), while those who lose their skills after graduation or do not attend college enter the labor market as unskilled workers ($e = l$). Skilled workers earn a higher hourly wage w_t^h than that earned by unskilled workers w^l.[13] Thus, the time-varying college wage premium is $w_t^c = \frac{w_t^h}{w^l}$. Skilled workers also receive higher labor market experience premia $l^h(j)$ over their working lives than unskilled workers $l^l(j)$. Finally, workers face idiosyncratic wage shocks $\varepsilon \sim Q_t(\varepsilon)$ every period.

Student debt is a non-dischargeable long-term debt in the U.S. Individuals need to make a total number of n_T repayments, following a fixed repayment schedule, to pay off their student debt. Each period, workers with existing student debt can be delinquent in their payments. However, those who delay their payments face a utility cost of delinquency χ_d. The remaining student debt after the delinquency choice accumulates interest r_b.

After retirement, individuals receive Social Security benefits $s_t^e(\theta, \varepsilon)$.[14] Retirees do not have the option to be delinquent and must pay off any remaining student debt following a fixed repayment schedule. All individuals retire at age $j = J_r$ and survive until age $j = J$.

1.4.2 College education

College education involves three types of costs. First, an individual has to pay yearly net tuition and fees:

$$\phi_t(\theta, a_0) = g_t^\phi(\phi_0 + \phi_1 log\theta + \phi_2 a_0), \text{ where } g_0^\phi = 1, \tag{2}$$

[12]For simplicity, we only study undergraduate students in the model. Student debt issued for graduate school accounts for a small fraction of the total debt in the U.S. For example, as shown in Figure D4, the fraction of the total federal student loans attributable to graduate study has been decreasing since 1980, and graduate students only represent 10% of the total federal loan borrowers.

[13]We normalize the hourly wage for unskilled workers to 1 in the model.

[14]This is time-varying as it is paid proportional to the wages that vary over time.

which is assumed to be a function of ability and parental transfers to be consistent with the heterogeneous education costs seen in the data.[15][16] Here, g_t^{ϕ} represents the growth rate of the education costs over time, with $g_0^{\phi} = 1$ at a steady state. Second, for every period, a college student needs to devote a fixed fraction of time $\bar{t}$ to studying, decreasing the total time endowment available for the labor supply during college education. Lastly, an individual who pursues a college degree faces a psychic education cost $\chi_{c,t}(\theta, a_0, \xi_\chi)$ that depends on its ability, parental transfer, and idiosyncratic preference shock ξ_χ, which is drawn from a standard normal distribution:

$$log\chi_{c,t} = (\chi_0 + \chi_1 log\theta + \chi_2 loga_0 + \chi_3\xi_\chi) - \Delta_t, \text{ where } \xi_\chi \sim N(0,1) \text{ and } \Delta_0 = 0. \tag{3}$$

This psychic cost of college education is essential to replicate the observed distribution of college completion rates (see Abbott et al. (2019), Cunha, Heckman, and Navarro (2004), and Heckman, Lochner, and Todd (2006)). In addition, to reproduce the empirically consistent increase in college completion rates in the U.S., we assume that the preference for college education becomes stronger over time, $\Delta_t = \varrho t > 0$.

A college student can finance its education through three different sources: parental transfers a_0, earnings from labor supply n, and government-subsidized student loans b.[17] Here, $b > 0$ is the cumulative student debt for the entire education period, and a college student chooses this debt amount in its first year of college. Under the Federal Student Loans Program (FLSP) in the U.S., students can borrow up to the cost of attendance (COA) minus the expected family contribution (EFC). To be consistent with this, the annualized borrowing amount $\frac{b}{J_c}$ cannot exceed a $\nu < 1$ fraction of the yearly tuition $\phi_t(\theta, a_0)$. Furthermore, the total amount borrowed cannot exceed the exogenous student loan limit $\underline{b}$ set by the government.[18] Finally, there is a marginal cost of borrowing student loans that follows a uniform distribution, $\xi_b \sim U(\underline{\xi_b}, \overline{\xi_b})$.

Without the stochastic components of the psychic cost of education and the marginal cost of borrowing, we do not have enough heterogeneity across individuals at the initial age. This makes the adjustment over the transition less smooth. If we instead assume a utility cost of borrowing, given the decreasing marginal

[15] In the NLSY97, the net education cost, measured as the total educational expenses paid minus grants and scholarships, is higher for wealthy or high-ability individuals than for wealth-poor or low-ability individuals. This reflects the fact that high-ability individuals and those from rich families are more likely to attend a top-quality college that charges higher tuition.

[16] For simplicity, we do not explicitly model differences in the quality of schooling. As a result, tuition is independent of the quality of education. However, by assuming tuition and earnings as a function of ability, we implicitly allow high-ability individuals who pay higher tuition for better education to earn more.

[17] In the data, we measure student loans using both subsidized and unsubsidized federal loans. In contrast, in the model, for simplicity, we assume that all loans are subsidized, as in F. Ionescu (2009). Except for the fact that the interest for subsidized loans is waived during college, there is little difference between the two types of loans.

[18] Note that students cannot borrow to consume.

utility in wealth, it disproportionately discourages high-ability individuals from borrowing relative to low-ability individuals. This makes the model inconsistent with the increasing amount of borrowing with ability in the data (see Figure 1.8). For ease of notation, we omit the time subscripts below.

Given the initial heterogeneities, the optimal college education decision at the beginning of life is

$$e(\theta, a_0, \xi_\chi, \xi_b) = \begin{cases} h & \text{if } V_1^h(\theta, a_0, \xi_b) - \chi_c(\theta, a_0, \xi_\chi) \geq EV_1^l(\theta, a_0, \varepsilon') \\ l & \text{otherwise,} \end{cases}$$

where $V_1^h - \chi_c$ is the value of an individual going to college, while $EV_1^l(\theta, a_0, \varepsilon')$ is the expected value of an individual entering the labor force as a high school graduate. Below, we explain the optimization problem of a college student in each school year in detail.

Freshmen

The optimization problem of the first-year college student is

$$\begin{aligned} V_1^h(\theta, a_0, \xi_b) &= \max_{c,n,b} u(c, n) + \beta V_2^h(\theta, a_0, b, \xi_b) \qquad (4) \\ \text{s.t. } c &= a_0 + \pi w^l \kappa_{j=1}^h(\theta, \varepsilon = 0)n - \phi(\theta, a_0) + \frac{b}{J_c} - \xi_b b \\ 0 &\leq \frac{b}{J_c} \leq \min\{\nu\phi(\theta, a_0), \underline{b}\} \\ c &\geq 0, \quad n \in (0, 1 - \bar{t}]. \end{aligned}$$

Here, $\pi < 1$ is a discounting factor for labor income during college, which is calibrated to match the average earnings during college. College students earn an unskilled hourly wage w^l times efficiency units of labor,

$$log\kappa_j^e(\theta, \varepsilon) = p^e log\theta + l^e(j) + \varepsilon, \qquad (5)$$

that is determined by the skill-specific labor market experience premium, ability, and idiosyncratic income shock.[19] Note that the optimization problems of sophomores and juniors are the same as that for freshmen, except that b is not a choice variable.

[19]We assume that work experience in college does not increase the labor market experience premium, $l^h(1)$. However, the number of years in college is counted as labor market experience once students graduate. In this way, we can use age as a proxy for a worker's labor market experience regardless of education level.

Seniors

The optimization problem of a senior is

$$
\begin{aligned}
V_{J_c}^h(\theta, a_0, b, \xi_b) &= \max_{c,n,a'} u(c,n) \\
&+ \beta\Big[\phi_d EV_{J_c+1}^l(i=0,\theta,a',b,\varepsilon') + (1-\phi_d)EV_{J_c+1}^h(i=0,\theta,a',b,\varepsilon')\Big] \\
&\text{s.t. } c + a' = a_0 + \pi w^l \kappa_{j=1}^h(\theta, \varepsilon=0)n - \phi(\theta, a_0) + \frac{b}{J_c} - \xi_b b \\
&a' \geq 0, \quad c \geq 0, \quad n \in (0, 1-\bar{t}],
\end{aligned}
\tag{6}
$$

where ϕ_d is the exogenous probability of losing skills after graduation. When college graduates lose their skills, they enter the labor market as unskilled workers, but still carry the accumulated student debt for college education. Here, i represents the number of repayments made for student debt. For example, if $i = 0$, an individual has not begun repaying their outstanding student debt. Finally, college seniors can save in financial assets a.[20]

1.4.3 Long-term student debt, delinquency choice, and workers

In the United States, a student loan is a non-dischargeable long-term debt.[21] Thus, individuals must make a total number of n_T repayments, following a fixed repayment schedule, to pay off their debt. Following this, in the model, individuals have to pay

$$\lambda_i = \frac{1}{n_T - i}, \quad i = 1, \ldots, n_T$$

fraction of existing debt in every repayment period.[22] Note that the values of λ_i increase with the number of repayments to guarantee a similar amount of repayment for every repayment period.[23]

[20] Individuals who do not attend college can save parental transfers in financial assets at the initial age. Unless the same thing is allowed, individuals with greater parental transfers may counterfactually prefer not to attend college. To avoid this, we allow senior students to save in financial assets.

[21] Under Chapters 7 and 13 bankruptcies in the U.S., a student loan is non-dischargeable. A student debt holder is considered to be in default once the payment is overdue for more than 270 days. In default, the line of credit is shut down, and the default status is reported to credit bureaus. Government agencies that guarantee student loans make a repayment plan, including penalties for the defaulter, such as wage garnishment and seized federal tax refunds. Including all these penalties, the level of debt under the new repayment plan can be as high as 125% of the original principal (see F. Ionescu (2009)). If a debtor cannot pay off all their student debt after 25 years of repayments, the remaining debt can be forgiven only if the debtor has a very low income.

[22] In the U.S., borrowers are required to start repaying their student debt 6 months after graduating or dropping out of college.

[23] Without interest accumulation, these λ_i values ensure the same repayment amount every repayment period.

As previously stated, for every period, workers with student debt choose whether to be delinquent on the payments that are due. Thus, the discrete choice problem of a worker is

$$V_j^e(i,\theta,a,b,\varepsilon) = \max\left\{V_j^{e,p}(i,\theta,a,b,\varepsilon),\ V_j^{e,np}(i,\theta,a,b,\varepsilon)\right\},$$

where $V_j^{e,p}$ is the value of repaying the balance that is due during this period, while $V_j^{e,np}$ is the value of delaying the payment.

To be specific, the optimization problem of an individual who decides to repay is

$$V_j^{e,p}(i,\theta,a,b,\varepsilon) = \max_{c,n,a'} u(c,n) + \beta E V_{j+1}^e(i+1,\theta,a',b',\varepsilon') \tag{7}$$

$$\text{s.t.} \quad c + a' = (1+r)a + w^e \kappa_j^e(\theta,\varepsilon)n - \lambda_i b$$

$$b' = (1+r_b)(1-\lambda_i)b$$

$$a' \geq 0, \quad c \geq 0, \quad n \in (0,1].$$

As this individual makes a repayment in this period, the total number of repayments made increases to $i+1$ for the next period. In addition, the interest accrues on the remaining balance of the student debt after repayment.[24]

The optimization problem of an individual who decides to be delinquent is

$$V_j^{e,np}(i,\theta,a,b,\varepsilon) = \max_{c,n,a'} u(c,n) - \chi_d + \beta E V_{j+1}^e(i,\theta,a',b',\varepsilon') \tag{8}$$

$$\text{s.t.} \quad c + a' = (1+r)a + w^e \kappa_j^e(\theta,\varepsilon)n$$

$$b' = (1+r_b)b$$

$$a' \geq 0, \quad c \geq 0, \quad n \in (0,1].$$

Note that a worker who does not attend college ($e = l$) solves the same problem as equation (7), with $b = 0$ and $i = 0$.[25]

[24]In Appendix A.3, we introduced wage garnishment when borrowers are delinquent on their payments.

[25]We assume that high school graduates and college dropouts face the same hourly wage w^l and labor market experience premia $l^l(j)$ during the working period.

1.4.4 Retirees

After retirement, individuals receive Social Security benefits $s^e(\theta, \varepsilon)$, proportional to their labor income at their last working age and education level. Retirees do not have the option to be delinquent and must pay their remaining debt according to the repayment schedule. To summarize, the optimization problem of a retiree is

$$V_j^e(i, \theta, a, b, \varepsilon) = \max_{c, a'} u(c, 0) + \beta V_{j+1}^e(i+1, \theta, a', b', \varepsilon) \tag{9}$$

$$\text{s.t.} \quad c + a' = (1+r)a + s^e(\theta, \varepsilon) - \lambda_i b$$

$$b' = (1 + r_b)(1 - \lambda_i)b$$

$$a' \geq 0, \quad c \geq 0.$$

1.5 Taking the model to the data

In this section, we show how we discipline our model using the data. First, we explain how we estimate the time-varying wage processes and education costs that are used as inputs for transitional dynamics. We also show how we estimate the distribution of ability and the gradient of ability on earnings. Finally, we discuss the calibration strategy.

1.5.1 Estimation

Time-varying wage processes

We estimate the wage process using the 1968–2017 PSID data.[26] The estimation procedure closely follows Kim (2022) and Heathcote, Storesletten, and Violante (2010). We first run the following OLS regression to estimate time-varying college wage premia w_t^c and skill-varying labor market experience premia

[26] We use survey data from 1968 to 2017, but estimate the variances of wage shocks only through 2015 because of the finite sample bias at the end of the sample period.

$l^e(j)$ over the sample period.[27]

$$\log w_{i,j,t,e} = \sum_{t=1}^{T} \beta_{t,0} D_t + \sum_{t=1}^{T} \beta_{t,1} D_t D_{h,t} + \sum_{e=l,h} \left(\beta_{e,2} D_{e,t} \iota_{i,j,t,e} + \beta_{e,3} D_{e,t} \iota^2_{i,j,t,e} \right) + \widehat{r}_{i,j,t,e}$$

Here, $w_{i,j,t,e}$ represents the hourly wage of an individual i at age j with education level e in year t. $D_{e,t}$ are the education dummies that take the value of 1 if the education level is $e \in \{l, h\}$, with l representing the non-college educated and h representing the college-educated. We regress the log hourly wage on time dummies D_t, an interaction term with time dummies and college education dummy $D_{h,t}$, an interaction term with education and labor market experience ι, and an interaction term with education and experience-squared ι^2.[28]

The estimated college wage premia and skill-specific labor market experience premia are shown on the left and right, respectively, in Figure 1.2. First, the college wage premia have almost doubled, increasing from 1.25 in 1979 to 1.42 in 2015. This is the first time-varying input in the model. Second, the wage growth is much steeper for the college-educated than for the non-college educated. For example, 25 years of labor market experience more than doubles the hourly wage for college-educated workers (red dashed line), while it only increases by 50% for non-college educated workers relative to their initial levels (black solid line).

Next, we estimate the time-varying wage shock process using the minimum distance method.[29] The regression residuals $\widehat{r}_{i,j,t,e}$ are assumed to be the sum of idiosyncratic wage shocks $\varepsilon_{i,j,t,e}$ and measurement error.[30] The idiosyncratic shocks consist of a persistent component η and a transitory component ε^v. Specifically,

$$\varepsilon_{i,j,t,e} = \eta_{i,j,t,e} + \varepsilon^v_{i,j,t,e}$$

$$\eta_{i,j,t,e} = \rho \eta_{i,j-1,t-1,e} + \varepsilon^p_{i,j,t,e},$$

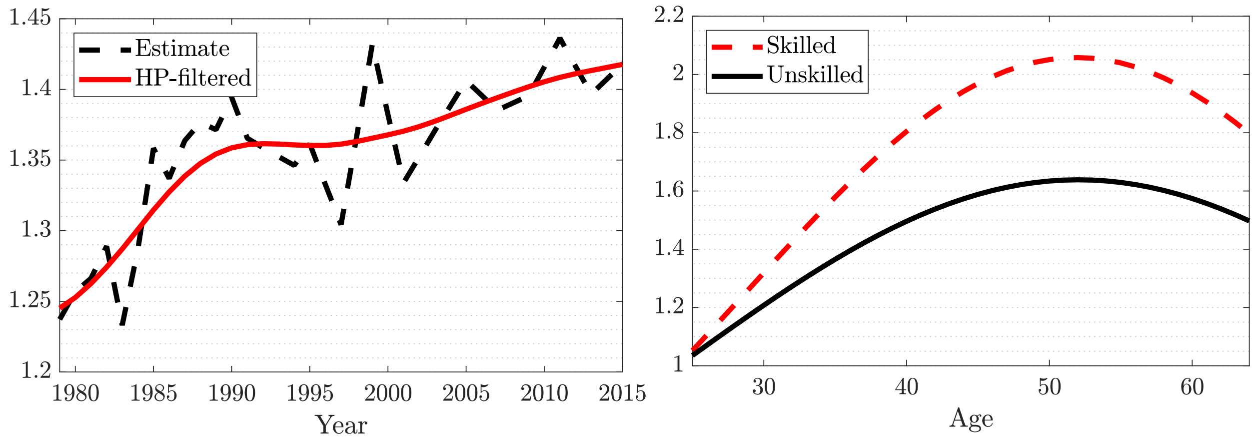
where $\varepsilon^p_{i,j,t,e} \sim N(0, \sigma^2_{pt})$ and $\varepsilon^v_{i,j,t,e} \sim N(0, \sigma^2_{vt})$. We estimate these year-varying variances of shock $\{\sigma^2_{pt}, \sigma^2_{vt}\}$ and the persistence of the shock $\{\rho\}$.

[27] As emphasized by Guvenen (2009), introducing heterogeneous income profiles (HIP) is important to capture realistic income inequality.

[28] The labor market experience ι is measured as the age minus years of schooling minus 6. In years missing the variable for years of schooling, we proxy years of schooling using the median of education brackets for individuals with less than a college degree. For example, if individuals responded that they finished grades 6–8, we approximate years of schooling for this individual as 7. For individuals with a college degree or more, we proxy their years of schooling as 16.

[29] See Appendix A.2 for the details of the estimation.

[30] We use the estimate from French (2004) for the variance of a measurement error in log hourly wages of 0.02.



(a) College wage premia

(b) Labor market experience premia

Figure 1.2: Estimated between-group wage dispersion

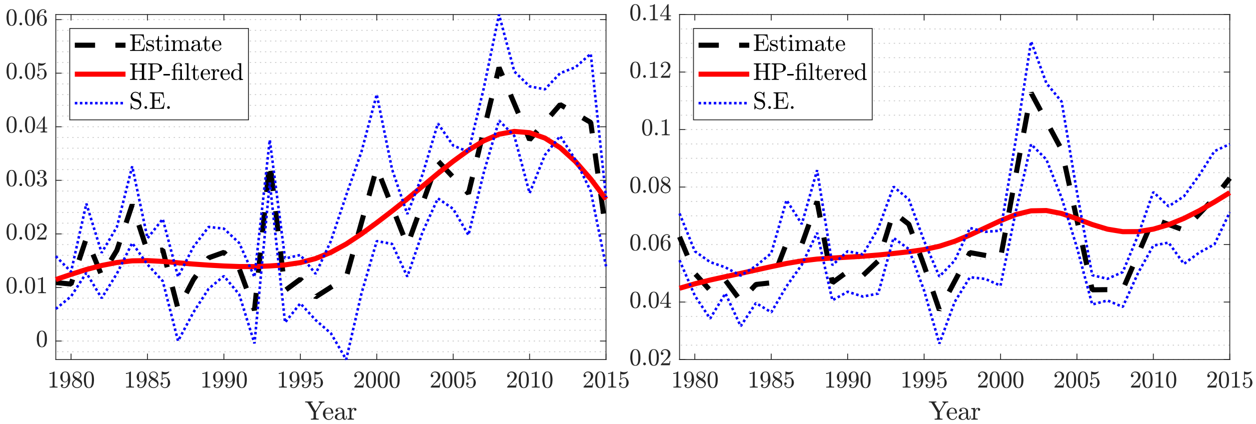
Note: The left panel shows the estimated time-varying college wage premia, $w_t^c = exp(\beta_{t,1})$. The black dashed line is a raw estimate and the red solid line is an Hodrick-Prescott (HP) filtered series with a smoothing parameter of 100. The right panel shows the labor market experience premia, $l^e(j) = exp(\beta_{e,3}\iota + \beta_{e,4}\iota^2)$, for college graduates (red dashed line) and non-college graduates (black solid line).

Figure 1.3 shows the estimated skill-varying variances of shocks, which is the second time-varying input in the model. It shows that there has been an increase in residual wage dispersion, reflecting the rising wage inequality in the U.S. For example, the variance of the persistent shock increases from 0.01 in 1979 to 0.03 in 2015, while that of the transitory shock rises from 0.045 to 0.08 over the same period. The estimated persistence of the shock is $\rho = 0.9792$.

Education costs

Following Gordon and Hedlund (2020), we calculate the average yearly net college cost in the NLSY97. The NLSY97 provides education expenses financed by financial aid and loans from family and friends, federal and other student loans, work-study financial aid, employer assistance financial aid, other financial aid, and out-of-pocket spending. We do not include college costs financed by grants and scholarships to measure the net education cost. The NLSY97 cohort, on average, paid $\$7,737$ for each year of a college education.

Figure 1.4 (a) shows gross and net TFRB indices, relative to the 1997 levels, between 1979 and 2015. To estimate changes in net college costs over this period, we multiply the average net college cost in the

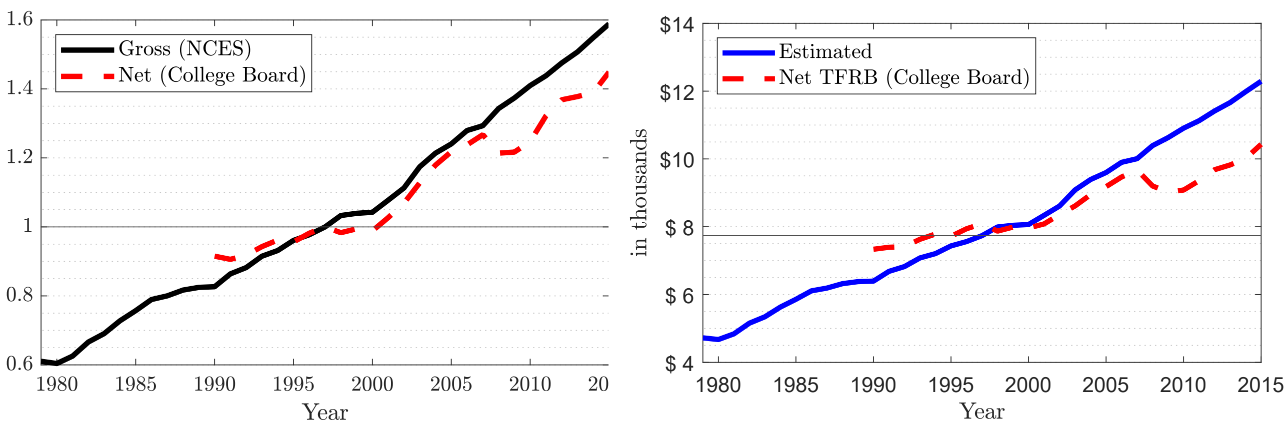


(a) Persistent shock

(b) Transitory shock

Figure 1.3: Variances of persistent and transitory shocks

Note: Minimum distance estimates of the wage shocks. The black dashed lines are the estimates, and the red solid lines are the HP-filtered trends with a smoothing parameter of 100. The blue dotted lines are standard errors estimated using a block bootstrapping with 300 replications.



(a) TFRB indices (base year = 1997)

(b) In 2004 U.S. dollars

Figure 1.4: Tuition and fees in the U.S.

Note: The gross TFRB data are from Table 330.10 in the 2017 NCES Digest of Education Statistics, while the net TFRB data are from the 2018 Trends in College Pricing published by the College Board. The data moment is the enrollment weighted average of 2-year public, 4-year public, and 4-year private non-profit institutions.

NLSY97 by the gross TFRB index in Figure 1.4 (a).[31] The estimated time-varying cost successfully reproduces the observed net TFRB in College Board for the available years, as shown in Figure 1.4 (b). Importantly, college costs have risen rapidly in the U.S., increasing from $\$4,723$ in 1979 to $\$12,000$ in 2015.

[31]Because the net TFRB data are available only since 1990, we use the gross TFRB index instead of the net TFRB index. However, as shown in Figure 1.4 (a), the gross and net TFRB indices are similar.

Ability distribution and ability gradient on earnings

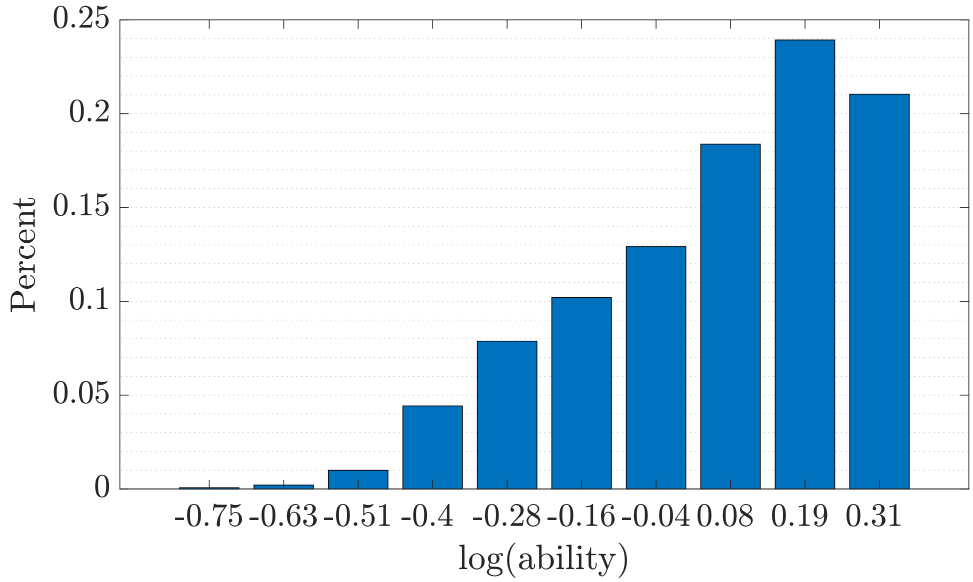


Figure 1.5: Ability distribution

Note: This figure shows the probability distribution of $log(AFQT80) - E[log(AFQT80)]$ over 10 discretized points, where $E[log(AFQT80)]$ is the average of $log(AFQT80)$. The AFQT80 scores are from the NLSY79

We estimate the distribution of ability from the Armed Forces Qualification Test (AFQT) score in the NLSY79 (see Figure 1.5). Next, to estimate the gradient of ability on earnings p^e in equation (5), we regress the log hourly wages in the NLSY79 on the time dummies and logAFQT80 scores, controlling for the age effect using the labor experience premia $l^e(j)$ estimated from the PSID data.[32] Table 1.1 summarizes the estimated results, which show that the effect of ability on earnings is almost two times higher for skilled workers than for unskilled workers.

Table 1.1: Estimated ability gradient on earnings

Education level	Gradient
College graduates	0.9724 (0.0929)
Less than college	0.5658 (0.0221)

Note: Standard errors are in parentheses.

[32]We use the PSID estimates to filter out age effects, as we do not have many individuals older than 57 in the NLSY79 sample.

1.5.2 Calibration

Parameters set externally

Demographics Assuming a model period of 1 year, we calibrate the model economy to the 1979 U.S. economy.[33] In the model, individuals begin their life at age 18 ($j = 1$), retire at age 65 ($J_r = 47$), and live until the age 85 ($J = 67$) with certainty.

Education and student loans During college education, individuals spend a fraction $\bar{t} = 0.25$ of their time studying (Abbott et al. (2019)). Individuals lose their skills acquired in college with the probability $\phi_d = 0.3$ at graduation.[34] Similar to the FSLP, we assume that the fixed loan payment period is $n_T = 13$ years. The interest rate for federal student loans is determined by the risk-free interest rate plus 3.1%. We choose a risk-free interest rate r as 3.0%, such that r_b becomes 6.1%, similar to the historical average of federal student loan rates.[35] The borrowing limit $\underline{b}$ is chosen to match the cumulative student loan limit of $\$23,000$ for the four years of college education in 2004 U.S. dollars.[36] Lastly, using the ratio of the average family contribution to the college expense in the NLSY97, we set $\nu = 0.7$.

Preferences Individuals face a standard separable utility

$$u(c,n) = \frac{c^{1-\sigma}}{1-\sigma} - (\psi + \psi_c \mathbf{1}_{j \le J_c,\, e=h}) \frac{n^{1+\eta}}{1+\eta} \tag{10}$$

with the coefficient of relative risk aversion σ set to 2. η is chosen to match the Frisch elasticity of labor supply for male, and there is an additional disutility of work during college ψ_c.

Social security payments Following Kim (2022), social security is paid proportional to the productivity shock in the last working age ε.[37] Specifically, the social security payment function is

$$s^e_t(\theta,\varepsilon) = \delta_s w^e_t \frac{\sum_{j=1}^{J_r-1} \kappa^e_j(\theta,\varepsilon)}{Jr-1} \overline{n} \tag{11}$$

[33] We choose 1979 as a benchmark year such that individuals who are 18 years old in 1979 reach their 50s by the end of the transition period. If we choose 1997 instead as a benchmark year, another available cohort for the NLSY data, an individual who enters college in 1997 only turns 30 in 2015.

[34] We measure the college drop-out rate as one minus the graduation rate among college-enrolled students.

[35] The loan rate is the same for both subsidized and unsubsidized loans. In 1992, the FSLP introduced a variable interest rate. In 2006, a fixed rate of 6.8% was introduced but reverted to a variable rate in 2013. During the variable-rate period, the loan rate ranged between 4% and 8%. Note that in the model with a fixed risk-free rate, there is no distinction between fixed and variable interest rates.

[36] This limit applies to the sum of subsidized and unsubsidized loans.

[37] In the U.S., social security benefits are paid based on the average of the highest 35 years of earning. In the model, calculating average earnings requires one more state variable, making computation more challenging. Given that the wage shock is highly persistent, we assume that social security is paid proportional to the wage shock in the last working age.

with a replacement rate $\delta_s = 0.4$ and the average hours worked in the economy $\overline{n} = 0.3$. Table 1.2 below summarizes the parameters set externally.

Table 1.2: Parameters set externally

Parameters	Description	Value
$\bar{t}$	A fixed fraction of time for studying (Abbott et al. (2018))	0.25
ϕ_d	College drop-out rate (OECD)	0.3
n_T	Student loan payment periods (FSLP)	13
σ	CRRA parameter	2.0
r	Risk-free interest rate	0.03
r_b	Student loan rate (FSLP)	0.061
$\underline{b}$	Student loan limit in 2004 dollars (FSLP)	$\$23,000$
$1-\nu$	Average Expected Family Contribution (NLSY97)	0.3
δ_s	Replacement rate for social security benefits	0.4
η	Frisch elasticity of male labor supply	0.50

Parameters set internally

Table 1.3 summarizes the parameters and moments calibrated inside the model. First, the time discount factor β is calibrated to match the capital-to-output ratio. ψ and ψ_c are calibrated to match the average hours worked for males between the ages of 25 and 65 and the average hours worked during college, respectively. We also reproduce the average yearly earnings of $\$8,936$ during college which is close to $\$8,377$ in the NLSY79. Second, we match the observed college attainment rates over ability terciles and parental transfer quartiles.[38] In the aggregate, 16% of the population has a college degree. Next, we target the average parental transfers of $\$4,720$ and the average annual college cost of $\$4,723$ in 2004 U.S. dollars, which is estimated in Section 1.5.1. Given these, 29% of undergraduate students graduate with student debt in the model, compared to $20-30\%$ in the data.

The calibrated economy generates an annual delinquency rate of 11%, which is measured as the fraction of borrowers who are delinquent on their payments. Similarly, using American Bankers Association data, Volkwein, Szelest, Cabrera, and Napierski-Prancl (1998) estimate the annual delinquency rates of student loan borrowers to be around 17–21$\%$ in the 1980s and early 1990s. Finally, we target the total outstanding

[38] Unlike the NLSY97, the NLSY79 does not have parental transfer information. Thus, in the NLSY79, we proxy parental transfers using family income.

student debt to the GDP, which amounts to \$45 billion in 2004 U.S. dollars.[39]

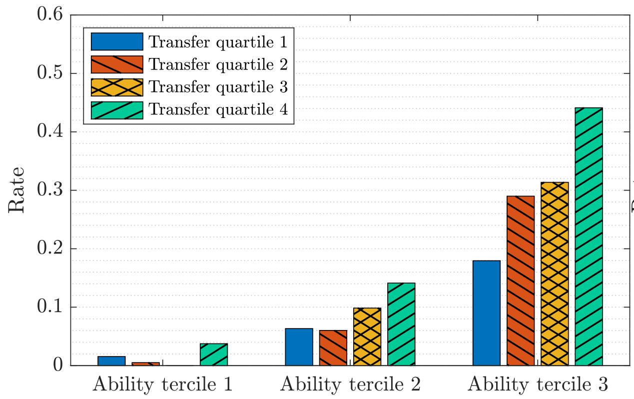
Table 1.3: Parameters set internally

	Values	Moment	Data	Model
β	0.95	Capital to output ratio	3.20	3.17
ψ	70	Average hours worked of male	0.30	0.31
ψ_c	480	Average hours worked during college	0.19	0.23
ψ_0	-2.64	College completion rate 1T	0.01	0.01
ψ_1	0.22	College completion rate 2T	0.09	0.12
ψ_2	0.23	College completion rate 3T	0.34	0.34
ϕ_0	-1.57	College completion rate 1Q	0.05	0.04
ϕ_1	0.50	College completion rate 2Q	0.11	0.11
ϕ_2	1.10	College completion rate 3Q	0.16	0.18
χ_0	3.00	College completion rate 4Q	0.27	0.29
χ_1	-29.8	Aggregate college completion rate (%)	15	16
χ_2	-12.1	Average annual parental transfers (\$)	$4,720$	$4,776$
χ_3	8.70	Average annual education cost (Figure 1.4) (\$)	$4,723$	$4,827$
χ_d	0.07	Delinquency rate (%)	10	11
$\overline{\xi_b}$	0.25	Total outstanding student debt to GDP (Figure 1.1)	0.0092	0.0081
$\underline{\xi_b}$	0.22	Fraction of undergraduates borrowing '93 (%)	20 - 30	29
π	0.47	Average earnings during college (\$)	$8,377$	$8,936$

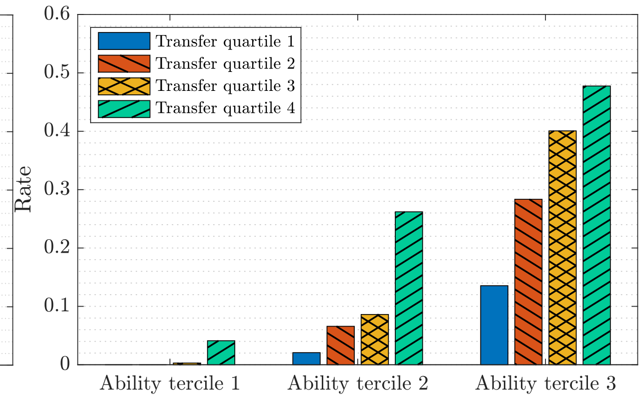
Note: The college completion rates are across ability terciles (T) and parental transfer quartiles (Q). The average annual parental transfers are from the NLSY97. The delinquency rate is from Volkwein et al. (1998). The fraction of undergraduates borrowing is from Cuccaro-Alamin and Choy (1998). The college completion rates, average hours worked, and average earnings during college are from the NLSY79.

Figure 1.6 further compares the college completion rates by ability terciles and parental transfer quartiles in the NLSY79 (left) to those in the model (right). Similar to Lochner and Monge-Naranjo (2011), the figure shows that the college completion rate is increasing in ability (parental transfers), conditional on parental transfers (ability).

[39]For student debt, data availability is limited. As in Looney and Yannelis (2015), the earliest data for the total outstanding student debt go back to 1985, and thus, we use their 1985 value for the calibration.



(a) Data



(b) Model

Figure 1.6: College completion rates in 1979

Note: The college completion rates over ability terciles and parental transfer quartiles, calculated from the NLSY79 (left) and the benchmark economy (right). In the NLSY79, we define a college completion rate as the percentage of individuals between the ages of 14 and 22 in 1979 who completed at least 16 years of education by 1988. We use family income as a proxy for parental transfers.

1.6 Quantitative results

This section presents quantitative results from transition dynamics with time-varying sources: college wage premia (Figure 1.2 (a)), variances of wage shock (Figures 1.3), and college costs (Figure 1.4).[40] In addition, as mentioned in Section 1.4.2, we calibrate ϱ in

$$\Delta_t = \varrho t \tag{12}$$

to match the observed rising college completion rate in the U.S. between 1979 and 2015 (see Figure 1.12 (a)).[41,42]

1.6.1 Comparison of the model to the data

Thus far, we have been agnostic about the model's performance on the distribution of college costs and students' borrowing amounts, which is essential for the model to be a valid framework for studying the dynamics of student debt. This is because, unlike the NLSY97 data, the NLSY79 data do not simultaneously provide information on student debt amounts, parental transfers, and college costs that are key to disciplining

[40] We used HP-filtered series for college wage premia and wage shock variances.

[41] Reproducing the observed changes in college choice is crucial for explaining the dynamics of student debt.

[42] Numerically, to ensure that the model converges to a new steady state, we solve transitional dynamics for 150 periods. We fix time-varying parameters to 2015 levels for periods after 2015.

our model. To overcome this problem, in this section, we compare the model's simulated moments in 1997 in the transition to data moments. Note that the model is calibrated for the steady state and that all the moments in the transition are not directly targeted.

The results in Table 1.4 show that the borrowing behavior of new college graduates in 1997 in the model is similar to that in the data, reproducing 58% of graduating seniors holding student debt and an average cumulative student debt of $\$10,507$ upon graduation, compared to 60% and $\$11,348$, respectively, in the data.

Table 1.4: Untargeted moments

Moment	Data	Transition (1997)
Graduating seniors with student debt (%)	60	58
Average cumulative student debt upon graduation	$11,348	$10,507

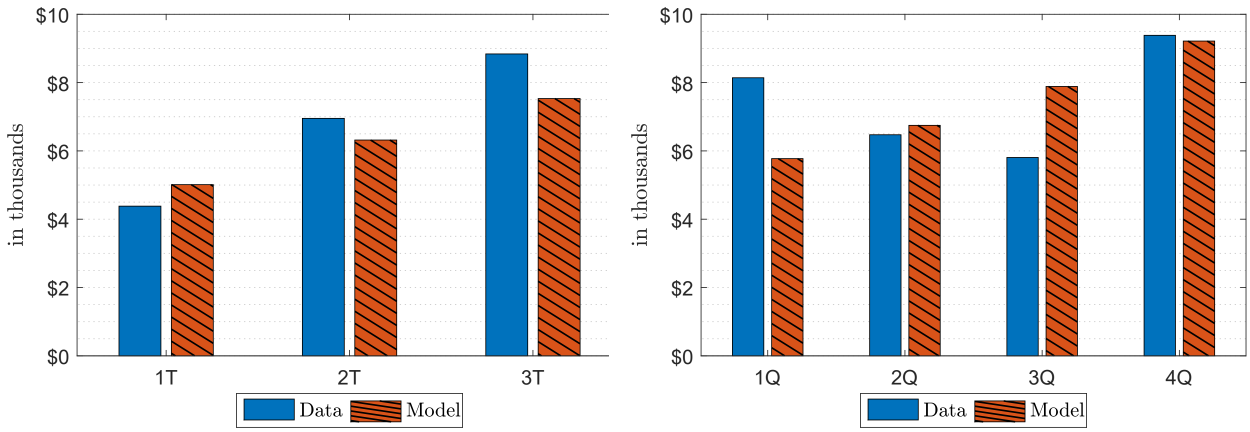
Note: The percentage of graduating seniors with student debt is obtained from the NPSAS and the average student debt upon graduation is obtained from the NLSY97.

In Figure 1.7, we also compare the distribution of the net college cost in the NLSY97 to that in the model across ability terciles (left) and parental transfer quartiles (right).[43] Importantly, in the data, the observed college expenses are higher for students in higher ability tercile group and the fourth parental transfer quartile group, reflecting heterogeneous costs of college education across individuals. Similarly, the model reproduces the distribution of college costs that increase with ability and parental transfers. This is driven by the fact that students with high ability or from rich families are more likely to attend a top-quality college that charges high tuition.

Figure 1.8 further exhibits the average cumulative student debt at graduation over ability terciles (left) and parental transfer quartiles (right).[44] In the NLSY97, a significant amount of student debt is held by individuals across all ability terciles and parental transfer quartiles. More importantly, the borrowing amount increases in student's ability and its parental transfer. Without targeting, the model successfully reproduces these facts.

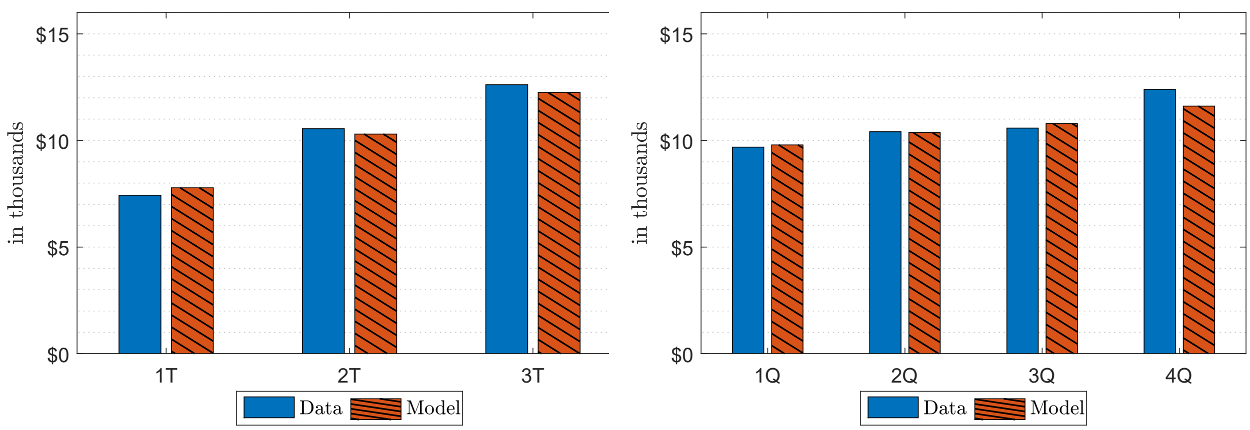
[43] Note that, in contrast to the NLSY79, parental transfer information is available in the NLSY97. See Appendix A.1 for more details of how we calculate parental transfers in the NLSY97.

[44] The NLSY97 reports the amount of student loans that an individual takes for each institution and academic term. For each individual, we calculate the yearly average by summing all the student loans across terms, years, and institutions until graduation and divide it by the number of years the individual is enrolled full-time in college. To calculate the cumulative student debt at graduation, we multiply the yearly average of student loans by 4. See Appendix A.1 for more details.



(a) over ability terciles (b) over parental transfer quartiles

Figure 1.7: Distribution of net college costs in 1997



(a) over ability terciles (b) over parental transfer quartiles

Figure 1.8: Distribution of student debt in 1997

1.6.2 Simulation

To show the implications of the model, we simulate the steady state and present the average net tuition, effective borrowing limit, and borrowing amount of college students over their initial abilities in Figure 1.9 (a). Bars show the fraction of college graduates at each ability level, circles indicate where each ability tercile group is divided, and purple dashed lines show each ability group's average earnings during college education.

Given our baseline parametrization, Figure 1.9 (a) shows that for each ability group, the average earnings are much larger than their tuition. This suggests that in 1979, most college students, if they want to, could finance their college education through labor income and were not constrained for their college education (see Carneiro and Heckman (2002), Keane and Wolpin (n.d.), and Lochner and Monge-Naranjo (2011) etc).

This also implies that low-ability students' decisions not to attend college were optimal, not because they were credit constrained for college investment.

Figure 1.9 (a) also shows that the borrowing amount increases with students' abilities. As ability is persistent through labor income, high-ability individuals wish to borrow more from future income. This is because they expect a higher lifetime income and wish to smooth out their consumption. As there is no other financial asset college students can use to borrow from the future, high-ability students use student loans as a means.

Figure 1.9 (b) shows how these implications change in 1997.[45] As tuition increases, so does the borrowing limit. As a result, a higher fraction of college students borrow from government student loans, and on average, they borrow a larger amount, compared to 1979. A stronger preference for college education encourages low-ability individuals to attend college, decreasing the average ability of college students.

Finally, although not shown in Figure 1.9, there is a 15 percentage points increase in the fraction of student borrowers who max out their government student loans between 1979 and 1997.[46] Note that students who max out government loans are not necessarily credit constrained for their schooling. As seen in the data, these are high-ability individuals who attend colleges with higher education costs and borrow from the high future income.[47]

The borrowing behavior of college students is less affected by parental wealth, as shown in Figure 1.10.[48] This is because the borrowing behavior is largely driven by the consumption smoothing motive rather than credit constraint for college investment.[49] Thus, with larger parental transfers, individuals reduce their labor supply instead of changing the borrowing amount to finance their college education.

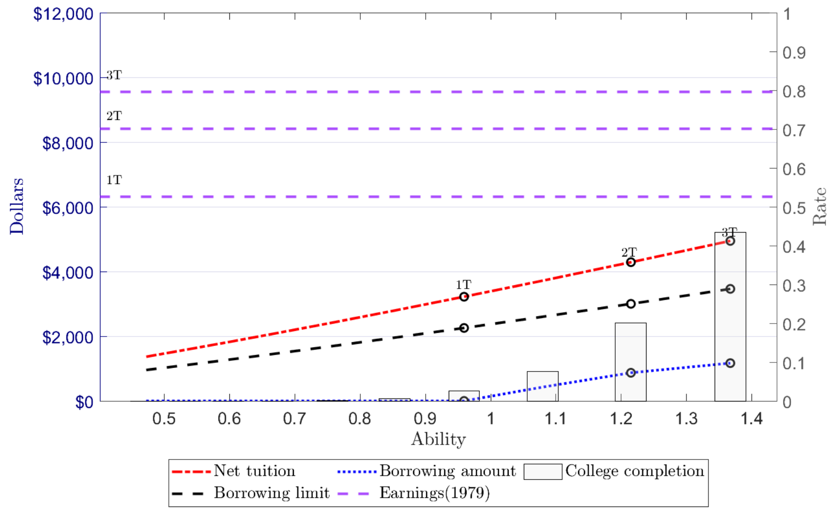
[45] Note that this is not a steady-state comparison. We simulate the model over the transition and calculate the moments for the year in transition, which is equivalent to 1997.

[46] Similarly, in the National Center for Education Statistics (NCES) data, the fraction of dependent student borrowers who max out their loans increased from 30% in 1989 to 50% in 1993 (Refer to Figure 6 in Trends in Undergraduate Borrowing: Federal Student Loans in 1989–90, 1992–93, and 1995–9). This may be why increasingly recent college students borrow from private lenders.

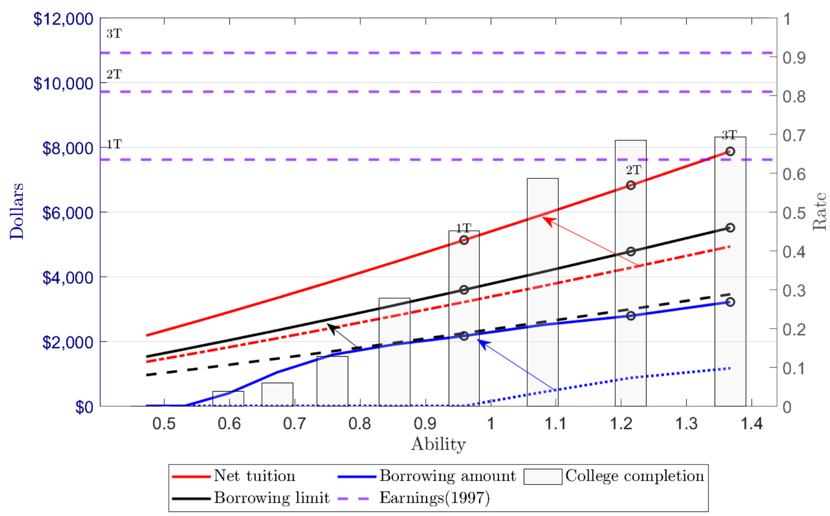
[47] Identifying whether individuals are credit-constrained for schooling in the data is not trivial. We can identify the effects of credit constraints on schooling when marginal students who did not go to college change their schooling decisions with better access to credit or larger parental transfers. In the model, when we relax the borrowing constraint allowing students to borrow up to the entire education cost ($\nu = 1.0$), the college completion rates rarely change from the benchmark economy ($\nu = 0.7$).

[48] See Figure D5 for the corresponding figure for the 1979 economy.

[49] The college enrollment rate is increasing in parental transfers because of a positive correlation between ability and parental wealth, not because individuals are credit constrained for their college investment.



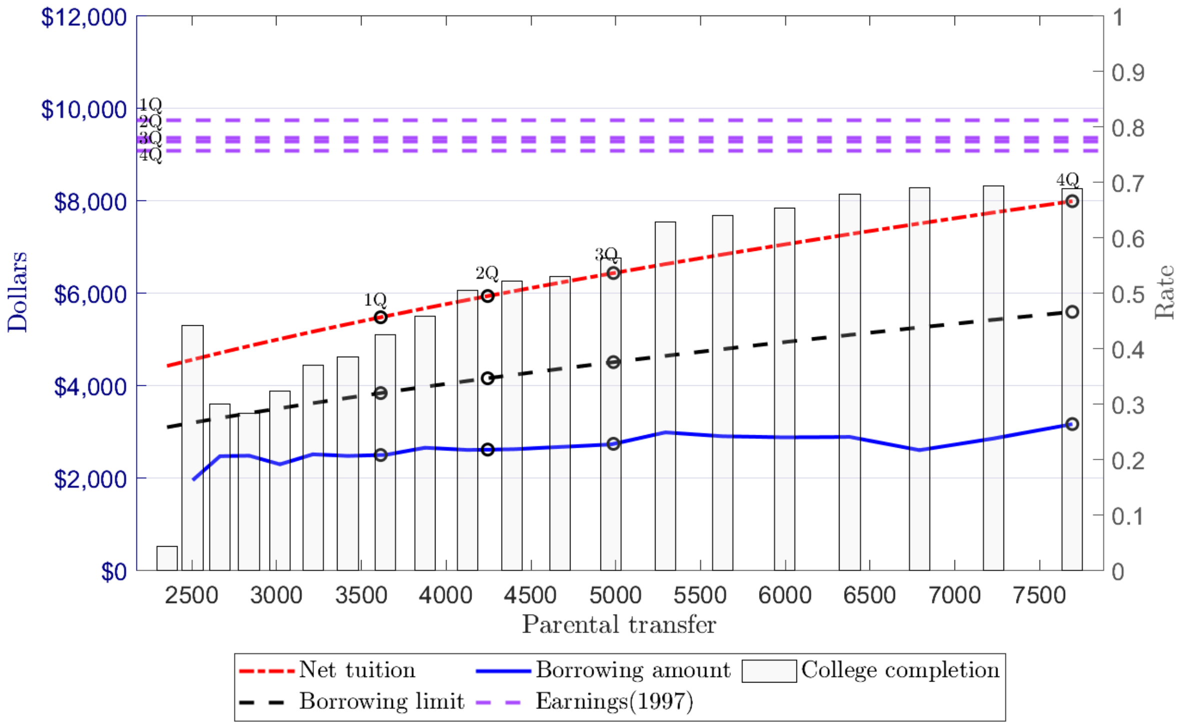
(a) 1979 economy



(b) 1979 vs. 1997 economies

Figure 1.9: Simulation over ability

Note: This figure shows the simulated annualized moments over the initial ability for the year 1979 (top) and compares these moments with those for the year 1997 (bottom). The red dash-dot line in the top figure represents the average net tuition and fees for the year 1979; the black dashed line represents the effective borrowing limit for the year 1979, and the blue dotted line represents the borrowing amount for the year 1979. The solid lines in the bottom figure represent the corresponding moments for the year 1997. The white bars represent the college completion rates, and the black circles indicate where the ability tercile is divided. The left y-axis indicates the dollar values of moments, while the right y-axis indicates the college complication rates.



(a) 1997 economy

Figure 1.10: Simulation over parental transfers

Note: This figure shows the 1997 simulated moments over parental transfers. The red dash-dot line represents the average net tuition and fees, the black dashed line represents the effective borrowing limit, and the blue line represents the borrowing amount. The white bars represent the college completion rates, and the black circled marks indicate where the parental transfer quartile is divided. The left y-axis indicates the dollar values of moments, while the right y-axis indicates the college complication rates.

1.6.3 Dynamics of student debt, college choice, and borrowing

In this section, we present the aggregate results from transitional dynamics. Figure 1.11 first shows the dynamics of the aggregate student debt in the benchmark economy (bars) and in data (line). Quantitatively, rising college costs, wage inequality, and delinquencies, in addition to a stronger preference for college education, increase the aggregate student debt balance from $45 billion in 1979 to $525 billion in 2015. This is approximately 50% of the observed increase in student debt in the U.S. Notably, these sources explain much of the dynamics of student debt until 2007. However, the aggregate student debt balance in the model starts to deviate from that in the data after 2007. This suggests that some other factors, such as the Great Recession or the 2006 reform, affect the dynamics of the U.S. student debt since 2007.[50]

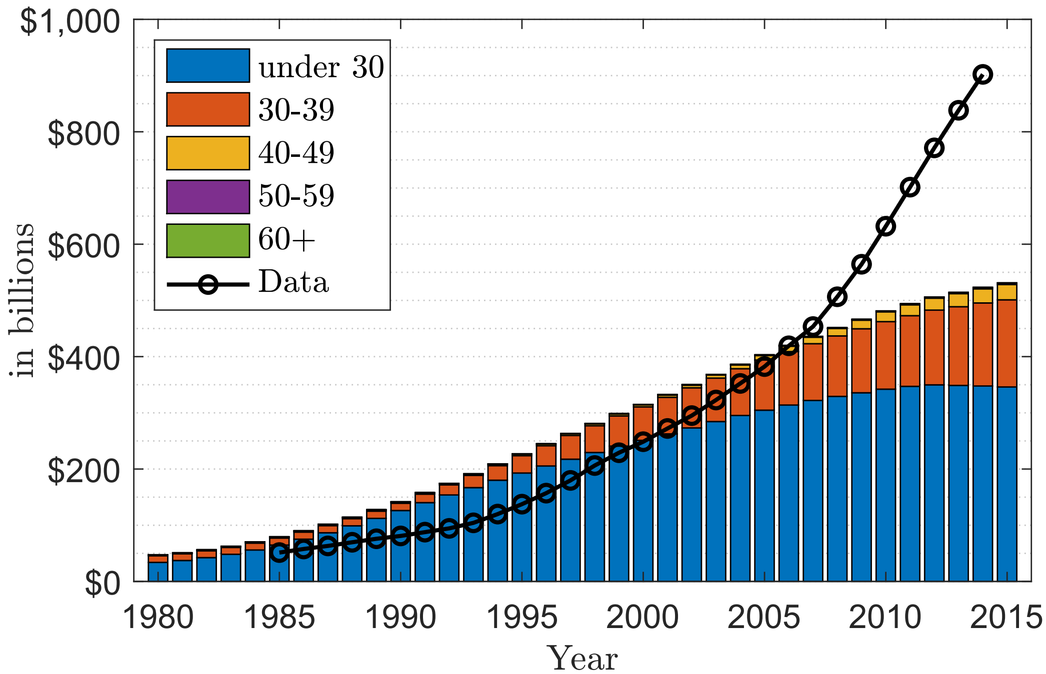


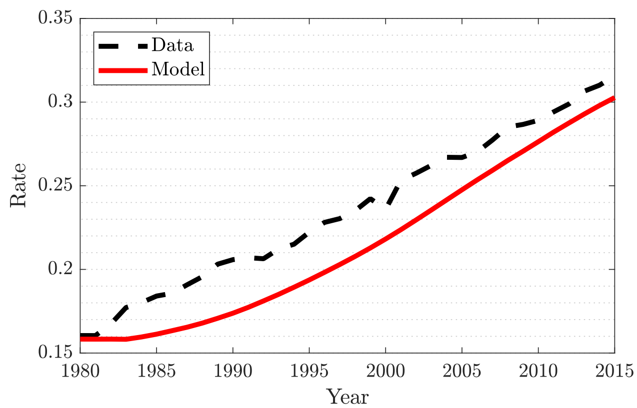
Figure 1.11: Dynamics of student debt

Note: In this figure, the dynamics of aggregate student debt over the transition (bars) are compared to those in the data (line).

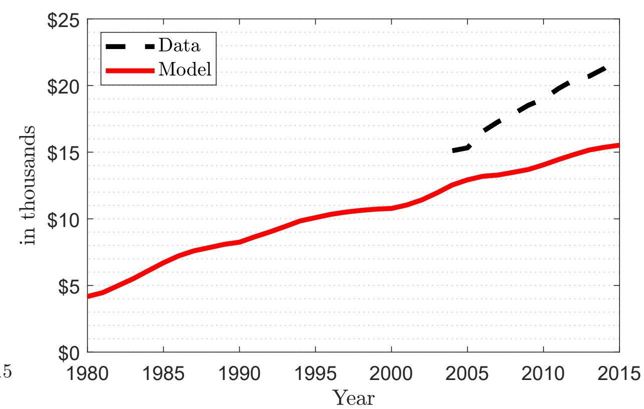
Figure 1.12 further compares changes in the college completion rates, average cumulative student debt amounts at graduation, and fraction of graduating seniors with student debt in the model to those in the data. It also shows the average ability of college students in Figure 1.12 (d). As mentioned before, the

[50]During the Great Recession, a number of individuals went back to college or graduate school. In addition, the adverse impact on the financial and labor markets increased the delinquency rates of student debt holders. Furthermore, in 2006, the Higher Education Reconciliation Act of 2005 removed the lock-in interest rate option for student debt and fixed the interest rate on student loans to 6.8%. As shown in F. A. Ionescu (2008), this reform has increased the overall delinquency rate of student debt holders.

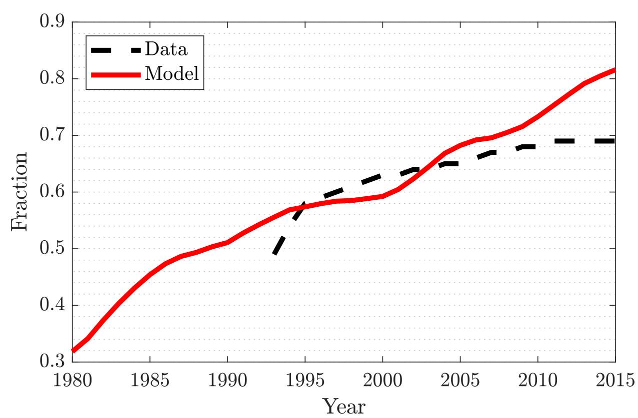
time-varying college completion rate is reproduced by the decreasing psychic cost of education. It is noteworthy that, without targeting, the model is also broadly consistent with the increasing borrowing amount by students and the fraction of students graduating with student debt.[51] Finally, an increasing preference for a college degree lowers the average ability of college students in the model. As we show in Section 1.6.5, this leads to a higher delinquency rate among borrowers.



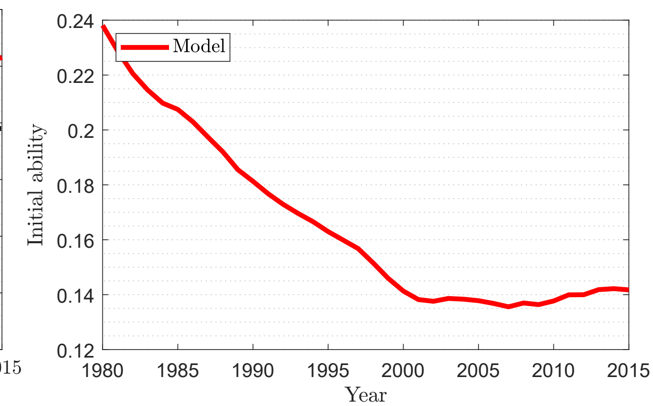
(a) College completion rate



(b) Average student debt at graduation



(c) Fraction of graduating seniors with student debt



(d) Average initial ability

Figure 1.12: Changes in college choice and borrowing behavior

Source: The college completion rates are from the CPS and the NLSY79. The average student debt at graduation and the fraction of graduating seniors with student debt are from the National Postsecondary Student Aid Study (NPSAS).

1.6.4 Decomposition exercise

In this section, we examine how each time-varying source affects the results by conducting counterfactual experiments in which we remove one of the time-varying components each time. We show the

[51]Note that, in 2008, the student loan limit increased to $\$31,000$, which is omitted from the model. This may explain the gap in the borrowing amount between the model and the data in Figure 1.12 (b).

dynamics of college completion rates, the average cumulative student debt amounts at graduation, delinquency rates, and the aggregate student debt in Figure 1.13. We further summarize these changes over the transition in the benchmark and counter-factual economies in Table 1.5.

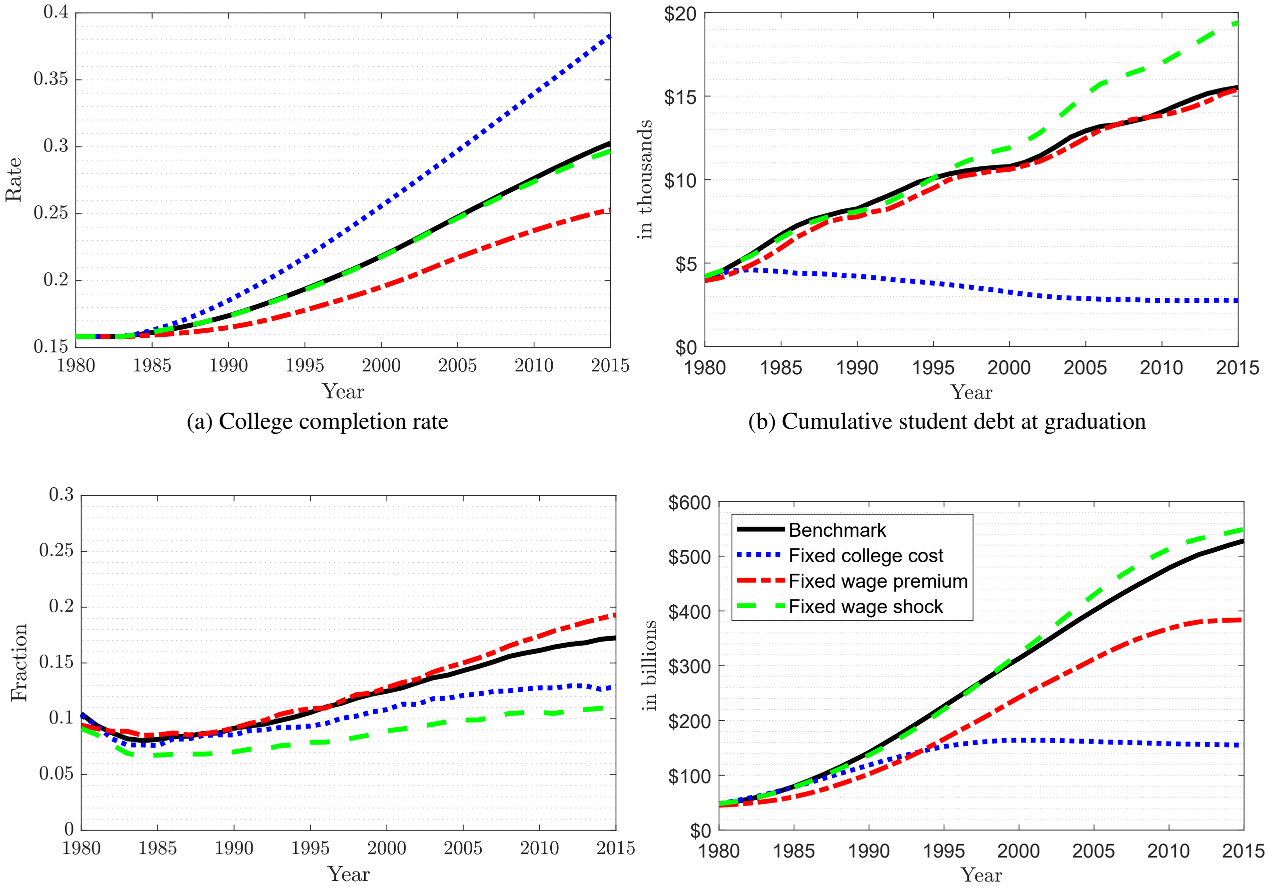
First, without rising college costs, the college completion rate further increases to 38%. However, students borrow less from government loans compared to the benchmark economy. The fraction of borrowers delinquent decreases by 5 percentage points as individuals carry a moderate amount of student debt over their lifetime. In net, the aggregate student debt increases by only \$105 billion, despite a larger fraction of people with a college degree. Second, as expected, the increasing college wage premia is important for explaining the rising college completion rate in the model. Without it, the college completion rate only increases by two-thirds of the rise in the benchmark economy. However, the mitigated increase in college wage premia barely affects the borrowing amount by students and delinquency rates in the economy. The aggregate student debt still increases significantly by \$335 billion, due to the increased college attendance and costs. These results imply that the increasing college costs plays an important role in the growth of student debt in the U.S., as it determines the borrowing behavior of college students. However, as we show in Section 1.6.5, if individuals pay off their debt without being delinquent, the aggregate student debt increases by only half of the growth in the benchmark economy, even with all the time-varying sources.

Although the time-varying wage shock process barely affects college choice, students borrow more from government loans and borrowers are less likely to be delinquent on their payments when variances of wage shocks, especially persistent components, do not rise. This is because a more volatile wage shock process implies a higher probability of a persistent negative wage shock. When such a shock is realized, individuals have a hard time paying off their student debt as planned.

Table 1.5: Changes over the transition

1979–2015	Δ (a)	Δ (b)	Δ (c)	Δ (d)
Benchmark	$+14\%$	+11.6K	$+6\%$	+480B
Fixed college costs	$+22\%$	-1.2K	$+2\%$	+105B
Fixed college wage premia	$+9\%$	+11.5K	$+10\%$	+335B
Fixed variances of wage shock	$+14\%$	+15.5K	$+1\%$	+500B

Note: Table 1.5 shows the changes in (a) college completion rates, (b) cumulative student debt at graduation, (c) fraction of borrowers delinquent, and (d) aggregate total student debt over the transition between the benchmark model and the counter-factual economies. The letter K indicates a thousand, and the letter B indicates a million.



(a) College completion rate

(b) Cumulative student debt at graduation

(c) Fraction of borrowers delinquent

(d) Aggregate student debt

Figure 1.13: Decomposition exercises

Note: The black solid line is the benchmark economy. The blue dotted line is the economy without time-varying college costs. The red dashed line is the economy without time-varying college wage premia. The green dash-dot line is the economy without the time-varying wage shock process.

1.6.5 Delinquency choice

An increasing delinquency rate over time and individuals rolling over their student debt are crucial for explaining the rapid growth of the student debt in the U.S. To show this, in this section, we explore the implications of delinquency in the model.

First, we present the fraction of workers who are delinquent on their payments after graduation over labor income quartiles in 1997 in Table 1.6. The model predicts that workers in the lower distribution of labor income face a higher chance of delaying their payments. This is in contrast to the fact that high-ability individuals, who are more likely to be in the top income distribution, are the ones borrowing the most from student loans (see Figure 1.8).

Table 1.6: Delinquency rate across earnings quartile in 1997

Earnings quartile	Q1	Q2	Q3	Q4
Delinquent borrowers	69.4	15.3	0.3	0.0

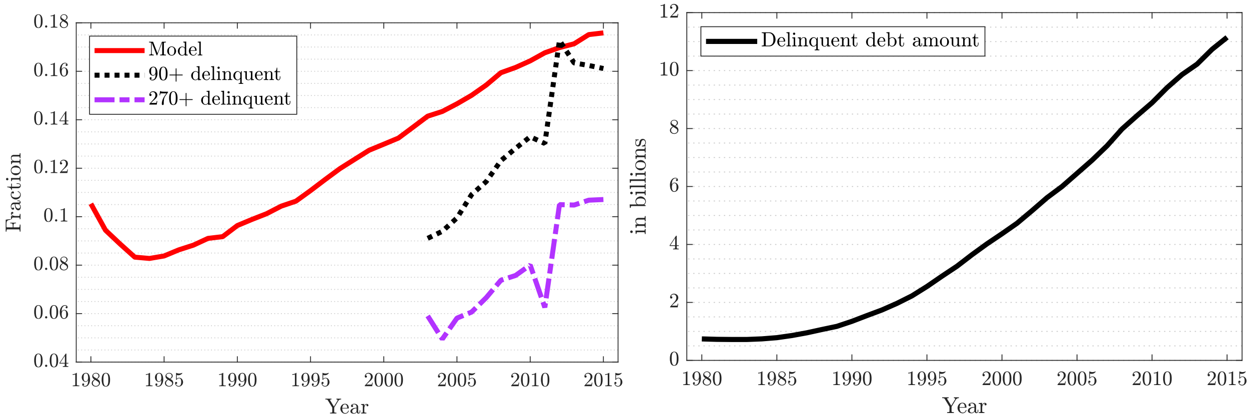
To show how the delinquency rate changes over time, Figure 1.14 exhibits the fraction of borrowers who are delinquent on their payments on the left and the fraction of student debt that is delinquent on the right. As seen in the left figure, without targeting, the model is qualitatively consistent with the observed dynamics of the aggregate delinquency rate-the fraction of borrowers who are delinquent on their payments 90 days or more.[52,53] The aggregate delinquency rate initially decreases as the college wage premia increase. However, as the residual wage dispersion increases, and the average ability of college graduates declines, the delinquency rate almost doubles between 1979 and 2015.

Finally, we quantify the role of delinquency choice in the growing student debt by solving the model without a delinquency option but with all the time-varying sources. We do not re-calibrate the model to make it comparable to the benchmark economy. Importantly, without delinquency choice, the total student debt only increases by $250 billion, compared to $480 billion in the benchmark economy, as shown in Table 1.7. This is because, as shown in Figure 1.15, in the counter-factual economy without delinquency, college students reduce their borrowing from student loans, as they do not have the option to defer their

[52] Despite the model period of 1 year, given that the model abstracts from additional penalties, such as wage garnishment, for 270 days or more delinquency, delinquency in the model represents 90 days or more delinquency.

[53] To compare, in Appendix A.3, we also solve the model with wage garnishment when borrowers are delinquent on their payments and calibrate the delinquency rate to the 270+ rate. All the quantitative results remain similar between the two models.

Figure 1.14: Fraction of delinquent borrowers and debt



(a) Fraction of delinquent borrowers (b) Fraction of delinquent debt

Note: In (a), the red solid line represents the moments from the model, and the black dashed line and purple dash-dot line represent 90+ delinquent and 270+ delinquent rates from the FRBNY data, respectively.

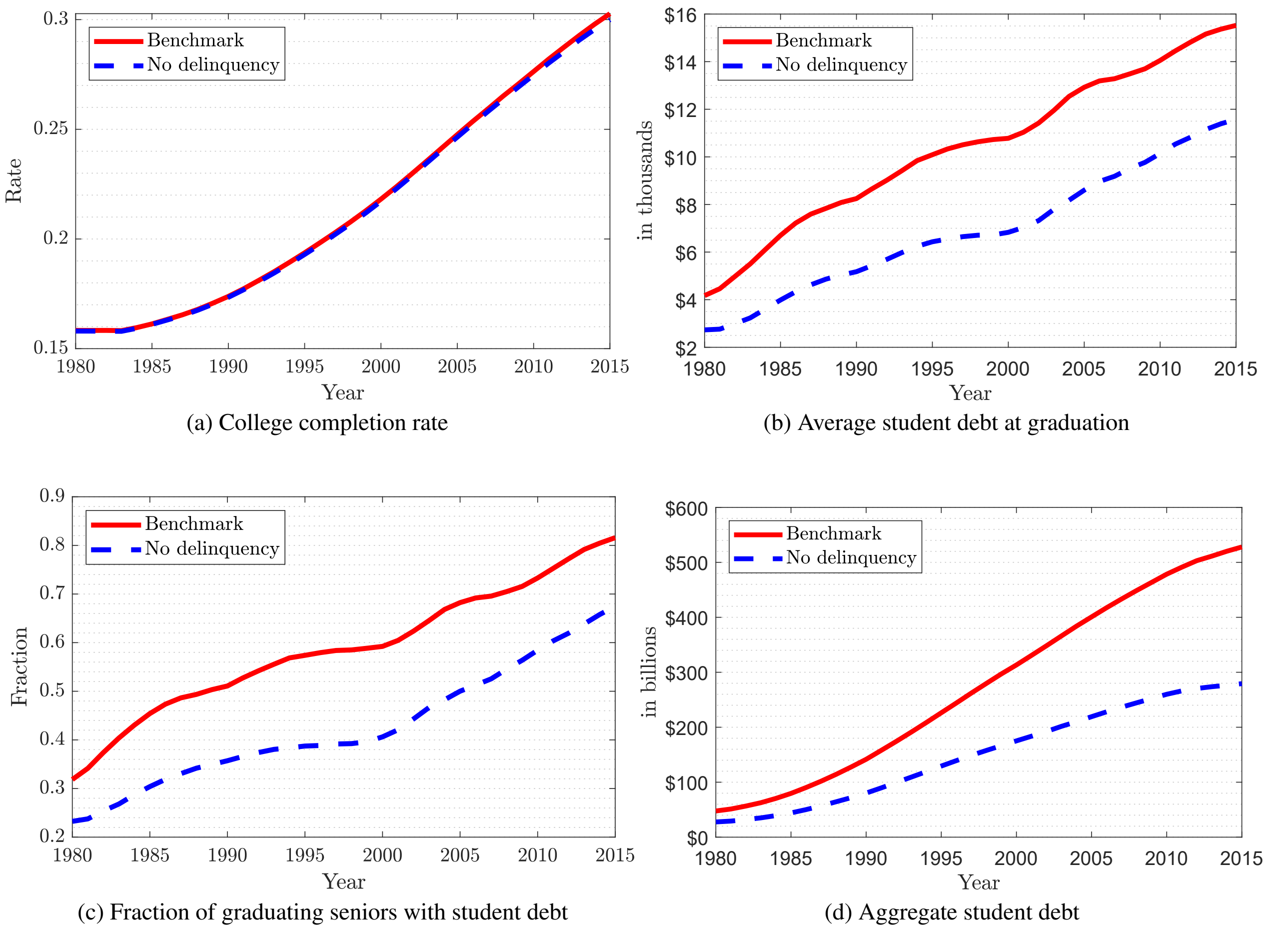
payments when they experience financial hardship.

In sum, although the rising college costs encourages students to borrow a larger amount for their college education, if individuals pay off the debt following the repayment schedule without being delinquent, the aggregate student debt does not increase as much as in the benchmark economy. However, a more volatile wage shock process and the lower ability of college graduates lead a higher fraction of borrowers to delay their payments. As a result, the total student debt balance rises sharply.

Table 1.7: Changes in total outstanding student debt

1979–2015	Δ Student debt
Benchmark	+480B
No delinquency choice	+250B

Note: The table shows the total changes in the student debt in the benchmark model and in the model without the delinquency option.



(a) College completion rate

(b) Average student debt at graduation

(c) Fraction of graduating seniors with student debt

(d) Aggregate student debt

Figure 1.15: Changes in college choice and borrowing behavior

Note: This figure shows (a) the college completion rate, (b) the average cumulative student debt at graduation, (c) the fraction of college graduates with student debt, and (d) the aggregate student debt in the benchmark economy (red solid) and no delinquency economy (blue dashed).

1.7 Conclusion

This paper examines the quantitative effects of rising college costs, wage inequality, and the delinquent behavior of borrowers on the growing student debt in the U.S. We study this by building an incomplete-markets OLG model with choices for a college education, student loans, and delinquency. Solving transitional dynamics with estimated increases in college costs and wage inequality, in addition to the increasing preference for college education over time, we find that the benchmark economy explains 50% of the observed growth in U.S. student debt between 1979 and 2015. Rising college costs largely determines the borrowing behavior of college students. The delinquency rate increases over time as the average ability of college students declines and the residual wage dispersion increases. Importantly, we find that the aggregate student debt balance only increases by half of that in the benchmark economy if borrowers can reimburse their debt without being delinquent. This suggests that increasing delinquency rates play a vital role in increasing aggregate student debt balances.

Chapter 2

An Empirical Analysis of Student Debt and Initial Labor Market Outcomes

The cost of higher education in the United States has soared over the past 20 years, and student debt, which allows financially constrained students to access higher education, has roughly quadrupled. More than 60% of U.S. college graduates enter the labor market with student debt, which is non-forgivable and has to be paid off according to a repayment schedule after a grace period of six months. The financial burden that comes from student debt repayment can be cruel to college graduates who have just entered the labor market, as they have to finance other expenditures, such as housing, cars, food, and more, with limited financial resources in the initial stage of working life. More specifically, it can influence college graduates' initial labor market decision, through which they finance repayment and other expenditures.

This paper empirically analyzes the effect of student debt on the initial labor market outcomes in the United States. Specifically, it examines how student debt affects college graduates' initial unemployment duration, occupational choice, and initial salary. For occupational choice, I focus on underemployment, which is defined as choosing an occupation that requires less than a 4-year college degree.[1] Studying this effect helps us understand whether college graduates with student debt face financial burdens from student debt repayment. Also, it gives policymakers insights into student loan policies that could reduce the student debt burden.

[1] Leuven and Oosterbeek (2011) illustrate the overview of the literature on overeducation since the 1970s and indicates that around a third of American workers are overeducated for their jobs and that overeducation results in a substantial reduction in the returns to schooling. Also, Figure 2.1 shows that around 40 to 50 percent of recent college graduates are underemployed since 1990.

Conducting an empirical study on the impact of student debt on the three initial labor market outcomes requires a dataset that simultaneously provides information on education, student debt, employment, and earnings. A dataset that has all this information is rare. To overcome this data challenge, I merge the following two datasets. First, I use the National Longitudinal Surveys of Youth 1997 (NLSY97), which provides college graduates' student loans and employment status, including the exact week at which they started working after college graduation as well as the title of their occupation. Then, I merge the NLSY97 dataset with the Occupational Information Network (O*NET) database. This database contains detailed information on more than $1,000$ occupations in the U.S., including the required education level for each occupation that is reported in the NLSY97. The merged dataset allows me to identify college graduates who take the first job that requires less than a 4-year college degree. Such an education mismatch in employment is also known as underemployment.

I examine the effect of student debt on initial unemployment duration using survival analysis. First, I estimate the Kaplan-Meier (KM) function to study the one-to-one relationship between holding student debt and unemployment duration. I find that student debt holders are more likely to become employed (i.e. shorter unemployment duration) in a given week compared to students with no debt. This result, however, does not specify whether the individuals become underemployed or non-underemployed after spending a certain initial unemployment duration after college graduation. Underemployment is different from employment because the latter does not specify the educational requirement of a job. Also, underemployment and non-underemployment are mutually exclusive. Thus, to further examine how holding student debt affects the initial unemployment duration to underemployment at the first job, I estimate the cause-specific Cumulative Incidence Function (CIF), which estimates the incidence of the occurrence of underemployment while taking non-underemployment into account. I find that student debt holders are more likely to become underemployed in a given week compared to student debt non-holder, taking into account that they can exit from unemployment to both underemployment and non-underemployment.

The above nonparametric estimations cannot control for other observed covariates that can affect the labor market outcomes. To incorporate other covariates into the analysis, I estimate a Cox-proportional hazards model. This model allows me to look at the effect of student debt on a college graduate's probability of exiting the unemployment period to the first job. I find that a 10% increase in student debt is associated with a 0.24% increase on average in the probability of getting the first job after exiting the initial unemployment period.

The above Cox-proportional hazards model can only study one event (i.e. employment) and cannot be used if there is more than one event. For example, college graduates can exit unemployment to underemployment if their first job requires less than a 4-year college degree. At the same time, they face the possibility of becoming non-underemployed. To study underemployment, I estimate the cause-specific hazards model. This model examines the effect of student debt on the probability of exiting unemployment to underemployment, taking into account the competing risk of exiting unemployment to non-underemployment. I find that, given that a college graduate is unemployed and can face one of two competing risks at a given week, a 10% increase in student debt is associated with a 0.18% increase on average in the probability of becoming underemployed at the first job after 4-year college graduation.

Endogeneity issue can arise in the above analyses because of unobserved factors that affect both student debt and labor market outcomes. This situation is referred to as endogeneity. For instance, students with high ability may expect higher future earnings by attending high-quality colleges and thus would borrow more student loans to smooth out consumption during college education.[2] These individuals would be less likely to be underemployed and more likely to earn a higher income. On the other hand, students with low ability and less financial resources to finance college education would want to borrow student loans. Given that they are more likely to earn low labor income in the future and that they need to repay outstanding student debt, they may choose to take a job as soon as possible, even if the job does not require a 4-year college degree, to keep up with the repayment. For these reasons, the results from OLS or hazards model analyses could likely underestimate (or overestimate) the impact of student debt.[3]

To overcome the endogeneity issue, I use the two-stage residual inclusion (2SRI) approach to reduce the endogeneity in the effect of student debt on the hazards of becoming underemployed (Terza, Basu, & Rathouz, 2008).[4] Furthermore, I use the two-stage least squares (2SLS) to examine the effect of student debt on the underemployment rate and initial salary. For the instrument, I construct the average amount of grants, excluding a self-amount, received by individuals graduating in a given year at a given institution type. The validity of the instrument holds for the following reasons. First, the average grant amount is strongly correlated to the amount of student loans taken to finance college. Second, by taking an institution-cohort-level average of individuals' grants that do not include a self-amount, I argue that the possible correlation

[2]Kim and Kim (2023) note that rich households are more willing to pay higher tuition for high-quality education, and such expenditure behavior gives colleges an incentive to raise the tuition.

[3]The importance of obtaining a causal impact of student debt on labor market outcomes is also emphasized in Luo and Mongey (2019b).

[4]Terza et al. (2008) shows that the 2SRI yields consistent estimates for non-linear models.

between the grant amount and the individual's unobserved factors is reduced.

The results from the proportional hazards model using the 2SRI approach indicate that holding student debt after college graduation leads to a 22.7% higher probability of exiting the unemployment period to underemployment after graduation, compared to college graduates without student debt. I also find that the change in the level of student debt has a positive, but insignificant, effect on the probability of exiting the unemployment period to underemployment. The results from running the 2SLS regressions show that holding student debt after college graduation, compared to not having student debt, leads to a 39.9% increase on average in the probability of becoming underemployed at the first job and a 54.4% decrease in average in the initial earnings at the first job. I also find that the effect of level-change in student debt is significant and has the same direction as that for holding student debt, but the magnitude is much smaller. For example, a $\$1,000$ increase in student debt held by college graduates leads to a 0.68% increase on average in the probability of becoming underemployed at the first job and a 1.07% decrease on average in the initial earnings at the first job, respectively. In sum, the results from IV regressions support that holding student debt leads individuals to exit an unemployment period faster, be more likely underemployed, and earn less initial labor income at their first job compared to those without student debt after college graduation. In other words, the results highlight that student debt forces college graduates to trade-off between the job search period and the job quality.

The remainder of the paper is organized as follows. Section 2.1 provides the literature review. Section 2.2 explains the data. Section 2.3 explains the hazards models and the results. Section 2.4 explains the IV regressions and the results. Lastly, Section 2.5 concludes the paper with a discussion of further work.

2.1 Literature review

This paper is related to the literature that explores the implications of student debt on the initial labor market outcomes after college graduation. The effect of student debt on labor income is controversial in the literature. For instance, Rothstein and Rouse (2011a) use a sample of undergraduate students from highly selective universities in the U.S. and find that student debt causes graduates to choose high-salary jobs and reduces the probability that they choose low-paid public-interest jobs.[5] This work, however, examines

[5] Field (2009) studies the impact of student debt on career choices through a random experiment data from New York University's Law School, where financial aid packages are randomly given to incoming students. This study finds that those with financial aid packages had a higher rate of getting the first job at public-interest law instead of in private sector than those without financial aid packages.

individuals from high-quality institutions, and the results do not apply to individuals who are not from highly selective universities. My paper studies the effect of student debt on initial labor market outcomes using the national representative dataset and find that student debt leads college graduate more likely to become underemployed with less initial income at their first job after college graduation. Minicozzi (2005) studies the effect of educational debt on job decisions using 1987 National Postsecondary Student Aid Survey (NPSAS) and find that college graduates with student debt are associated with higher initial wage but lower wage growth over the next 4 years. Also, Gervais and Ziebarth (2019) study the causal effect of student debt on labor outcome after graduation using regression kink and find that student debt has a negative effect on earnings, driven by working hours rather than wages. Luo and Mongey (2019b) finds that college graduates with high student debt are likely to accept jobs with higher wages and lower job satisfaction.[6] Ji (2021a), however, finds that college graduates with student debt spend less time on job search and end up with low-pay jobs. While these works focus on unemployment duration and initial salary, my paper also examines the effect of student debt on the underemployment rate of college graduates at the first job.

This paper is also related to the literature that looks at underemployment. Feldman (1996) provides an extensive review of the literature on underemployment. Abel and Deitz (2017) examines underemployment of college graduates across different types of jobs, as well as the factors associated with a greater likelihood of becoming underemployed. They find that underemployed college graduates are more likely to work in higher-paying noncollege jobs than similarly aged workers without a college degree. Also, they find that college graduates who major in more quantitatively oriented and occupation-specific fields have lower underemployment than those with majors that are more general. Lee, Shin, and Lee (2015) use a quantitative model with an educational choice and show that the interaction between college dropout and underemployment after college graduation is important in understanding the increase in wage inequality in the U.S. since 1980. Clark, Joubert, and Maurel (2017) study the overeducation of US workers using the NLSY79 and find that the overeducation is associated with persistently and sizably lower wages. However, none of them takes into account the possible effect of student debt on the underemployment rate after college graduation.

[6] Similar results are found in other studies. For instance, Sieg and Wang (2018) study the impact of student debt on career and marriage market choices of female lawyers using a dynamic model and find that women with more student debt tend to work in private sector jobs.

2.2 Data

This paper utilizes two datasets. First, I use the National Longitudinal Survey of Youth 1997 (NLSY97), which consists of $8,984$ individuals between the ages of 12 and 18 when they were first surveyed in 1997. I focus on the representative individuals who obtained a 4-year college degree by 2007 and exclude the supplemental sample.[7] I use the information from the year 1997 to 2007 to exclude the possible effect of the Great Recession on labor market outcomes. I also exclude individuals who served military or whose occupational information for the first job after college graduation is missing. The final sample size for the empirical analysis consists of $1,041$ individuals. Second, I use the Occupational Information Network (O*NET) database. This dataset contains detailed information on more than 1,000 occupations that cover the U.S. economy. More importantly, it has the information on the required education level for each occupation that is reported in the NLSY97. I merge these two data to have different information in one merged dataset, through which I identify college graduates who take a job that requires less than a 4-year college degree (i.e. underemployment).

The main independent variables of interest for this study are the total amount of student debt accumulated at the time of 4-year college graduation from a 4-year college and the student debt holder dummy which is equal to 1 if a college graduate holds a student debt and 0 otherwise. The main outcome variables of interest are the initial unemployment duration, the underemployment status of the first job after college graduation, and the initial earnings. Other variables used in this study include the ASVAB score, unemployment rate at the year of the first job, and dummy variables including college type, majors, mother's college education, gender, race, regions, and work experience while in college (see Table 2.1 for summary statistics).[8]

2.2.1 Underemployment status

Underemployment has been defined in different ways in the literature. More specifically, Feldman (1996) categorizes the underemployment into following five dimensions:

(1) more formal education than the job requires

(2) involuntarily employed in a field outside the area of formal education

[7]Out of $8,984$ individuals, $2,236$ oversampled individuals are dropped, leaving $6,748$ representative individuals. The oversample individuals are a supplemental sample designed to oversample Hispanic or Latino and Black respondents living in the U.S.

[8]Throughout the paper, dollar amounts are expressed in 2004 U.S. dollars using the IPUMS CPI.

Table 2.1: Summary statistics of final sample

Variable	Mean	(Std. Dev.)	N
Total student debt at graduation(in 2004 dollars)	14,596	(21,106)	1,041
Student debt holder ratio (%)	60.90	(0.4882)	1,041
Underemployment ratio (%)	49.47	(0.5002)	1,041
Unemployment duration (in weeks)	12.82	(16.02)	1,041
ASVAB score (%)	73.43	(20.94)	919
Public school ratio (%)	69.62	(0.4601)	1,027
Major: Science (%)	27.47	(0.4466)	983
Major: Technology (%)	11.39	(0.3179)	983
Major: Engineering (%)	5.80	(0.2338)	983
Major: Math (%)	1.63	(0.1266)	983
Major: Business management (%)	20.24	(0.4020)	983
Major: Others (%)	33.47	(0.4721)	983
Mother has college degree (%)	40.73	(0.4916)	1,041
Male (%)	41.98	(0.4938)	1,041
Non Black and non Hispanic (%)	83.57	(0.3707)	1,041
Region: North East (%)	21.04	(0.4078)	1,041
Region: North Central (%)	27.47	(0.4466)	1,041
Region: South (%)	30.74	(0.4616)	1,041
Region: West (%)	19.98	(0.4000)	1,041

Note: Underemployment ratio corresponds to the underemployment rate for the first job among employed individuals after college graduation. The total student debt at graduation includes individuals who did not borrow student loans. The mean total student debt at graduation, excluding non-student debt holders, is $23,967$.

(3) higher-level work skills and more extensive work experience than the job requires

(4) involuntarily engaged in part-time, temporary, or intermittent employment, and

(5) earning wages 20% or less than in the previous job (for new graduates, wages 20% or less than the average of graduating cohort in same major or occupational track).

In this paper, given that I can use the NLSY97 and O*NET database to identify the educational requirement for the first job that individuals take after college graduation, I use the first definition for underemployment. In other words, I define college graduates as underemployed if they get a job that requires less than a 4-year bachelor's degree.[9]

The NLSY97 does not provide direct measures of the required educational level for the jobs held by individuals, which is critical information for identifying underemployment status at the first job. However, it provides every occupation that an individual holds in a given year at the three-digit Census occupational level. With this information, I merge the NLSY97 data with the educational requirement information from

[9] Here, the educational level less than a 4-year bachelor's degree includes less than a high school diploma, high school diploma, post-secondary certificate, some college courses, and associate's degree (or other 2-year degree).

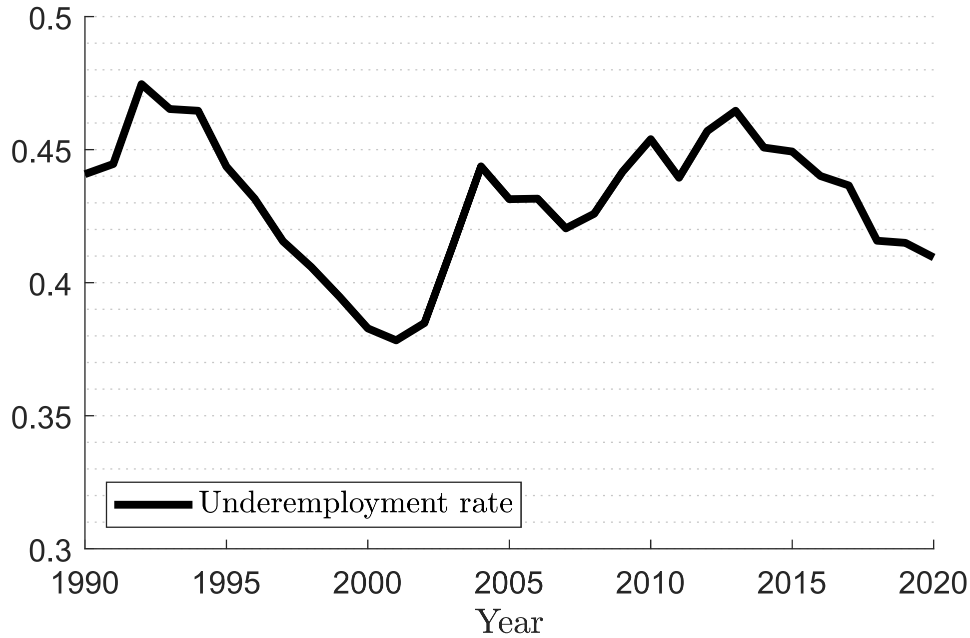


Figure 2.1: Underemployment rate of recent college graduates
Source: the Labor Market for Recent College Graduates by Federal Reserve Bank of New York. Note: The data provides monthly underemployment rates for recent college graduates. The average annual values are calculated by the author.

the O*NET database. This information is collected through the following survey question: "*If someone were being hired to perform this job, indicate the level of education that would be required*". Among the final sample in the NLSY97, 49.51% are underemployed, which is close to the historical underemployment rates for recent college graduates shown in Figure 2.1.[10]

2.2.2 Total student debt at college graduation

The NLSY97 provides the amount of student loan that an individual borrows to finance the education for each academic term of every school attended in a given survey year. I estimate the individual's total student loan amount taken for college education by summing up the student loans borrowed for each term of every college attended until graduation. Table 2.1 shows that average student debt accumulated by the time of graduation from a 4-year college is $14,026$ U.S. dollars.

[10]The definition of underemployment rate in this data is 'the share of graduates working in jobs that typically do not require a college degree'. As seen in Figure 2.1, underemployment rates have been floating between 40% and 50% since 1990.

Table 2.2: Initial labor market outcomes

	Whole sample	Student debt holder	Non holder
Unemployment duration (in weeks)	12.82	11.01	15.65
Underemployment ratio (%)	49.51	49.52	49.50
Initial earnings ($)	18,003	18,182	17,732
Total sample(N)	1,041	634	407
	Whole sample	Underemployed	Non-underemployed
Unemployment duration (in weeks)	12.82	11.41	14.21
Student debt at college graduation($)	14,596	14,652	14,540
Initial earnings ($)	18,003	16,133	19,896
Total sample(N)	1,041	515	526

2.2.3 Initial unemployment period

The initial unemployment period for the first job is the period between 4-year college graduation and the first job. Using the information on individual's weekly employment status, it is calculated by counting the number of weeks spent between college graduation and full-time first job employment. [11] In the sample, some individuals started working full-time before 4-year college graduation. In this case, I assume that their initial unemployment period is zero.[12]

Figure 2.2 describes the distribution of the transition period between student debt holders and non-holders, where each bar shows the fraction of college graduates who spend the corresponding weeks to find their first job after college graduation. It shows that the distribution for student debt holders is generally more right-skewed, suggesting that a larger fraction of student loan borrowers take a shorter period to find their first job after college graduation compared to non-holders.

Table 2.2 shows the descriptive statistics of the variables of interest between student debt holders and non-holders, as well as between underemployed and non-underemployed groups. For instance, student debt holders take 11.01 weeks to find their first job after 4-year college graduation, while non-student debt holders take 15.65 weeks. Also, we see that the underemployed individuals spend less unemployment period until they find their first job and earn fewer initial earnings than employed individuals.

While Table 2.2 shows that the initial unemployment duration is shorter for student debt holders and underemployed workers, we do not see a significant difference in the underemployment rate and initial earnings

[11] Here, I assume that an individual is employed for the full-time first job if the reported work hours in a given week are greater than or equal to 30 hours. For more information about the continuous week scheme, please refer to NLSY97's Appendix 7.

[12] For 4-year college graduation, the NLSY97 provides the month at which an individual graduates. However, it does not provide the week of the graduation. For simplicity, I assume that the individuals in the sample graduated on the first week of the reported month of graduation. For the distribution of years and months of graduation, please see Figure D1.

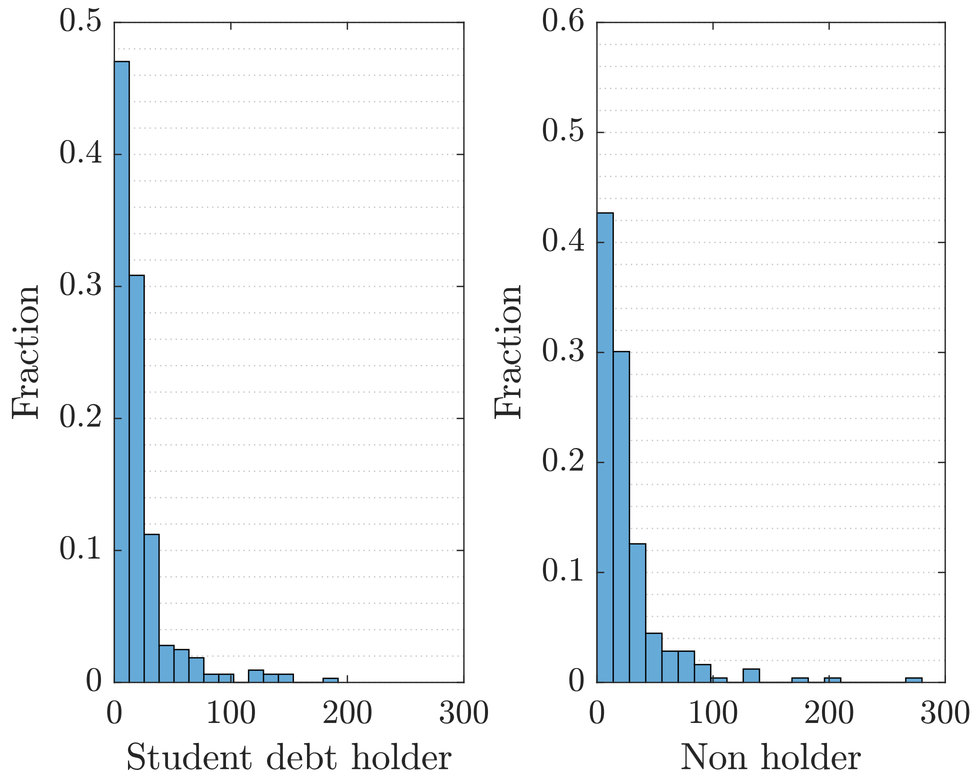


Figure 2.2: Distribution of transition: student debt holder and non holder

Note: Figure 2.2 compares the distribution of unemployment duration between student debt holder (left) and non-student debt holder (right). The x-axis indicates the number of weeks it takes to get the first job since college graduation. Each bar represents the faction of individual spending a given number of weeks of finding the first job after college graduation.

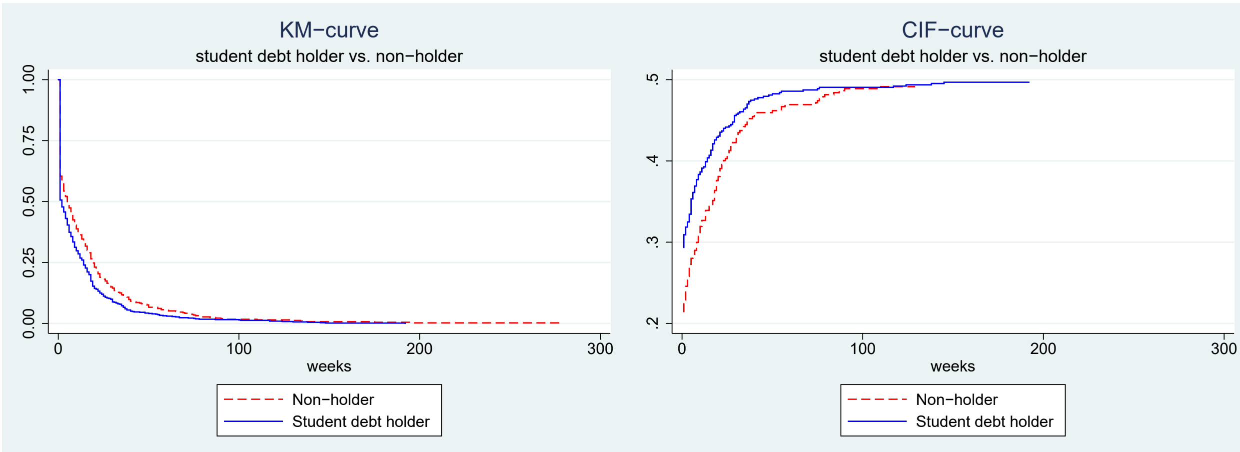
between student debt holders and non-holders, as well as in the average student debt amount between underemployed and non-underemployed groups. One possible reason is that college students graduate at different points in time, and such cohort effects, not reflected in Table 2.2, can lead college graduates to behave differently in terms of employment choices. Also, this could be due to unobserved underlying processes that affect borrowing behaviors and employment decisions but cannot be captured by descriptive statistics. The next sections take account of these issues.

2.3 Survival analysis

I first conduct survival analyses to examine the effect of student debt on initial unemployment duration since college graduation. The survival analysis is used to study the time until the occurrence of an event of interest. The main advantage of using survival analysis for studying the time-to-event is that it takes into account the cohort effects that arise from different points in time at which individuals graduate from college(i.e. Figure D1). Given that the graduation time and the date of first employment for each individual are available from the NLSY97, survival analysis can be used to examine how student debt affects the initial unemployment period.

The survival analysis for this study consists of non-parametric and semi-parametric approaches. In the non-parametric approach, I look at the relationship between holding student debt and the initial unemployment period, without controlling for other covariates. I first calculate Kaplan-Meier (KM) survival estimates, which is the probability of surviving from the event of interest in a given a length of time. Then, to look at how holding student debt is correlated to exiting initial unemployment to underemployment, I calculate cause-specific Cumulative Incidence Function (CIF).

In the semi-parametric approach, I first conduct a Cox-proportional hazard model, which can include other covariates, to further investigate the relationship between student debt and the unemployment period. Here, the unemployment period is interpreted by the conditional probability of exiting unemployment to first employment after college graduation for an individual who is still unemployed. Then, I conduct a cause-specific hazards model to study underemployment. This model can be used when there is more than one event arising. For example, after spending a different number of weeks of unemployment after college graduation (i.e. T_i), an individual can exit the unemployment period and find a job that does not require a 4-year college degree (i.e. underemployment), while he or she also has a possibility of exiting the unemployment period but finding the job that requires a 4-year college degree (i.e. non-underemployment). The event of interest for this analysis is underemployment while the competing event is the non-underemployment, and one event prevents the other event from ever happening. Hence, taking into account the competing event, the cause-specific hazards model can study the effect of student debt after college graduation on the hazard of exiting unemployment to underemployment for the first job.



(a) K-M curve

(b) CIF curve

Figure 2.3: Kaplan-Meier survival estimates and cause-specific CIF estimates

Note: Kaplan-Meier survival estimates (left) and Cumulative Incidence Function estimates (right). KM survival estimate shows the rate at which individuals are still looking for their job at a given week t. The cause-specific CIF estimate is the probability of exiting the unemployment duration to underemployment on or before time t in the presence of a competing event of non-underemployment. The blue line indicates student debt holder, and the red line indicates non-student debt holder. The x-axis indicates the number of weeks spent to find the first job after college graduation.

2.3.1 Non-parametric estimation: survival functions

I first look at the association between holding student debt and unemployment duration for the first job using a nonparametric approach. Figure 2.3 (a) shows Kaplan-Meier survival estimates for student debt holders and student debt non-holders. The estimates are calculated as follows:

$$S(t+1) = S(t)\frac{N(t+1) - D(t+1)}{N(t+1)},$$

where $N(t)$ is the number of individuals who are still unemployed at week t and $D(t)$ is the number of individuals who are employed at week t. In other words, the survival function is the ratio of individuals who are still unemployed and are looking for the first job at each t. As shown in Figure 2.3 (a), the survival estimates are lower for student debt holders after college graduation than those for college graduates without student debt. This implies that, after college graduation, individuals who hold student debt are more likely to get their first job in a given week t compared to non-student debt holders.

The Kaplan-Meier estimate assumes that competing risks are absent. In other words, Kaplan-Meier functions cannot distinguish individuals who face different types of employment outcomes. In the context of this study, college graduates exit unemployment to either underemployment or non-underemployment. To study underemployment while taking into account the competing event of non-underemployment, I calculate the cause-specific cumulative incidence function (CIF):

$$CIF_k(t) = P(T \leq t, D = k) = \int_0^t S_k(u) h_k^{cs}(u) du$$

where $k = \{l, h\}$ indicates the employment types (l = underemployment and h = non-underemployment), $h_k^{cs}(t)$ is a cause-specific hazard function, and $S_k(u)$ is survival function:

$$h_k^{cs}(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, D = k | T \geq t)}{\Delta t}$$

$$S_k(t) = \exp\left(-\int_0^t h_k^{cs}(u) du\right).$$

For example, $CIF_l(t)$ tells us the probability of exiting the unemployment to underemployment l on or before time t, given that the individual is still looking for the first job. In this setting, only one event can occur, and two events are mutually exclusive. Figure 2.3 (b) shows the cause-specific CIF estimates between student debt holders and student debt non-holders, with underemployment as the event of interest, and the estimates are higher for the student debt holders (i.e. blue line). This indicates that, after college graduation, the student debt holders have a higher probability of becoming underemployed at their first job at a given t in the presence of the competing risk of becoming non-underemployed.

2.3.2 Semi-parametric estimation

The non-parametric estimations above do not control for other covariates that can also affect the initial unemployment period. To control for other explanatory variables in the analysis, in this section, I estimate the Cox proportional hazards model. This studies the effect of student debt on the hazard of exiting the unemployment period to employment after college graduation.[13] Then, I estimate the cause-specific hazards model to further examine whether student debt lead to underemployment after college graduation, in the presence of competing event of non-underemployment.

[13]Please note that I am using the terminology *hazard* instead of *probability*(i.e. incidence of event occurring). According to Austin and Fine (2017), there is a one-to-one relationship in terms of the direction of the effect between hazard ratio and probability, but the magnitude of the relative effect of the covariate on the hazard ratio(or cause-specific hazard ratio) is only approximately correct in terms of the probability.

Cox proportional hazards model

The hazard function measures the probability of experiencing the event, given that an individual has survived up to a specific time. In the context of this study, the hazard function,

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t},$$

can be interpreted as the instantaneous rate of exiting the unemployment period to employment within the interval $[t, t + \Delta t]$ for an individual who is still unemployed and looking for their first job in the week t after college graduation. Here, T is the number of unemployed weeks spent by an individual finding their first job after college graduation.

To investigate the relationship between student debt and the hazard of becoming employed after college graduation, I conduct the Cox proportional hazards regression (Cox, 1972). More specifically, it can be estimated as follow:

$$h(t, x) = h_0(t) \exp(\boldsymbol{x'\beta}),$$

where t represents the survival time(i.e. initial unemployment period), $h_0(t)$ is the baseline hazard, and coefficients $\boldsymbol{\beta}$ measure the effect of covariates $\boldsymbol{x}$.[14] The above equation can be re-written as a multiple linear model:

$$\ln\left(\frac{h(t, x)}{h_0(t)}\right) = \boldsymbol{x'\beta}.$$

Here, $\frac{h(t, x)}{h_0(t)}$ is the hazard ratio, and $\boldsymbol{x'\beta}$ become the linear combination. The coefficients $\boldsymbol{\beta}$ are estimated by Cox proportional hazards regression. The proportionality assumption of the Cox model allows us to assume no specific shape of baseline hazards $h_0(t)$ and to generate constant estimates over time:

$$\exp(\beta(x^* - x)) = \frac{\frac{h(t, x^*)}{h_0(t)}}{\frac{h(t, x)}{h_0(t)}} = \frac{h(t, x^*)}{h(t, x)}.$$

For interpretation, if the covariate x is the log-level of student debt, $x^* - x = \log(1.1)$ the event of interest is employment, then $\exp(\beta(x^* - x))$ tells us the hazard ratio for an individual to exit underemployment to employment with an increase in 10% student debt level, holding other covariates constant. If x is the

[14]The baseline hazard function is the hazard of exiting unemployment to employment for an individual with covariate values 0.

Table 2.3: Cox proportional hazards model regressions

	(1)	(2)	(3)	(4)
log(Student debt)	1.023**	1.024**		
Student debt holder			1.254**	1.267**
Unemployment rate	1.198**	1.200**	1.200**	1.202**
ASVAB Score	0.998	0.997	0.998	0.997
School type(D = 1) if public)	1.055	1.06	1.053	1.056
Major: Technology	1.340*	1.347*	1.337*	1.342*
Major: Engineering	1.318+	1.300+	1.318+	1.300+
Major: Math	1.243	1.238	1.239	1.234
Major: Business management	1.216+	1.204+	1.212+	1.200+
Major: Others	1.017	1.019	1.02	1.022
Mother has a college degree		1.002		1.001
Male		1.055		1.054
Non Black and non Hispanic		1.107		1.102
Region: North central		1.001		0.998
Region: South		1.013		1.009
Region: West		0.944		0.944
χ^2	28.45	30.66	28.55	30.63
Proportional Hazards test	0.72	0.16	0.72	0.16

Note: The estimates are exponentiated coefficients (i.e. Hazards ratios). The base dummy for Major is Science. The base dummy for Region is North East.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

binary that equals 1 if college graduates hold student debt and 0 otherwise, then $\exp(\beta)$ becomes the hazard ratio of college graduates with student debt relative to non-holders. A hazard ratio greater than 1 (i.e. $\beta > 0$) indicates that the covariate is positively correlated with the hazard of experiencing the event (i.e. employment) and thus negatively associated with the survival length (i.e. initial unemployment duration).

Table 2.3 shows the results from conducting Cox proportional hazards regressions. The estimates reported in this table are the hazard ratios $\exp(\beta)$, and the event of interest is the employment at the first job. Columns (1) and (2) show the results where the variable of interest is the log of student debt at graduation, and columns (3) and (4) show the results where the variable of interest is the student debt dummy, which is equal to 1 if an individual holds student debt at college graduation and 0 otherwise. Columns (1) and (3) control for the unemployment rate at the year of the first job, ability, and educational variables such as school type and majors, and columns (2) and (4) additionally control for gender, race, and region.

Column (2) shows that the hazard ratio for the log of student debt is 1.024, indicating that a 10 $\%$ increase in student debt holding everything else constant is associated with a 0.23% increase in the hazard of getting the first job after 4-year college graduation.[15] The hazard ratios for a technology major (and an engineering major) are 1.347 (and 1.300), implying that having a technology major (or an engineering major) is associated with a 35 percent (30 percent) higher hazard of getting the first job at a given t, compared to

[15] A 10% increase in student debt is same as $\ln(1.1) = 0.095$ increase in the student debt. Hence, the hazard ratio is $\exp(0.095 * 0.024) = 1.0023$.

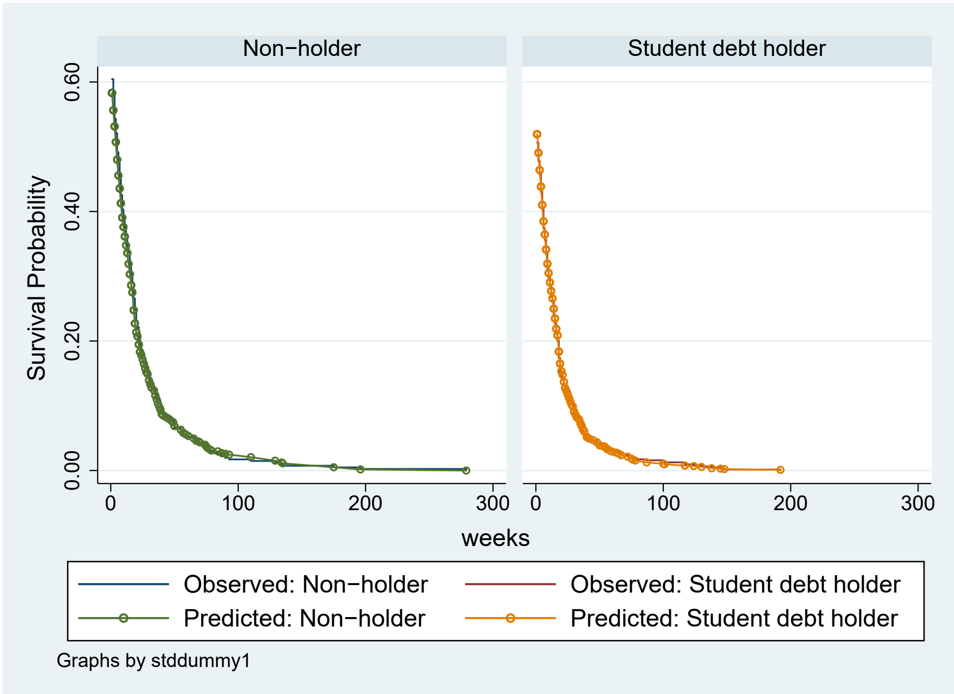


Figure 2.4: Observed v.s. Predicted survival estimates

Note: Figure 2.4 compares Kaplan-Meier observed survival estimates with the Cox-predicted estimates for the student debt dummy variable. The closer the observed values are to the predicted, the less likely it is that the proportional-hazards assumption has been violated.

having a science major. The estimate of the student debt holder dummy in column (4) is 1.267, indicating that holding student debt is associated with 27% higher hazard of getting the first job at a given t, compared to having no student debt. The estimates for other covariates in column (4) are close to those in column (2) in terms of the magnitude of effects.

I also conduct the test for the proportional-hazards assumption, a fundamental assumption for the Cox models and obtain the test statistics of 0.72 and 0.16 for columns (1) and (2). The values of test statistics are small, failing to reject the null hypothesis that the hazards are proportional. The proportional assumption is also tested by comparing the actual KM survival estimates with the predicted survival estimates from the Cox proportional hazards model. As in Figure 2.4, the observed and predicted values are close together, thereby further supporting that the proportional assumption is not violated.

In short, the results suggest that having more student debt leads to a higher hazard of exiting the unemployment period to employment after college graduation. Also, I find that the effect is smaller when we look at the change in the level of student debt than when we compare the student debt holders and non-student debt holders. In other words, having student debt has a larger effect on the hazard of becoming employed at the first job compared to those without student debt, while the effect is smaller when we look at the change in the level of student debt held by college graduates.

Cause-specific hazards model

To further examine the relationship between student debt and underemployment while taking into account the competing event of non-underemployment, I use a cause-specific hazards model. Let $k = \{l, h\}$ refer to the employment type where l is underemployment and h is non-underemployment. Then, the cause-specific hazard function, $h_k^{cs}(t)$ is

$$h_k^{cs}(t) = \lim_{\Delta \to 0} \frac{P(t < T \leq t + \Delta t, D = k | T \geq t)}{\Delta t},$$

which is an instantaneous rate of exiting the unemployment period at week t to underemployment among those who are still unemployed and looking for their first job after college graduation. The corresponding regression model is

$$h_k^{cs}(t, x) = h_{0,k}^{cs}(t) \exp(\boldsymbol{x'\beta})$$

and the coefficient measure is the cause-specific hazard ratio,

$$\exp(\beta(x^* - x)) = \frac{h_k^{cs}(t, x^*)}{h_k^{cs}(t, x)}.$$

As mentioned above, the difference between the Cox proportional hazards model and the cause-specific hazards model is that the latter examines the effect of student debt of college graduates on the hazard of exiting unemployment to underemployment, holding other covariates constant.

Table 2.4 shows the results from cause-specific hazard regressions. As in Table 2.3, the estimates are the hazard ratios. The event of interest is underemployment after 4-year college graduation, and the competing event is non-underemployment. Columns (1) and (2) show the results where the variable of interest is the log of student debt at graduation, and columns (3) and (4) show the results where the variable of interest is the student debt dummy, which is equal to 1 if an individual holds student debt after college graduation. Columns (1) and (3) control for the unemployment rate at the year of the first job, ability, and educational

Table 2.4: Cause-specific hazards model regressions

	(1)	(2)	(3)	(4)
log(Student debt)	1.015	1.017+		
Student debt holder			1.168	1.188+
Unemployment rate	1.395***	1.374***	1.397***	1.376***
ASVAB Score	0.996+	0.994*	0.996+	0.994*
School type (D = 1 if public)	1.029	1.048	1.027	1.046
Major: Technology	1.008	1.016	1.007	1.014
Major: Engineering	0.902	0.879	0.902	0.879
Major: Math	0.792	0.773	0.792	0.772
Major: Business management	1.311+	1.281+	1.309+	1.278+
Major: Others	1.192	1.193	1.194	1.195
Mother has a college degree	1.007	1.007	1.007	1.007
Male		1.104		1.104
Non Black and non Hispanic		1.149		1.147
Region: North central		1.001		0.999
Region: South		0.807		0.805
Region: West		0.838		0.837
chi2	24.18	31.00	24.34	31.18
Proportional Hazards test	0.66	0.28	0.68	0.28

Note: The estimates are exponentiated coefficients (i.e. Hazards ratios). The base dummy for Major is Science. The base dummy for Region is North East.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

variables such as school type and majors, and columns (2) and (4) additionally control for gender and regions.

Column (2) shows that the cause-specific hazard ratio for the log of student loan covariate is 1.017, indicating that a 10% increase in student debt holding everything else constant is associated with a 0.16% increase in the hazard of exiting unemployment to underemployment at the first job after 4-year college graduation. The result, however, is less significant, only at 10%. The proportional-hazards tests in columns (1) and (2) give the test statistics of 0.66 and 0.28, respectively, failing to reject the null hypothesis that hazards are proportional. Column (4) shows that the coefficient for the student debt dummy is 1.188, indicating that holding student debt after college graduation is associated with a 18.8% higher hazard of exiting unemployment to underemployment at the first job, compared to student debt non-holders. As in the Cox proportional hazards models, the magnitude of the effect of student debt is larger when we compare the student debt holder to student debt non-holders than when we look at the change in the level of student debt.

2.4 IV regressions

The above results from survival analysis can still face the endogeneity issue. Suppose we have the following proportional hazards model:

$$h(t, x) = h_0(t) \exp(\boldsymbol{x'\beta} + u)$$

where $\boldsymbol{x}$ is a vector of observed variables, and u includes unobserved variables such as ability. Because unobserved variables cannot be included in the model, the endogeneity would be present if $\boldsymbol{x}$ is correlated to u. For instance, the amount of student loans a college student takes may be correlated with his or her unobserved ability, which may affect initial labor market decisions. Consider college students whose inherent ability is high. They may expect a higher lifetime income and thus take larger student loans to smooth their consumption while in college. These individuals are less likely to be underemployed and are more likely to earn a higher income. Such a positive correlation could bias the effect of student debt on labor market outcomes upward. On the other hand, students with low ability and less financial resources for college education would borrow student loans. Given that they are likely to earn low labor income in the future and that they need to repay outstanding student debt, they may choose to take a job as soon as possible, even if they become underemployed, to keep up with the repayment.[16]

To adjust for the endogeneity, I conduct IV regressions using the average amount of grants received to finance college costs at the year of college graduation by different institution types, excluding the self-amount when calculating the average.[17] Hence, it is calculated as below

$$IV_{isy} = \frac{\sum_{j=1}^{N_{sy}} (g_{jsy} | j \neq i)}{(N_{sy} - 1)}$$

where s is the type of institution, g_{isy} is the amount of grant received by individual i at the year of college graduation y, and N_{sy} is the number of individuals who belong to the institutional type s at y. For institution type, it is categorized into public, private non-profit, and private for-profit.[18]

[16] The endogeneity of student debt is also discussed in Luo and Mongey (2019b). In Table 2.7, I also conduct the endogeneity test (i.e. Durbin-Wu-Hausman) and obtain the test statistics that are large enough to reject the null hypothesis that the student debt is exogenous.

[17] I look at the amount of grant received at the year of college graduation because the institutional type is defined based on the last year of college that individuals attended. Also, this approach gives more exogenous variation, presuming that the graduation year is not fully deterministic and differs across individuals.

[18] The publicly available NLSY97 dataset only provides institutional types and does not provide information on actual institutions

I argue that this IV is valid for the following reasons. First, the average amount of grants by institutional type is correlated with the accumulated amount of student loans taken to finance the education, as the maximum student loan amount that a college student can take is limited to the net college cost. Table 2.5 shows the results of the first stage regression where the student loan is regressed on the average amount of grants and other control variables. As you can see in Column (1), the amount of student loan is significantly correlated with the amount of grants received to finance college costs at the year of college graduation. Column (2) shows the first regression where I use the student debt holder dummy that equals 1 if an individual holds student debt after college graduation. As in Column (1), holding student debt at graduation is strongly correlated with the instrument. The test statistics from the weak identification tests (Stock and Yogo 2005) in Table 2.7 also support that the instruments are not weak. Second, to be a valid instrument, it should only affect the employment outcomes indirectly through the student debt and not by other possible channels, also known as the exclusion restriction. This condition cannot be directly proved and has to be argued. By constructing the institution-average amount of grant that does not include the self-amount, the instrument is less correlated with each individual's unobserved characteristics.

For the proportional hazards model which examines the unemployment duration (i.e. probability of exiting the unemployment period at a given point of time after college graduation), the 2SLS estimators would be inconsistent because of the non-linearity of the model.[19] In this study, I estimate the two-stage residual inclusion (2SRI) estimator, an alternative method for 2SLS proposed by Terza et al. (2008) who show that the 2SRI estimators are consistent for non-linear models. In the context of the proportional hazards model, the 2SRI approach is straightforward. In the first stage of 2SRI, instead of getting the fitted values of endogenous variables, the fitted residuals are obtained from estimating the first stage through a nonlinear least square. The fitted residuals serve as control functions that captures possible unobserved factors that could affect both student debt and initial labor makret outcomes. For a student debt holder dummy, the logit model is used for the first stage. Then, in the second stage of 2SRI, I regress the outcome variable on the endogenous variable, exogenous variables, and fitted residuals. By including the fitted residuals in the second stage, this approach makes the student debt appropriately exogenous. To look at the effect of student debt after college graduation on the underemployment rate and initial earnings, I use the two-stage least squares (2SLS) for estimation.

---

that individuals attended. However, I still obtain a significant correlation between the IV variable and the accumulated student debt held at college graduation.

[19]In linear model settings, both 2SLS and 2SRI yield the consistent estimates.

Table 2.5: First stage regression

	(1) Student debt ($1,000)	(2) Student debt holder
IV_grant ($1,000)	0.134***	0.002***
Unemployment rate	0.351	-0.081***
Work experience in college	1.256**	0.089***
ASVAB Score	-0.0287**	0.0001
Major: Technology	-1.855*	-0.090***
Major: Engineering	0.145	0.0146
Major: Math	-7.233***	-0.215***
Major: Business management	-3.576***	-0.137***
Major: Others	-1.438***	-0.0942***
Mother has a college degree	-5.501***	-0.158***
Male	-1.663***	-0.0667***
Non Black and non Hispanic	-5.372***	-0.125***
Region: North central	-3.375***	0.0607***
Region: South	-5.014***	-0.0165
Region: West	-2.686**	-0.0106
Constant	23.52***	1.177***
F	47.25	56.43

Note: The dummy for work experience in college indicates whether an individual worked at the last year of college education. The omitted dummy for Major is Science. The omitted dummy for Region is North East.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.6 shows the results from running 2SRI regression for the proportional hazards models. The event of interest is underemployment after 4-year college graduation, express by the unemployment period before the first job, and the competing event is non-underemployment. Columns (1) and (2) shows the results where the variable of interest is the log of student debt at graduation, and columns (3) and (4) show the results where the variable of interest is the student debt dummy, which is equal to 1 if an individual holds student debt after college graduation. Columns (1) and (3) control for the unemployment rate at the year of the first job, ability, and majors, and columns (2) and (4) additionally control for gender race, and region. Column (4) shows that the coefficient for the student debt holder dummy is 1.227, indicating that holding student debt after college graduation leads to a 22.7% higher probability of exiting the unemployment period to underemployment after college graduation, compared to college graduates without student debt. However, the result is less significant, at 10% level. The statistics for the proportional hazards test show that the proportional hazards assumption is not violated.

Table 2.7 shows the results from running IV regressions on the two outcome variables. The dependent variable for columns (1), (2), (5), and (6) is the underemployment rate of individuals at the first job after college graduation, and the dependent variable for columns (3), (4), (7), and (8) is the initial earnings at the first job. Columns (1), (3), (5), and (7) control for the unemployment rate at the year of the first job, work experience while enrolled in college, ASVAB score, and majors. Columns (2), (4), (6), and (8) additionally

Table 2.6: 2SRI regressions summary

	(1)	(2)	(3)	(4)
log(Student debt)	1.114	1.079		
Student debt holder			1.228	1.227+
Unemployment rate	1.384***	1.364***	1.399***	1.375***
ASVAB Score	0.996+	0.995*	0.997+	0.995*
Major: Technology	0.991	1.015	0.992	1.013
Major: Engineering	0.87	0.855	0.875	0.859
Major: Math	0.775	0.755	0.796	0.769
Major: Business management	1.228	1.215	1.234+	1.22
Major: Others	1.18	1.19	1.186	1.195
Mother has a college degree	0.847+	0.855	0.842+	0.851
Male		1.096		1.094
Non Black and non Hispanic		1.172		1.162
Region: North central		1.057		1.058
Region: South		0.83		0.826
Region: West		0.888		0.885
Fitted residual (log(Student debt))	0.909	0.94		
Fitted residual (student debt holder)			0.965	0.972
chi2	37.18	48.27	36.8	48.4
Proportional Hazards test	0.98	0.92	0.98	0.94

Note: The estimates are exponentiated coefficients (i.e. Hazards ratios). The base dummy for Major is Science. The base dummy for Region is North East.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7: IV regression summary

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Student debt($\$1,000$)	0.00639*	0.00679*	-0.00745	-0.0107+				
Student debt holder					0.426*	0.399*	-0.44	-0.544+
Unemployment rate	0.0653***	0.0654***	-0.0881***	-0.0925***	0.0995***	0.101***	-0.117***	-0.135***
Work experience in college	0.141***	0.138***	0.0320+	0.0225	0.108***	0.112***	0.0523*	0.0450*
ASVAB Score	-0.000668*	-0.00136***	-0.00352***	-0.00299***	-0.000620+	-0.00160***	-0.00350***	-0.00266***
Major: Technology	-0.0952***	-0.0904***	-0.0696+	-0.0775+	-0.0599*	-0.0674**	-0.108*	-0.106*
Major: Engineering	-0.146***	-0.158***	0.228***	0.206***	-0.145***	-0.163***	0.214***	0.202***
Major: Math	-0.103*	-0.116**	-0.00968	-0.0152	-0.0544	-0.0808	-0.0555	-0.0505
Major: Business	0.0603***	0.0553**	0.0993**	0.0677+	0.0977***	0.0857**	0.0771+	0.0498
Major: Others	0.105***	0.111***	-0.0873***	-0.107***	0.135***	0.137***	-0.119***	-0.142***
Mother has a college degree		0.0339+		-0.161***		0.0606*		-0.193***
Male		0.0419***		0.0105		0.0581***		-0.0103
Non Black and non Hispanic		0.0678**		0.0118		0.0799**		0.00319
Region: North central		0.0353+		-0.125**		-0.0118		-0.0628*
Region: South		-0.0638**		-0.0874*		-0.0912***		-0.0586+
Region: West		-0.0189		-0.0717+		-0.0356+		-0.0653+
Constant	0.0318	0.0014	10.41***	10.59***	-0.322+	-0.315	10.73***	10.97***
F	64.91	47.35	20.64	14.43	60.32	46.15	19.66	13.51
Weak identification test	79.124	67.985	61.499	49.281	36.706	40.872	35.765	38.784
Endogeneity test	7.265	6.936	3.953	5.435	7.476	7.202	3.491	5.157

Note: The dummy for work experience in college indicates whether an individual worked at the last year of college education. The omitted dummy for Major is Science. The omitted dummy for Region is North East.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

control for gender, mother's education, race, and region.

Column (2) shows that the coefficient for the student loan amount is 0.00679, indicating that a $\$1,000$ increase in student debt held by college graduates leads to 0.68% increase in the probability of becoming underemployed. Also, in column (4), the coefficient for the log of the student loan amount is -0.107, indicating that a $\$1,000$ increase in student debt after graduation will lead to a 1.07% decrease in the initial earnings at the first job. The negative effect of student debt on initial earnings contradicts to the result from

Luo and Mongey (2019b) which shows that college graduates with high student debt are likely to accept jobs with higher wages. Overall, the results from IV regressions imply that holding more student debt after college graduation leads individuals to become more likely underemployed and earn less initial earnings at the first job compared to those with less student debt after college graduation. Similar results are found when I change the variable of interest to the student debt holder dummy which equals 1 if college graduates hold student debt after college graduation and equals 0 otherwise. However, the magnitude of the effect is larger when we compare the student debt holders and non-student debt holders than when we compare the level of student debt.

2.5 Conclusion and Discussion of Further Work

This paper studies the relationship between student debt and initial labor market outcomes of college graduates in the United States using the NLSY97 and the O*NET data. Specifically, I show that a higher amount of student debt after college graduation leads individuals to have a shorter initial unemployment duration, a higher probability of becoming underemployed, and lower initial earnings. I merge the NLSY97 and the O*NET data and identify the individual's underemployment status at the first job after college graduation. Then, I use the survival analysis, both nonparametric and semi-parametric approaches, and IV regressions to investigate the causal relationship between student debt and initial labor market outcomes.

In sum, this study reveals that student debt has a negative effect on college graduates' initial labor market outcomes. More specifically, student debt forces college graduates to trade-off between the job search period and the job quality. While this paper studies the short-term effect of student debt on initial labor market outcomes, for future research, it would be insightful to examine the long-term persistent effects of student debt on college graduate' labor market outcomes during their work life. For instance, some college graduates could choose to be underemployed for a short period of time while at the same time looking for a better job. For them, underemployment would be a temporary stage through which they transit into non-underemployment for the next job. These individuals would have higher lifetime earnings and less financial burden in terms of student debt repayment. On the other hand, the duration of underemployment could be more persistent among student debt holders compared to non-student debt holders. Identifying the effect of student debt on the persistence of the underemployment of college graduates could give policymakers the insights into student loan policies that could mitigate the financial burden that student debt borrowers face.

Appendix A

Appendix for Chapter 1

A.1 National Longitudinal Survey of Youth (NLSY79 and NLSY97)

The NLSY79 consists of 12,686 individuals between the ages of 14 and 22 in 1979, and the NLSY97 consists of 8,984 individuals between the ages of 12 and 18 in 1997. We only use the nationally representative sample and exclude the supplemental sample.[1] We drop individuals with 1) no ability information, 2) no family income information for age 16 or 17 in the NLSY79, 3) no parental transfer information in the NLSY97, and 4) the highest degree achieved higher than the bachelor's degree.[2] We use the data waves from 1979 to 1988 for the NLSY79 and from 1997 to 2008 for the NLSY97. These data waves are chosen such that the youngest cohort in each survey becomes the age 23 by the last year in the data. The final sample size is 2,159 for the NLSY79 and 4,838 for the NLSY97.

We measure the ability using the Armed Forces Qualification Test (AFQT) score for the NLSY79 and the Armed Services Vocational Aptitude Battery (ASVAB) test score for the NLSY97. The AFQT is a test that consists of arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operations, and it is a part of the ASVAB test. Family income includes military income; wages, salary, and tips; net business income; net farm income; unemployment compensation; child support; AFDC payments; food stamps; other welfare and social security income; education benefits and grants; inheritance; other income (interest, dividends, rent); income from parents and other household members; and rental subsidy. An

[1] Supplemental samples are designed to oversample Hispanic or Latino and black respondents living in the United States. In the NLSY79, 6,575 individuals are supplemental samples, while, in the NLSY97, 2,236 individuals are supplemental samples.

[2] For family information in the NLSY79, we use the family income for age 17 if the information is available for both age 16 and 17.

individual is measured as college educated if he or she completed at least 16 years of education.

Following Abbott et al. (2019), we measure the parental transfer as all transfers, including allowances, that an individual receives from parents or guardians when he or she is between 16 and 22 years old. The Income section of the NLSY97 reports the three sources of parental transfers: from both parents (or guardians), from a living mother figure (or female guardian), and from a living father figure (or male guardian). We measure the parental transfers for respondents who live with both parents using transfers from both parents. If respondents do not live with both parents (or guardians), then we sum the amounts from both a living mother figure and a living father figure. If respondents live with any parent (or guardian), we include the average amount of imputed rent by age groups. We compare the estimated parental transfers with the amount of family aid, adjusted for rent payments, in the Education section of the NLSY97 and use the maximum of the two.[3]

The Education section of the NLSY97 provides the amount of student loans that an individual borrows for every school, academic term attended, and year.[4] Given that the student loan amounts include both federal and private loans, we top-code the annual amount that exceeds $35,000$ dollars. Then, we calculate the annualized average student debt by summing up the borrowed amount across terms, schools, and years of education and dividing it by the total number of full-time equivalent years in college. To calculate the cumulative student debt until graduation, we multiply the yearly average of student loans by 4.

We estimate the college education cost following Gordon and Hedlund (2020). The NLSY97 provides eight different sources of financing education for every school, academic term attended, and year:

(1) Financial aid from family and friends ($YSCH\_24600$)

(2) Loans from family/friends ($YSCH\_24700$)

(3) Grants and scholarship ($YSCH\_25400$)

(4) Federal subsidized and other student loan ($YSCH\_25600$)

(5) Work study financial aid ($YSCH\_26000$)

(6) Employer assistance financial aid ($YSCH\_26200$)

[3]As noted in Abbott et al. (2019), the family aid in the Education section is not fully consistent with the parental transfers in Income section, has many skips, and does not cover all transfers.

[4]If it is reported as a bracket, we use the median to approximate the actual amount.

(7) Other financial aid ($YSCH\_26200$)

(8) Out of pocket spending ($YSCH\_26500$)

With these measures at a given year and school, we first identify whether an individual reports any change in how he or she finances the education from the last term. If there is no change, we carry over the amount from the previous term for all eight financing sources. If an individual reports the change from the previous term but does not report a specific amount for any financing source for the current term, we update that amount to zero. Second, we assign the weight of 1 to the reported amount if an individual was a full-time student and 0.5 if an individual was a part-time student at the given term. Finally, for each school and year, we compute the average amount of each financing source across terms and schools. This gives us the individual-specific average amount per term across schools for each financing source. To get an annual value, we multiply this amount by the number of full-time equivalent terms that an individual attended schools in a given year. Then, we sum up the annual average values of eight financing sources and define it as the annual average sticker price that an individual paid for college education. The net college cost is obtained by subtracting the average amount of grant and scholarship from the average sticker price.

Lastly, in the NLSY97, we calculate the weekly average hours worked in college as follows. First, we calculate individuals' total hours of work by summing the hours of work while enrolled in 4-year or 2-year college. Then, we divide this by the number of years of college enrollment. To express it as weekly hours, we divide the annual average by 52.

A.2 Minimum Distance Estimation

We estimate the year-varying variances of shock $\{\sigma^2_{pt}, \sigma^2_{vt}\}$, the persistence of the shock $\{\rho\}$, and the variance of the initial value for the persistent shock using minimum distance methods.[5] The theoretical moment is defined as:

$$m^j_{t,t+n}(\mathcal{P}_{102\times 1}) = E(r_{i,j,t}r_{i,j+n,t+n}),$$

[5] Given that the PSID has conducted a biennial survey starting from 1997, the estimation of annual shock processes must confront the problem of observations missing for every other year. As Heathcote et al. (2010) indicate, although the variance for the persistent shock for the missing years can theoretically be found using the available information from adjacent years, the resulting estimates are downward-biased because of insufficient information. Therefore, I follow their approach and estimate the variances for missing years by taking the weighted average of the two closest surrounding years.

which is the covariance between the wages of individuals at age j in year t and $t+n$. To calculate empirical moments, I group individuals into 50 years and 26 overlapping age groups. For example, the first age group contains all observations between 25 and 34 years old, and the second group contains those between 26 and 35 years old. The empirical moment conditions are

$$\hat{m}^{j}_{t,t+n} - m^{j}_{t,t+n}(\mathcal{P}) = 0,$$

where $\hat{m}^{j}_{t,t+n} = \frac{1}{I_{j,t,n}} \sum_{i=1}^{I_{j,t,n}} \widehat{r}_{i,j,t}\widehat{r}_{i,j+n,t+n}$ and $I_{j,t,n}$ is the number of observations of age j at year t existing n periods later.

The minimum distance estimator solves

$$\min_{\mathcal{P}} [\hat{\mathbf{m}} - \mathbf{m}(\mathcal{P})]'[\hat{\mathbf{m}} - \mathbf{m}(\mathcal{P})],$$

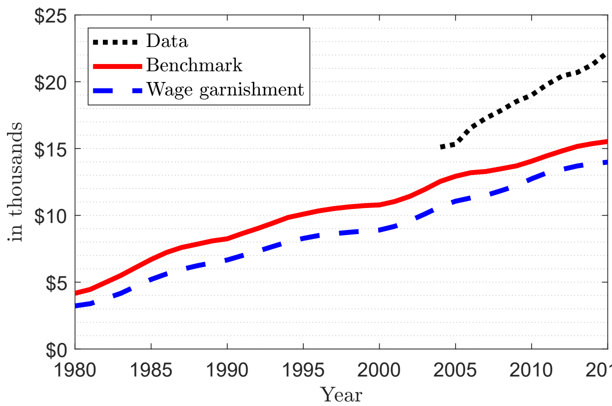
where the vectors $\hat{\mathbf{m}}$ and $\mathbf{m}$ represent empirical and theoretical moments of dimension $10,070 \times 1$. The identity matrix is used as the weighting matrix.

A.3 A Model with Wage Garnishment

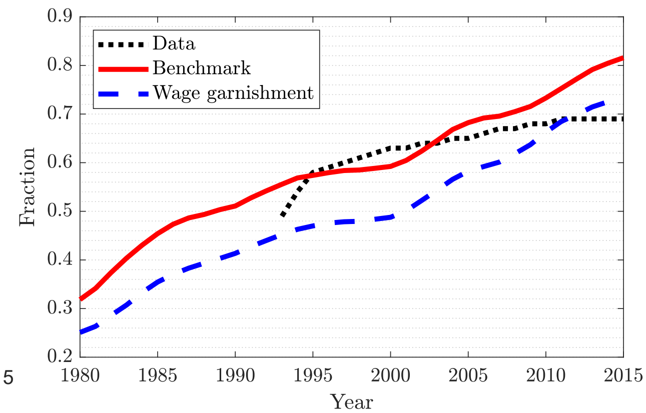
In the benchmark economy, delinquency is considered as missing payments for 90 days or more. Thus, delinquency only involves a psychic cost. However, given that the model period is one year, we solve a model with wage garnishment when defaulting in this section to be consistent with the penalty of missing payments for 270 days or more. Following F. Ionescu (2009), we assume that 3% of the labor income is garnished when borrowers are delinquent on their payments. Though we do not re-calibrate the model, as seen in FigureC1 (c), introducing wage garnishment brings the delinquency rate in the model close to the observed 270 days or more delinquency rate.

FigureC1 shows that the aggregate student debt increases \$110 billion less than the benchmark economy. With an extra penalty on delinquency, college students rely less on government loans to finance college. Thus, on average, fewer students borrow from government loans, and they borrow less amount than the benchmark economy. Also, fewer student loans are rolled over time, reducing the increase in total student debt in the economy.

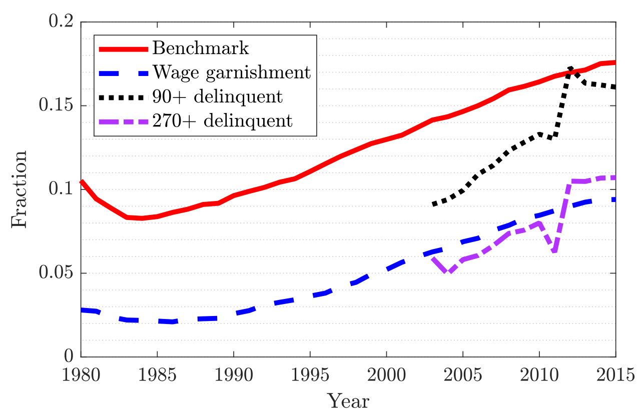
However, as seen in Table C1, when we do not allow delinquency, the aggregate student debt only



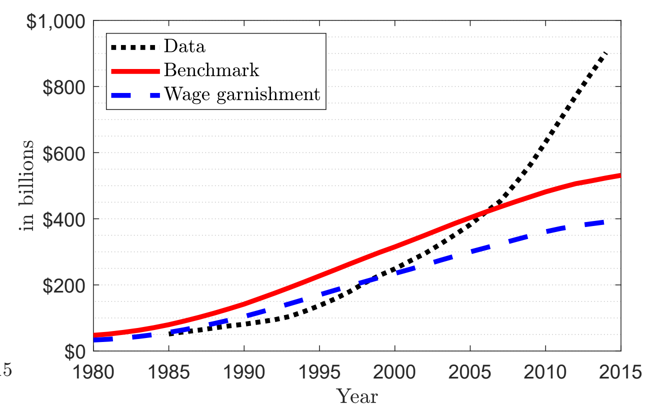
(a) Cumulative student debt at graduation



(b) Fraction of graduating seniors with student debt



(c) Fraction of delinquent borrowers



(d) Aggregate student debt

Figure C1: Benchmark v.s. Economy with Wage Garnishment

Note: This figure shows (a) cumulative student debt at graduation, (b) fraction of graduating seniors with student debt, (c) fraction of delinquent borrowers, and (d) aggregate student debt under the benchmark economy and the economy with wage garnishment. The red solid line is the benchmark economy. The blue dashed line is the economy with wage garnishment. The black dashed line and purple dash-dot line in (c) represent 90+ delinquent and 270+ delinquent rates from the FRBNY data, respectively.

increases by $\$210$ billion, compared to $\$364$ billion in the model with delinquency that accompanies wage garnishment. This emphasizes that even a low delinquency rate can significantly contribute to the growth of student debt in the U.S.

Table C1: Changes in total outstanding student debt

1979–2015	Δ Student debt
Wage garnishment	+364B
No delinquency choice	+210B

Note: This shows the total changes in student debt in a wage garnishment model with and without the delinquency option.

A.4 Additional Figures and Tables

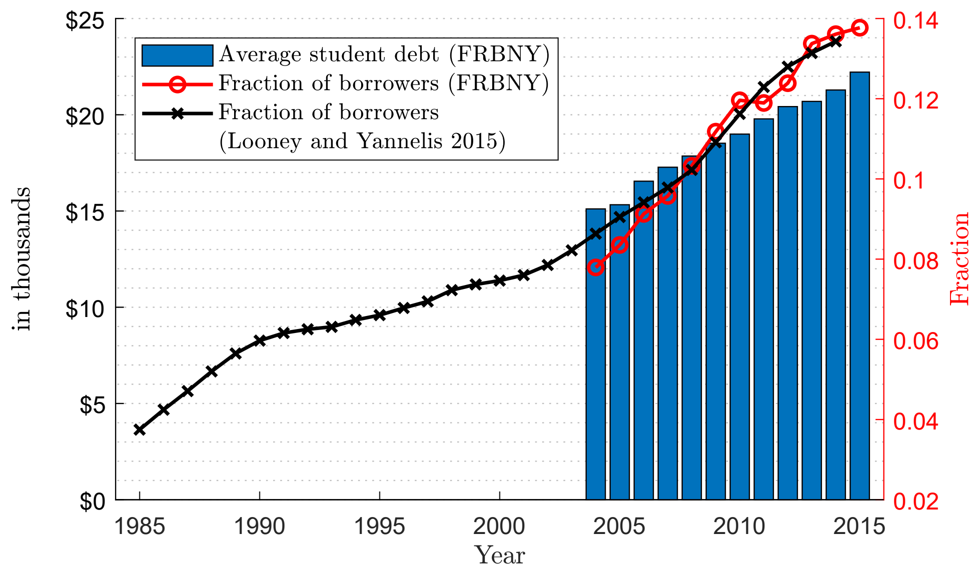
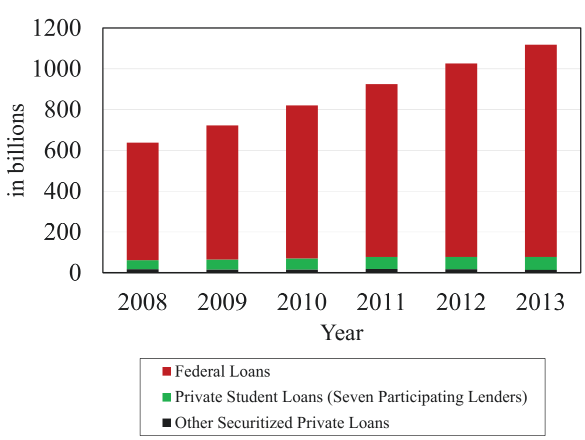


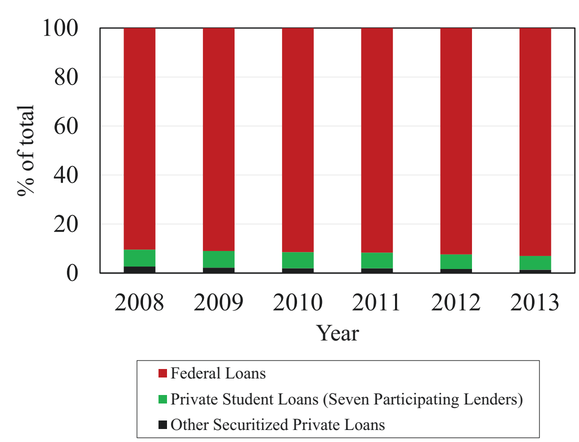
Figure D1: Average student debt and the number of borrowers

Source: The average student debt per borrower is from the National post-secondary student aid study (NPSAS) and the New York Fed Consumer Credit Panel. The total number of borrowers is from the New York Fed Consumer Credit Panel data. All the values are expressed in 2004 U.S. dollars.

Figure D2: Composition of outstanding student loans



(a) In current dollars



(b) As a fraction of total debt

Source: The 2013 MeasureOne Private Student Loan report. Notes: In July 2012, the Department of Education and the Consumer Financial Protection Bureau (CFPB) released a study of the private education loan market based, in part, on data submitted by nine major private education lenders. The 2013 MeasureOne Private Student Loan report provides an update to and extends the CFPB study to 2013. The MeasureOne collected data from the nation's seven largest active private student lenders, including Discover Bank; The First Marblehead Corporation; PNC Bank; RBS Citizens; Sallie Mae; SunTrust Banks; and Wells Fargo Bank.

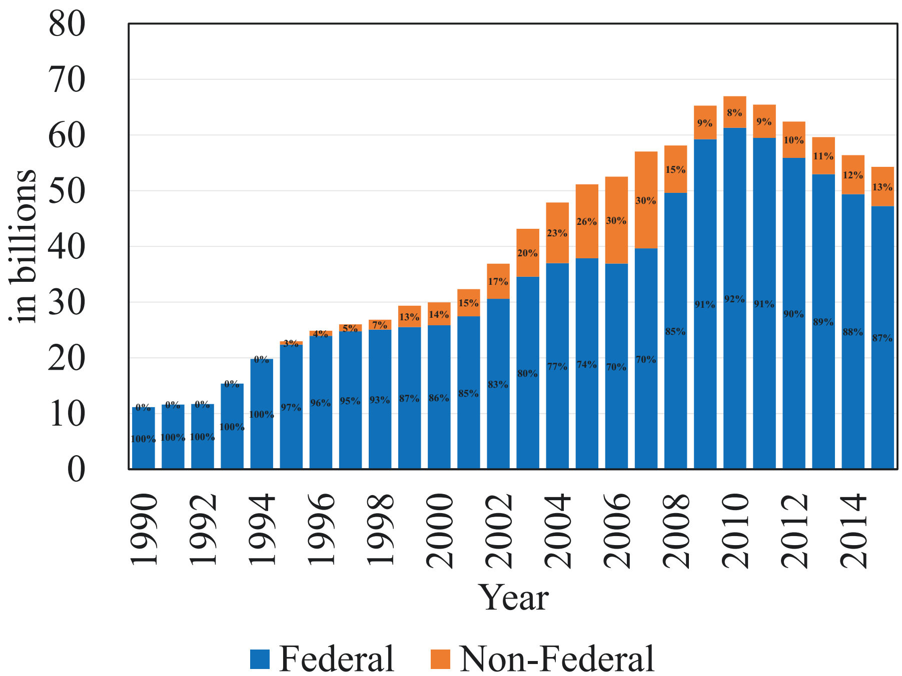
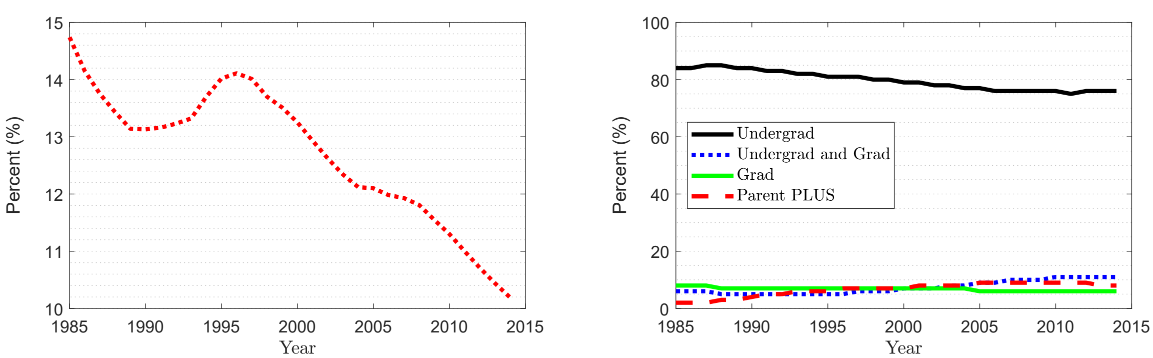


Figure D3: Annual amount of disbursed loans to undergraduate students
Source: College Board (2018). Expressed in 2004 U.S. dollars.

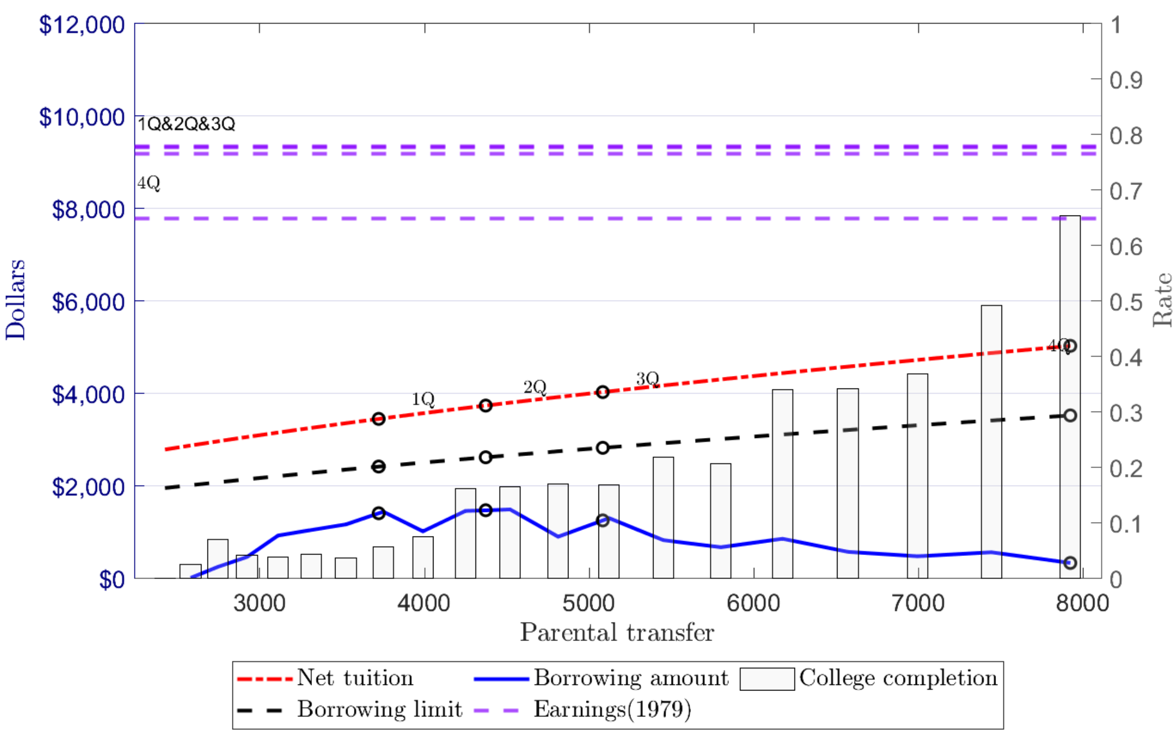


(a) Percentage of federal student loan for graduate study

(b) Percentage of federal student loan borrowers by loan type

Figure D4: Percentages of aggregate federal student debt for graduate study and student loan borrowers by loan type
Source: Looney and Yannelis (2015).
Note: Figure D4 shows the percentage of total outstanding federal student loan balances attributable to graduate school institutions (top) and the percentage of federal student loan borrowers by different types of federal loans (bottom). The data is taken from 4 percent of the National Student Loan Data System sample, which includes the annual information on student loans and institutions attended for about 4 million federal student loan borrowers.

Figure D5: Simulation over parental transfers



(a) 1979 economy

Note: This figure shows the 1979 simulated moments over parental transfers. The red dash-dot line represents the average net tuition and fees, the black dashed line represents the effective borrowing limit, and the blue line represents the borrowing amount. The white bars represent the college completion rates, and the black circled marks indicate where the parental transfer quartile is divided. The left y-axis indicates the dollar values of moments, while the right y-axis indicates the college complication rates.

Appendix B

Appendix for Chapter 2

B.1 OLS Regression

To examine the correlation between student debt after college graduation and initial labor market outcomes, I run pooled OLS regressions, and the results are shown in Table D1. The variables of interest are the student loan amounts (in $\$1,000$) in columns (1) to (6) and the dummy for holding student debt after college graduation in columns (7) to (12). The dependent variables are as below:

- Columns (1), (2), (7), and (8): duration of unemployment for the first job
- Columns (3), (4), (9), and (10): underemployment dummy
- Columns (5), (6), (11), and (12): log of initial earnings at the first job

Columns (1), (3), (5), (7), (9), and (11) control for the unemployment rate at the year of the first job, work experience while enrolled in college, ASVAB score, and majors. Columns (2), (4), (6), (8), (10), and (12) further control for gender, mother's education, race, and regions.

In column (2), the estimated coefficient for the student debt variable is -0.0394 and significant, indicating that an $\$1,000$ increase in student debt after college graduation is associated with 0.0394 weeks decrease in an unemployment period for the first job.[1] In column (4), the estimated coefficient for the student debt variable is -0.00064, indicating that an $\$1,000$ increase in student debt after college graduation is associated with 0.064% decrease in the probability of becoming underemployed. Lastly, in column (6), the

[1] The 0.02 weeks decrease is approximated by multiplying the coefficient -0.23 by $\ln(1.1)$. The negative significant effect of student debt on the unemployment period using OLS regressions is also found in Ji (2021a) and Luo and Mongey (2019b).

Table D1: Pooled OLS regression summary

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Student debt($\$1,000$)	-0.0413***	-0.0394***	-0.000587*	-0.000640**	0.00169***	0.00143***						
Student debt holder							-3.034***	-3.007***	-0.0327**	-0.0339**	0.0666***	0.0552**
Unemployment rate	-3.525***	-3.565***	0.0746***	0.0704***	-0.0994***	-0.102***	-3.755***	-3.822***	0.0729***	0.0683***	-0.0943***	-0.0970***
Work experience in college	-19.30***	-19.47***	0.152***	0.147***	0.0295	0.0247	-19.05***	-19.26***	0.156***	0.150***	0.0266	0.0224
ASVAB Score	0.0240**	0.0317***	-0.00116***	-0.00153***	-0.00300***	-0.00281***	0.0237**	0.0333***	-0.00112***	-0.00147***	-0.00304***	-0.00285***
Major: Technology	-5.311***	-6.154***	-0.102***	-0.101***	-0.0471	-0.0388	-5.566***	-6.345***	-0.104***	-0.104***	-0.0427	-0.0371
Major: Engineering	-5.028***	-5.191***	-0.144***	-0.155***	0.244***	0.233***	-5.031***	-5.146***	-0.145***	-0.156***	0.245***	0.232***
Major: Math	-6.813***	-7.642***	-0.159***	-0.165***	0.0725	0.0852	-7.198***	-7.989***	-0.163***	-0.169***	0.0743	0.0858
Major: Business	-3.118***	-2.905***	0.0381**	0.0293*	0.140***	0.127***	-3.397***	-3.175***	0.0356*	0.0269+	0.141***	0.127***
Major: Others	-1.209*	-1.304*	0.102***	0.102***	-0.0836***	-0.0920***	-1.440*	-1.515**	0.0982***	0.0981***	-0.0790***	-0.0889***
Mother has a college degree		2.966***		-0.00689		-0.108***		2.703***		-0.00766		-0.106***
Male		0.383		0.0293**		0.0154		0.242		0.0290**		0.0173
Non Black and non Hispanic		-5.075***		0.0285+		0.0660*		-5.230***		0.0261+		0.0653*
Region: North central		0.489		0.00558		-0.0563*		0.797		0.00972		-0.0646*
Region: South		-1.258**		-0.104***		-0.0181		-1.116*		-0.102***		-0.023
Region: West		-2.048***		-0.0430**		-0.0216		-1.964***		-0.0444**		-0.0237
Constant	39.49***	42.84***	0.126*	0.188***	10.28***	10.29***	42.01***	45.51***	0.143**	0.206***	10.24***	10.26***
F	553	343.6	76.59	57.08	22.61	15.84	537.5	342.3	76.57	57.08	21.61	15.39

Note: The dummy for work experience in college indicates whether an individual worked at the last year of college education. The omitted dummy for Major is Science. The omitted dummy for Region is North East.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

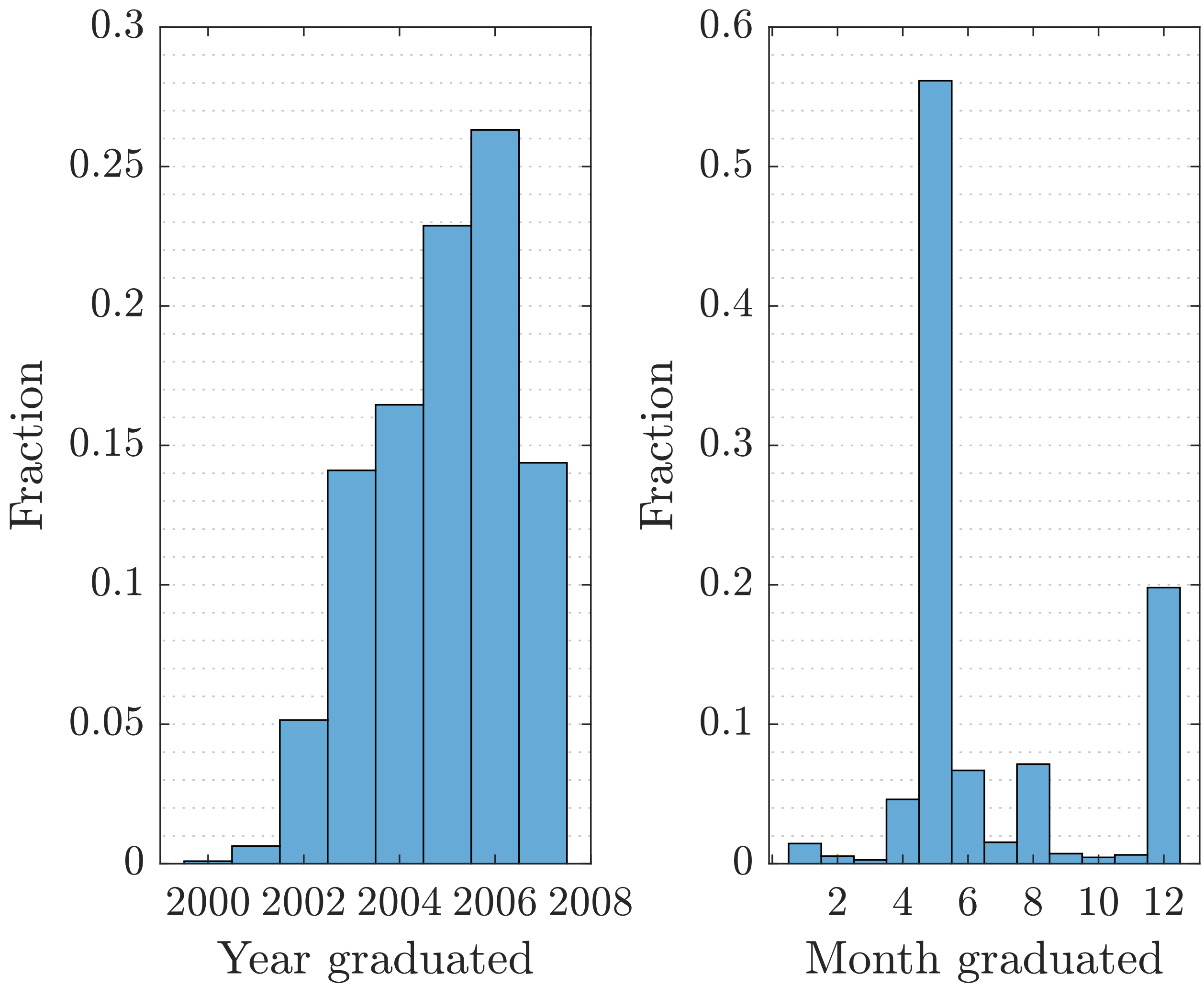
estimated coefficient for the student debt variable is 0.00143, indicating that an $\$1,000$ increase in student debt after college graduation is associated with 0.143% increase in the initial earnings at the first job.

The results are similar and significant, with a greater magnitude of effect, when I run the regressions using the dummy for holding student debt after college graduation as the variable of interest. For instance, in column (8), the coefficient for the student debt dummy is -3.007, indicating that holding student debt after college graduation is associated with 3 weeks decrease in an unemployment period for the first job. In column (10), the estimated coefficient for the student debt dummy is -0.0327, indicating that holding student debt after college graduation is associated with 3.27% decrease in the probability of becoming underemployed at the first job. Lastly, in column (12), the estimated coefficient for the student debt dummy is 0.0552, indicating that holding student debt after college graduation is associated with 5.52% higher initial earnings at the first job.

Overall, the results imply that student debt after 4-year college graduation is associated with shorter period of unemployment before finding the first job, a lower probability of becoming underemployed, and a higher initial income at the first job. Also, the results indicate that the effect of student debt is larger when we compare student debt holders with non-student debt holders than when we look at the change in the level of student debt. These evidences are, however, limited to suggesting the association between student debt and initial labor market outcomes. In fact, the results from the OLS regressions do not imply causation and could be biased due to the endogeneity of student debt.

B.2 Year and Month of graduation

Figure D1: Distribution of years and months of graduation



Note: Figure D1 shows the years (left) and months(right) at which individuals from the NLSY97 sample graduate from college.

References

Abbott, B., Gallipoli, G., Meghir, C., & Violante, G. L. (2019). Education Policy and Intergenerational Transfers in Equilibrium. *Journal of Political Economy*, *127*(6), 2569–2624.

Abel, J. R., & Deitz, R. (2017). Underemployment in the early careers of college graduates following the great recession. In *Education, skills, and technical change: Implications for future us gdp growth* (pp. 149–181). University of Chicago Press.

Austin, P. C., & Fine, J. P. (2017, 09). Practical recommendations for reporting fine-gray model analyses for competing risk data. *Journal of the Royal Statistical Society. Series B (Methodological)*, *36*, 4391–4400.

Belley, P., & Lochner, L. (2007, December). The Changing Role of Family Income and Ability in Determining Educational Achievement. *Journal of Human Capital*, *1*(1), 37–89. Retrieved 2023-03-15, from http://www.journals.uchicago.edu/doi/10.1086/524674 doi: 10.1086/524674

Brown, M., Haughwout, A., Lee, D., Scally, J., & Klaauw, W. v. d. (2014). *Measuring Student Debt and Its Performance* (Tech. Rep.).

Cai, Z., & Heathcote, J. (2022, January). College Tuition and Income Inequality. *American Economic Review*, *112*(1), 81–121. Retrieved 2023-03-15, from https://pubs.aeaweb.org/doi/10.1257/aer.20181027 doi: 10.1257/aer.20181027

Carneiro, P., & Heckman, J. J. (2002, October). The Evidence on Credit Constraints in Post-Secondary Schooling. *The Economic Journal*, *112*(482), 705–734. Retrieved 2023-03-15, from https://academic.oup.com/ej/article/112/482/705/5079515 doi: 10.1111/1468-0297.00075

Chetty, R., Friedman, J., Saez, E., Turner, N., & Yagan, D. (2017, July). *Mobility Report Cards: The Role of Colleges in Intergenerational Mobility* (Tech. Rep. No. w23618). Cambridge, MA: National

Bureau of Economic Research. Retrieved 2023-03-15, from http://www.nber.org/papers/w23618.pdf doi: 10.3386/w23618

Clark, B., Joubert, C., & Maurel, A. (2017). The career prospects of overeducated americans. *IZA Journal of Labor Economics*, *6*(3).

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 187–220.

Cuccaro-Alamin, S., & Choy, S. P. (1998). *Postsecondary Financing Strategies: How Undergraduates Combine Work, Borrowing, and Attendance* (Tech. Rep.). National Center For Education Statistics.

Cunha, F., Heckman, J., & Navarro, S. (2004, December). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers*, *57*(2), 191–261. Retrieved 2023-03-15, from https://academic.oup.com/oep/article-lookup/doi/10.1093/oep/gpi019 doi: 10.1093/oep/gpi019

Feldman, D. C. (1996). The nature, antecedents and consequences of underemployment. *Journal of Management*, *22*(3), 385-407. Retrieved from https://www.sciencedirect.com/science/article/pii/S0149206396900306 doi: https://doi.org/10.1016/S0149-2063(96)90030-6

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*, 1–21.

French, E. (2004). The Labor Supply Response to (Mismeasured but) Predictable Wage Changes. *The Review of Economics and Statistics*, *86*(2), 602–613.

Gervais, M., & Ziebarth, N. L. (2019). Life after debt: postgraduation consequences of federal student loans. *Economic Inquiry*, *1*(3), 1342–1366.

Gordon, G., & Hedlund, A. (2020). Accounting for Tuition Increases Across U.S. Colleges. *SSRN Electronic Journal*. Retrieved 2023-03-15, from https://www.ssrn.com/abstract=3720860 doi: 10.2139/ssrn.3720860

Guvenen, F. (2009, January). An empirical investigation of labor income processes. *Review of Economic Dynamics*, *12*(1), 58–79. Retrieved 2023-03-15, from https://linkinghub.elsevier.com/retrieve/pii/S1094202508000306 doi: 10.1016/j.red.2008.06.004

Hai, R., & Heckman, J. J. (2017, April). Inequality in human capital and endogenous credit constraints. *Review of Economic Dynamics*, *25*, 4–36. Retrieved 2023-03-15, from https://linkinghub.elsevier.com/retrieve/pii/S1094202517300029 doi: 10.1016/j.red.2017.01.001

Heathcote, J., Storesletten, K., & Violante, G. L. (2010, August). The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, *118*(4), 681–722. Retrieved 2023-03-15, from https://www.journals.uchicago.edu/doi/10.1086/656632 doi: 10.1086/656632

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. In *Handbook of the Economics of Education* (Vol. 1, pp. 307–458). Elsevier. Retrieved 2023-03-15, from https://linkinghub.elsevier.com/retrieve/pii/S1574069206010075 doi: 10.1016/S1574-0692(06)01007-5

Ionescu, F. (2009, January). The Federal Student Loan Program: Quantitative implications for college enrollment and default rates. *Review of Economic Dynamics*, *12*(1), 205–231. Retrieved 2023-03-15, from https://linkinghub.elsevier.com/retrieve/pii/S1094202508000550 doi: 10.1016/j.red.2008.09.004

Ionescu, F., & Simpson, N. (2016, March). Default risk and private student loans: Implications for higher education policies. *Journal of Economic Dynamics and Control*, *64*, 119–147. Retrieved 2023-03-15, from https://linkinghub.elsevier.com/retrieve/pii/S0165188915002183 doi: 10.1016/j.jedc.2015.12.003

Ionescu, F. A. (2008, January). Consolidation of Student Loan Repayments and Default Incentives. *The B.E. Journal of Macroeconomics*, *8*(1). Retrieved 2023-03-15, from https://www.degruyter.com/document/doi/10.2202/1935-1690.1682/html doi: 10.2202/1935-1690.1682

Irwin, M., & Kim, H. (2023). *The impact of bankruptcy reform on education and student loans* (Work in progress).

Ji, Y. (2021a). Job search under debt: Aggregate implications of student loans. *Journal of Monetary Economics*, *117*, 741-759. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304393220300672 doi: https://doi.org/10.1016/j.jmoneco.2020.05.002

Ji, Y. (2021b, January). Job Search under Debt: Aggregate Implications of Student Loans. *Journal of Monetary Economics*, *117*, 741–759. Retrieved 2023-03-15, from https://linkinghub.elsevier.com/retrieve/pii/S0304393220300672 doi: 10.1016/j.jmoneco.2020.05.002

Keane, M. P., & Wolpin, K. I. (n.d.). The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment. *International Economic Review*, *42*(4), 1051–1103.

Kim, H. (2022, January). Education, wage dynamics, and wealth inequality. *Review of Economic Dynamics*, *43*, 217–240. Retrieved 2023-03-15, from https://linkinghub.elsevier.com/retrieve/pii/S1094202521000107 doi: 10.1016/j.red.2021.02.006

Lee, S. Y. T., Shin, Y., & Lee, D. (2015). *The option value of human capital: Higher education and wage inequality* (Tech. Rep.). National Bureau of Economic Research.

Leuven, E., & Oosterbeek, H. (2011). Chapter 3 - overeducation and mismatch in the labor market. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 4, p. 283-326). Elsevier. Retrieved from https://www.sciencedirect.com/science/article/pii/B9780444534446000031 doi: https://doi.org/10.1016/B978-0-444-53444-6.00003-1

Lochner, L. J., & Monge-Naranjo, A. (2011, October). The Nature of Credit Constraints and Human Capital. *American Economic Review*, *101*(6), 2487–2529. Retrieved 2023-03-15, from https://pubs.aeaweb.org/doi/10.1257/aer.101.6.2487 doi: 10.1257/aer.101.6.2487

Looney, A., & Yannelis, C. (2015). A Crisis in Student Loans?: How Changes in the Characteristics of Borrowers and in the Institutions They Attended Contributed to Rising Loan Defaults. *Brookings Papers on Economic Activity*, *2015*(2), 1–89. Retrieved 2023-03-15, from https://muse.jhu.edu/article/616850 doi: 10.1353/eca.2015.0003

Luo, M., & Mongey, S. (2019a, May). *Assets and Job Choice: Student Debt, Wages and Amenities* (Tech. Rep. No. w25801). Cambridge, MA: National Bureau of Economic Research. Retrieved 2023-03-15, from http://www.nber.org/papers/w25801.pdf doi: 10.3386/w25801

Luo, M., & Mongey, S. (2019b, May). *Assets and job choice: Student debt, wages and amenities* (Working Paper No. 25801). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w25801 doi: 10.3386/w25801

Mezza, A., Ringo, D., Sherlund, S., & Sommer, K. (2020). Student loans and homeownership. *Journal of Labor Economics*, *38*(1), 215–260.

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

Moschini, E. G., Raveendranathan, G., & Xu, M. (2023). *Over-optimism About Graduation and College Financial Aid* (Working Paper).

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

Rothstein, J., & Rouse, C. E. (2011b, February). Constrained after college: Student loans and early-career occupational choices. *Journal of Public Economics*, *95*(1-2), 149–163. Retrieved 2023-03-15, from https://linkinghub.elsevier.com/retrieve/pii/S0047272710001337 doi: 10.1016/j.jpubeco.2010.09.015

Sieg, H., & Wang, Y. (2018). The impact of student debt on education, career, and marriage choices of female lawyers. *European Economic Review*, *109*, 124–147.

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, *27*(3), 531-543. Retrieved from https://www.sciencedirect.com/science/article/pii/S0167629607001063 doi: https://doi.org/10.1016/j.jhealeco.2007.09.009

Volkwein, J. F., Szelest, B. P., Cabrera, A. F., & Napierski-Prancl, M. R. (1998, March). Factors Associated with Student Loan Default among Different Racial and Ethnic Groups. *The Journal of Higher Education*, *69*(2), 206–237. Retrieved 2023-03-15, from https://www.tandfonline.com/doi/full/10.1080/00221546.1998.11775133 doi: 10.1080/00221546.1998.11775133

Weidner, J. (2016). Does Student Debt Reduce Earnings? *Princeton University*.