10-26-2020 9:30 AM

Essays on Student Loans and Returns to Skill

Qian Liu, *The University of Western Ontario*

Supervisor: Lochner, Lance, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Economics

Abstract

This thesis consists of three studies, which explore topics related to labor economics. Chapters 2 and 3 examine the returns on student loans and student loan repayment policy, respectively. Chapter 4 examines the returns to skill and the evolution of skills at older ages.

In Chapter 2 (co-authored with Lance Lochner), we study rates of return on government student loans in Canada using novel administrative data from the Canada Student Loans Program. We exploit rich information on personal characteristics, loan amounts, field of study, and institution of attendance to explain differences in rates of return across different types of borrowers. We find that field of study is a particularly important determinant of rates of return, explaining 60-70% of the variation in predicted returns across borrowers, while institution differences explain only about 10% of the variation. We also show that if private lenders were to cream-skim borrowers with predicted returns above 10% (5%), the average return would fall from the current -5% to -6.4% (-9.4%), raising the cost of the government student loan program and adverse selection concerns.

In Chapter 3, I study the effects of introducing an income-based student loan repayment (IBR) plan when considering labor market risks and the insurance provided by parents. I develop a dynamic life-cycle model with endogenous parents-to-children transfers, together with children's education, borrowing, repayment, and labor supply decisions. After estimating the model using the National Longitudinal Survey of Youth 1997, I quantify the impacts of introducing an IBR while keeping the government budget constant. IBR crowds out savings and parental transfers as it provides more insurance to borrowers. Interestingly, a weak labor supply response to IBR suggests that moral hazard is not a concern. Further, the college enrollment rate increases, and the largest gains are for low-income and low-ability families. Finally, aggregate welfare increases with relatively low-income families benefiting the most.

In Chapter 4 (co-authored with Lance Lochner, Youngmin Park, and Youngki Shin), we show that repeated cross-section data with multiple skill measures (one continuous and repeated) available each period are sufficient to nonparametrically identify the evolution of skill returns and cross-sectional skill distributions. With panel data and the same available measurements, the dynamics of skills can also be identified. Our identification strategy motivates a multi-step nonparametric estimation strategy. We further show that if any continuous repeated measurement is shown to be linear in skills, a much simpler GMM estimator can be used. Using Health and Retirement Survey data on men ages 52+ from 1996-2016, we show that one of the available (continuous and repeated) skill measures (word recall) is linear in skills and implement our GMM estimation approach. Our estimates suggest that the returns to skill were fairly stable from the mid-1990s to the Great Recession and rising thereafter. We document considerable differences in skills and lifecycle skill profiles over ages 52–70 across cohorts, with more recent cohorts possessing lower skills in their mid-50s but experiencing much weaker skill declines with age. We also document skill differences by education and race, which are stable across ages and explain roughly one-third and one-half, respectively, of the corresponding differences in wages.

Keywords: Student loans, return on student loans, income-based repayment, parental transfers, postsecondary education, returns to skill, skill dynamics.

Summary for Lay Audience

This thesis consists of three studies, which explore topics related to labor economics. Chapters 2 and 3 examine the returns on student loans and student loan repayment policy, respectively. Chapter 4 examines the returns to skill and the evolution of skills at older ages.

In Chapter 2 (co-authored with Lance Lochner), we calculate the net revenue the government receives by lending to undergraduate students in Canada. We also look at what types of borrowers are the most "profitable" to the government. We find that, on average, the government has lost 5 cents for every dollar it lent out. We also find that field of study is the most important predictor for repayment. The government gets more revenue by lending to students studying in law, education, health-related, and engineering-related majors compared to other majors. We also quantify the government revenue losses if private lenders were to siphon away the more profitable borrowers.

In Chapter 3, I study the impacts of introducing an income-based student loan repayment (IBR) plan for university graduates. There have been concerns among policy makers that increasing debt levels and uncertainty around finding a well paying job has made it increasingly difficult for current graduates to repay their student loans. This has led to policy proposals that the student loan repayment program should be income-based, i.e., repayment should be based on borrowers' income instead of their debt. However, IBR may reduce the revenue that the government can collect, and even encourage some people to work less. In this chapter, I study the effects of introducing an IBR while keeping the government budget fixed. My analysis considers one of the most important aspects of college financing — parental transfers, and examines how parental transfers and students with different levels of parental transfers respond to the new policy. I find that IBR leads to less savings and fewer parental transfers. Two important behavioural responses to IBR are that (i) very few borrowers work less because of IBR and (ii) IBR encourages more youth from low-income families to go to college. Social welfare increases because of the introduction of IBR.

In Chapter 4 (co-authored with Lance Lochner, Youngmin Park, and Youngki Shin), we study the returns to skill and the evolution of skills for older men in the U.S. The literature is not yet settled on whether the rising wage inequality in the U.S. is due to higher returns to skill or greater dispersion of skills across workers. Because skill is not directly observed, previous literature has had to make strong assumptions to differentiate between these two channels. In this paper, we develop a novel strategy to separate the returns to skill from the evolution of skills, without making any strong assumptions. Using data that repeatedly measures the cognitive abilities of men ages 52+, we show that one of the measures (word recall) is linear in skills. This finding is critical for us to separate the returns to skill from the evolution of skills. Our estimates suggest that the returns to skill were fairly stable from the mid-1990s to the Great Recession and rising thereafter. We also find that more recent cohorts possess lower skills in their mid-50s, but they experience much weaker skill declines with age. We observe substantial skill differences by education and race. These differences are stable across ages and explain roughly one-third and one-half, respectively, of the corresponding differences in wages.

Co-Authorship Statement

This thesis contains co-authored material. Chapter 2 is co-authored with Lance Lochner. Chapter 3 is co-authored with Lance Lochner, Youngmin Park, and Youngki Shin. All authors are equally responsible for the work.

Acknowledgements

I will be forever grateful to my supervisor Lance Lochner for his support and guidance. Without his mentorship and encouragement, this thesis would not be possible. I learned from him what a true scholar is and working with him has been invaluable in preparing me for a career as a researcher. I am also greatly indebted to the members of my supervisory committee, Nirav Mehta and Todd Stinebrickner, for their comments and suggestions that have helped to improve my research.

I would like to thank the professors who offered me valuable suggestions during my time at Western, including, but certainly not limited to, Audra Bowlus, Salvador Navarro, David Rivers, and Terry Sicular. I am also grateful to administrators Sandra Augustine, Karin Feulgen, Gary Kim, Debra Merrifield, Maureen O'Connell, and Sharon Phillips for their help. Support provided by Charlie Brown and Mohammad Mushtaq for accessing the NLSY97 Geo data at the Institute for Social Research (University of Michigan) is also greatly appreciated.

I owe a great deal of thanks to my cohort, Miguel Cardoso, Galyna Grynkiv, Tomasz Handler, Antonella Mancino, Sergii Pypko, and Diego Salazar, and also to my Chinese friends, Yifan Gong, Meng Li, Zhuang Liu, Sha Wang, Wenya Wang, and Yuxi Yao, for their companionship and support.

I would also like to thank Professor Ximing Yue, who was my supervisor during my study at Renmin University of China. He inspired me to pursue a career in research, and without his encouragement and support, I would not have had the opportunity to study at one of Canada's top economics departments.

Finally, I would like to thank my parents for their unconditional love and support.

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

This thesis consists of three studies which explore topics related to student loans and returns to skill. Chapters 2 and 3 contribute to the literature by examining topics related to student loans, and Chapter 4 contributes to the literature on wage inequality by studying returns to skill and the evolution of skills.

In Chapter 2 (co-authored with Lance Lochner), we quantify the returns on government student loans in Canada (purely from the government lender's perspective) and empirically examine the determinants of those returns. Both governments and private lenders are concerned about the expected returns on the loans they disburse and whether the amounts they collect in repayments will cover the amounts they lend out. In addition, the potential heterogeneity in the expected return across different types of borrowers raises an adverse selection concern. For example, private creditors may undercut the government loan program by cream skimming, i.e. poaching profitable borrower types and leaving the government with unprofitable ones, raising the cost of the government student loan program.

Using novel administrative data from the Canada Student Loans Program, we exploit rich information on personal characteristics, loan amounts, field of study, and institution of attendance to explain differences in rates of return across different types of borrowers. We find that field of study is a particularly important determinant of rates of return, explaining 60-70% of the variation in predicted returns across borrowers, while institution differences explain only about 10% of the variation. We also show that if private lenders were to cream-skim borrowers with predicted returns above 10% (5%), the average return would fall from the current -5% to -6.4% (-9.4%), raising the cost of the government student loan program and adverse selection concerns.

In Chapter 3, I study the effects of introducing an income-based student loan repayment (IBR) plan on college enrollment, student loan repayment, parental transfers, labor supply, government budget, and aggregate welfare. Increasing borrowing for post-secondary education, along with high student loan default rates, increasing labor market risk, and the need for parental assistance, indicates that youth may be inadequately insured through the current student loan repayment scheme. Some researchers and policy makers suggested that student loan repayment schemes should provide more insurance against labor market risk through more accessible IBR plans. However, there is concern that the provision of more insurance for student loan repayment could lead to large revenue losses for the student loan program. To make the program self-sustaining, interest rates may need to increase, raising the costs of post-secondary education for many who are currently paying their loans in full. This, in turn, may discourage some youth from attending college altogether.

To study the effects of student loan repayment policy when considering labor market risks and the insurance provided by parents, this paper develops a dynamic life-cycle model with endogenous parents-to-children transfers, together with children's education, borrowing, repayment, and labor supply decisions. After estimating the model using the National Longitudinal Survey of Youth 1997, I quantify the impacts of introducing an IBR plan while keeping the government budget constant. IBR crowds out savings and parental transfers as it provides more insurance to borrowers. Interestingly, a weak labor supply response to IBR suggests that moral hazard is not a concern. The college enrollment rate increases, and the largest gains are for low-income and low-ability families. Aggregate welfare increases with relatively low-income families benefiting the most.

Chapter 4 (co-authored with Lance Lochner, Youngmin Park, and Youngki Shin) studies the returns to skill and the evolution of skills for older men in the U.S. To explain the growing wage inequality in the U.S., researchers are typically forced to make strong assumptions to separate the rising returns to skill from the growth in the variance of skills across workers, due to a lack of direct measures of skill.

This paper shows that repeated cross-section data with multiple skill measures (one continuous and repeated) available each period are sufficient to nonparametrically identify the evolution of skill returns and the cross-sectional skill distributions. With panel data and the same available measurements, the dynamics of skills can also be identified. Our identification strategy motivates a multi-step nonparametric estimation strategy. We further show that if any continuous repeated measurement is shown to be linear in skills, a much simpler GMM estimator can be used.

Using Health and Retirement Study data on men ages 52+ from 1996-2016, we show that one of the available (continuous and repeated) skill measures is linear in skills and implement our GMM estimation approach. Our estimates suggest that the returns to skill were fairly stable from the mid-1990s to the Great Recession and rising thereafter. We document considerable differences in skills and lifecycle skill profiles over ages 52–70 across cohorts, with more recent cohorts possessing lower skills in their mid-50s but experiencing much weaker skill declines with age. We also document skill differences by education and race, which are stable across ages and explain roughly one-third and one-half, respectively, of the corresponding differences in wages. We observe substantial differences in skills for men in their mid-50s choosing to retire at different ages, but no clear evidence of sharp declines in skills surrounding retirement ages.

Chapter 2

Return on Student Loans in Canada

2.1 Introduction

As an increasing number of students borrow to finance their higher education, and many of them delay payments or even default, both governments and private lenders are concerned about the expected returns on the loans they disburse and whether the amounts they collect in repayments will cover the amounts they lend out. This is particularly important when considering adjustments to loan programs or shifts in the underlying pool of students that borrow. In this paper, we quantify the returns on government student loans in Canada (purely from the government lender's perspective) and empirically examine the determinants of those returns.

Under government student loan schemes, borrowers are generally offered the same contract with the same interest rates and repayment plans. While there are some exceptions, contracts do not depend on other aspects of borrowers like their field of study or institution of attendance.[1] Given that post-school earnings and, therefore, the potential to repay is closely related to these factors (Rumberger and Thomas, 1993; Loury and Garman, 1995; Altonji et al., 2012), the expected return on student loans is likely to depend on these characteristics as well and would be heterogeneous across individuals.

Heterogeneity in predictable returns across students raises concerns about the efficiency of student loan programs and implies *ex ante* cross-subsidization, where borrowers with high expected returns effectively subsidize those with low expected returns. Moreover, there is an

[1] For example, the U.S. offers subsidized and unsubsidized loans with different interest terms depending on the financial need of borrowers. Loan limits tend to vary more across borrowers based on factors like dependency status or financial need, year in school, and educational costs.

adverse selection concern that private creditors may undercut the government loan program by cream skimming, i.e. poaching profitable borrower types and leaving the government with unprofitable ones (Bachas, 2019).[2] This can raise the cost of the program, forcing the government to raise the interest rate on student loans or subsidize the remaining pool. A higher interest rate would likely drive additional high-return borrowers away, leading to an even more negatively selected pool.

A major challenge to estimating returns on student loans is a lack of publicly available data on repayment records over the full repayment period. Most data provide very limited snapshots of repayment status (e.g., the National Graduates Survey (NGS) in Canada and the Baccalaureate and Beyond Longitudinal Studies in the U.S.). As a result, previous studies generally examine student loan repayment status, such as default rates or the proportion of loan repaid, typically a few years after students leave school (Gross et al., 2009). Schwartz and Finnie (2002) use Canadian NGS data for the class of 1990 to study student loan repayment by characteristics such as gender and field of study. However, the NGS only reports the amount owed two years after graduation and self-identified problems with loan repayment. As discussed in Lochner and Monge-Naranjo (2015), knowing whether borrowers had ever defaulted or are currently in default as of some arbitrary date (especially early in the repayment process) does not necessarily provide an accurate picture of returns on student loans and how those returns differ across borrowers.

Using novel administrative data, this paper is the first to quantify the realized returns on government student loans, link them to the characteristics of borrowers, and estimate predicted returns for different borrowers based on characteristics (and choices) observed by the government at the time loans are disbursed. We use data from the new Education and Labor Market Longitudinal Platform (ELMLP) provided by Statistics Canada. Our main data source comes from the Canada Student Loans Program (CSLP) administrative records on student loan disbursements, needs determination, and repayments. These data provide extensive information on the repayment and characteristics of each borrower and are available from 2003-2004 to

[2]This practice is already underway in the U.S. by companies such as SoFi (Kosir et al., 2015). Using data from a student loan refinancing firm, Bachas (2019) provides a careful analysis of these issues in the U.S., including the impact of offering risk-based interest rates in the student loan market on borrower welfare and government revenue.

2015-2016, allowing us to observe nearly complete repayment streams for borrowers entering repayment in the mid-2000s. The ELMLP also allows us to link the CSLP data to administrative records from the Postsecondary Student Information System (PSIS), which includes additional institutional and individual characteristics.

For reasons discussed below, we use these data to study rates of return on student loans for the cohort of undergraduate borrowers that attended public post-secondary schools in the Atlantic provinces of Canada and received their last undergraduate loan disbursement in 2005. We use eleven years of data to calculate the realized rate of return for each borrower in our sample. With these individual-specific realized returns, we estimate average returns to CSLP for the full cohort of borrowers, as well as for different observable groups based on their demographic characteristics, field of study, and post-secondary institution — all factors known by the CSLP at the time loans are disbursed. We then regress individual-specific returns on all of these characteristics, as well as the amount borrowed, to estimate the importance of each of these factors when taking into account the influence of others. These results demonstrate that field of study is a particularly important determinant of rates of return, with differences in returns between fields as high as 35 percentage points. Differences between ranked and unranked universities are more modest, generally less than 10 percentage points. While differences across demographic groups are generally modest, we do estimate a 10 percentage point lower return for permanent residents compared to Canadian citizens.

Finally, our estimates enable us to calculate predicted rates of return for borrowers based on the amount borrowed, demographic characteristics, field of study, and institution of attendance. Altogether, the predictors we include explain as much as 14% of the variation in realized returns, so even the very simple specification we consider predicts a sizeable amount of the variation in returns on student loans. Most interestingly, differences in field of study explain 60-70% of the variation in predicted returns across borrowers, while institution indicators (for each of the 12 ranked universities in the Atlantic provinces and an indicator for two-year college attendance) explain only about 10% of the variation in predicted returns. Using the full distribution of predicted returns, we also explore the extent to which average rates of return to the CSLP would change if private lenders were to cream-skim different high-end borrowers. These results suggest that losing all borrowers with predicted returns above 10% (5%) would reduce the rate of return on CSLP loans from the current -5% to -6.4% (-9.4%). Thus, concerns about adverse selection are not unfounded.

The chapter proceeds as follows. Section 2.2 provides an overview of the CSLP, while Section 2.3 discusses our approach to quantifying the rate of return on student loans. Section 2.4 describes the CSLP and PSIS data, as well as our sample selection criteria. Section 2.5 presents our empirical results on average rates of return on CSLP loans, the estimated importance of observable factors for rates of return, and the distribution of predicted rates of return. We also discuss the potential impacts of cream-skimming by private lenders on CSLP returns. Section 2.6 concludes.

2.2 Overview of CSLP

CSLP provides loans and grants to help Canadian students pay for their post-secondary education based on their financial needs. The federal government works in collaboration with most provincial/territorial governments to coordinate the delivery of student loans and grants.[3]

Loans are interest free while borrowers are in-study. After borrowers leave school, the loans received during school are combined into a single repayment plan; this process is called loan consolidation. Each loan is consolidated after a six month grace period following the study end date. Borrowers begin to repay after consolidation. However, interest begins to accrue as of the first day after the study end date, and interest accumulated during the grace period is added to the principal balance at the time of loan consolidation.

Consolidation establishes the repayment agreement, including amortization period and interest rate. The default repayment plan is a standard debt-based payment plan with an amortization period of 9.5 years. Borrowers can extend their amortization period to a maximum of 14.5 years. Borrowers have the choice of a floating interest rate (default option) or a fixed interest rate.[4] According to Office of the Chief Actuary (2019), 99% of borrowers choose the

[3] All provinces and territories participate in the CSLP, except Quebec, the Northwest Territories and Nunavut, which operate their own financial assistance programs and receive alternative funding from the federal government.

[4] Prior to 2019, the floating interest rate was the prime rate plus 2.5 percentage points and the fixed rate was prime plus 5 percentage points. Effective on 1 November 2019, the floating interest rate was lowered to prime and the fixed rate was lowered to prime plus 2 percentage points.

floating rate.

The CSLP also provides repayment assistance to borrowers who have difficulty repaying their student loans. In 1983, the CSLP introduced an Interest Relief (IR) program to provide short-term relief to borrowers in financial difficulty while repaying the loan. Borrowers receiving IR were not required to make any payments for 3 months (extended to 6 months in 1998) at a time. Interest did not accrue on the loans while a borrower was on IR.[5]

In 2009, the federal government introduced the Repayment Assistance Plan (RAP) to replace the previous IR. Students experiencing financial difficulty and needing assistance must apply for RAP, as enrollment is not automatic. If approved, borrowers receive repayment assistance for a period of six months. If assistance is needed after the six-month period, borrowers must reapply. Under RAP, financial difficulty is determined by comparing a calculated "affordable" monthly payment to a calculated required monthly payment. The affordable payment is determined as a fraction (ranging from zero to 20%) of income above a pre-defined threshold based on family size, while the required payment is calculated based on the outstanding loan amount. If the affordable payment is less than the required monthly payment, the borrower qualifies for RAP and need only make the lower affordable payment. In most cases, borrowers enrolled in RAP make no payment at all.

There are two stages in RAP: Interest Relief (Stage 1) and Debt Reduction (Stage 2). Stage 1 is similar to the older IR: the Government covers all interest that is not covered by the reduced borrower's payment. A borrower eligible for RAP could receive this benefit for up to 60 months. In Stage 2, the Government covers all interest and principal amounts not covered by the borrower's payment, so that student debt is eliminated within a maximum of 15 years from leaving school (or 10 years for borrowers with permanent disabilities).

A loan is considered in default when the borrower has missed 9 months of payments, and the loan is sent to the Canada Revenue Agency (CRA) for collection. Once in collection, borrowers cannot receive any form of repayment assistance and must rehabilitate their loans in order to have access to repayment assistance. Actuarial reports on the CSLP show a recovery rate (i.e. the fraction of the loan recuperated after a default) of roughly 30% in recent years (Office of the Chief Actuary, 2019).

[5] See Situ (2006) for more details about IR.

Some borrowers enter bankruptcy during their repayment period. In principle, borrowers filing for bankruptcy are not shielded from student loan payments if they file within seven years of leaving school; however, student loans can be discharged beyond the seven year point.

2.3 Measuring Rates of Return on Student Loans

This section describes our measure of rates of return and our use of CSLP data to calculate individual-specific returns. We also discuss our method for estimating predicted return based on characteristics observed by the government at the time loans are disbursed.

2.3.1 Realized Returns

We define the net rate of return on a loan to borrower i, R_i, as follows:

$$R_i = \frac{\sum_{t=1}^{\overline{T}_i} \frac{P_{i,t}}{(1+d)^t}}{L_{i,0}} - 1, \tag{2.1}$$

where $\overline{T}_i$ is the number of years over which repayment takes place, $\{P_{i,t}\}_{t=1}^{\overline{T}_i}$ is the sequence of payments, $L_{i,0}$ is the initial loan amount, and d is the constant discount rate used to calculate the return from the lender's perspective.

In practice, the discount rate d would typically reflect the government's cost of funds, potentially with an adjustment factor to account for aggregate risk as discussed in Lucas and Phaup (2010) and Lucas and Moore (2010a). It may differ from the nominal interest rate, r, charged on the loan and used to determine payment amounts.[6] Note that if $d = r$ and the borrower repays the loan in full, then $R_i = 0$ and the lender breaks even as the discounted flow of payments equals the initial loan amount. Furthermore, the timing of payments is irrelevant for the rate of return when $d = r$.

If there is any chance that some borrowers will not repay in full, one would expect the

[6]The standard debt-based payment is based on traditional annuity formula based on the current balance $L_{i,t}$, remaining years on the loan τ, and interest rate r: $L_{i,t}\left[\frac{r(1+r)^\tau}{(1+r)^\tau-1}\right]$. The Office of the Chief Actuary publishes annual actuarial reports on the CSLP, which report the Government's cost of borrowing and nominal interest rates. For example, see Office of the Chief Actuary (2019).

lender to set $r > d$ to cover its losses on those who do not repay. When $r > d$, the return $R_i > 0$ for borrowers that fully repay their loan (assuming they do not repay before any interest is accumulated). The higher the nominal rate relative to the discount rate, the higher the rate of return, keeping all else constant. When $r \neq d$, the timing of repayments also affects the rate of return to the lender. For $r > d$ ($r < d$), the return is higher when payments are shifted later (earlier) as interest accumulates on the loan at a higher (lower) rate than the discount rate. Of course, the return is reduced when the borrower does not repay in full, as is often the case with default or bankruptcy. The earlier a borrower enters default or bankruptcy, the lower the return if the borrower never makes another payment. By contrast, if $r > d$, then borrowers that default for a period of time before returning to their standard payments (and eventually repaying their loans in full) will generate a higher return than borrowers who make the standard payments, since delayed payments lead to higher returns when $r > d$. These cases make clear that standard measures of default at some point in time (or whether someone had ever defaulted up until some point) are not necessarily informative about the rate of return from that borrower, especially when $r > d$ as is commonly the case. Someone who defaulted in year two of repayment may yield a very low return if they never make another payment, but they could also yield a high return if their default is temporary and they eventually pay off their loan in full. From the lender's perspective, the rate of return R_i appropriately captures all of these possibilities and provides an accurate reflection of a borrower's creditworthiness. Of course, calculating R_i requires observing borrowers over their entire repayment periods.

In practice, we may not observe borrowers for their entire repayment history, making equation (2.1) impossible to implement. While we observe repayments for up to 11 years in our sample, many borrowers have not repaid their loans in full by that time (despite a "standard" repayment period of 10 years). We, therefore, make several informed assumptions about payments after the observed period of repayment based on what we know about the CSLP.

Suppose we only observe payments for $T_i \leq \overline{T}_i$. Letting $\tilde{P}_{i,T_i+1}$ reflect the discounted present value of all remaining payments after T_i, the empirical rate of return can be calculated as follows:

$$R_i^e = \frac{\sum_{t=1}^{T_i} \frac{P_{i,t}}{(1+d)^t} + \tilde{P}_{i,T_i+1}}{L_{i,0}} - 1. \tag{2.2}$$

We make different assumptions about $\tilde{P}_{i,T_i+1}$ depending on the loan status in period T_i and, in some cases, the last observed loan amount L_{i,T_i} and payment P_{i,T_i}. We consider five possible loan statuses at the end of period T_i: (1) the loan is fully repaid; (2) the loan is not fully repaid but in good standing and standard repayment; (3) the loan is not fully repaid and the borrower is in income-contingent repayment (i.e. RAP); (4) the loan is in default; (5) the borrower has declared bankruptcy. We discuss our assumptions about $\tilde{P}_{i,T_i+1}$ in each of these cases.

First, if the borrower fully repays the loan by period T_i, the outstanding balance at T_i is $L_{i,T_i} = 0$ and, therefore, $\tilde{P}_{i,T_i+1} = 0$.

Second, if the loan is not fully repaid and the borrower is making standard debt-based payments at the end of T_i, we assume that she will continue to make the same annual payments as she made in year T_i until the loan is paid off, unless that would take more than 15 years from the date of consolidation for their last loan.[7] If that would require payments beyond 15 years from consolidation, we assume that she makes a constant payment (determined by the standard annuity formula) to repay the loan within that time frame.[8]

Third, if the loan is not fully repaid and the borrower is making income-based payments (under RAP) at the end of T_i, we assume that they will continue to make the same annual payment as they made in year T_i until their time on RAP is over or the loan is fully repaid, whichever occurs sooner.[9]

Fourth, if the borrower has defaulted by T_i, we assume that 30% of the remaining balance will be collected, i.e., $\tilde{P}_{i,T_i+1} = 0.3L_{i,T_i}/(1+d)^{T_i}$. This assumption is consistent with recovery

[7]Unfortunately, the nominal interest rate r is not directly available from our data. To calculate the accumulation of interest after T_i, we assume $r = 5.5\%$, which is the floating rate of student loans from year 2010 to 2014 (and did not change much until 2018).

[8]We also consider an alternative assumption that full repayment is made in year $T_i + 1$: $\tilde{P}_{i,T_i+1} = \frac{L_{i,T_i}}{(1+d)^{T_i}}$. For $r > d$, this implies a slightly lower return than our baseline assumption. The main findings presented in Section 2.5 remain the same under this alternative assumption; although, the average rate of return is slightly lower.

[9]An alternative (extreme) assumption is that zero payments are collected beyond year T_i (i.e., $\tilde{P}_{i,T_i+1} = 0$). The main data findings presented in Section 2.5 remain the same under this alternative assumption with slightly lower average returns.

rates in recent years (Office of the Chief Actuary, 2019).[10]

Fifth, if the borrower has declared bankruptcy by T_i, we assume that no additional payments will be collected (i.e., $\tilde{P}_{i,T_i+1} = 0$). While information on collections among students entering bankruptcy is scarce, conversations with officials at CSLP suggest that, in practice, little is collected from those entering bankruptcy.

2.3.2 Average and Predicted Returns

To calculate average or predicted returns for borrowers, we assume that repayments after period T_i do not systematically differ across subgroups of the population (conditional on remaining debt and repayment status). Under this assumption, we can estimate *ex ante* expected returns for any population subgroup by using the average realized return, R_i^e, for that subgroup.

We can also estimate the expected return for borrowers conditional on observed characteristics $\boldsymbol{X}_i$, $E(R_i|\boldsymbol{X}_i)$, using a regression approach. Consider a simple model of returns $R_i = \boldsymbol{X}_i\boldsymbol{\beta}+\varepsilon_i$, where the vector $\boldsymbol{X}_i$ includes observable characteristics such as the amount borrowed, demographic characteristics, field of study, and institution of attendance, while ε_i reflects idiosyncratic variation in returns conditional on $\boldsymbol{X}_i$ (with $E(\varepsilon_i|\boldsymbol{X}_i) = 0$). In this case, regressing R_i^e on $\boldsymbol{X}_i$ yields the estimated coefficient vector $\hat{\boldsymbol{\beta}}$, which can be used to calculate predicted returns $\hat{R}_i \equiv \boldsymbol{X}_i\hat{\boldsymbol{\beta}}$ for a borrower with characteristics $\boldsymbol{X}_i$. In the absence of aggregate risk, these predicted returns provide an estimate of *ex ante* expected returns for lenders with rational expectations, i.e., $E(R_i|\boldsymbol{X}_i)$.

2.4 Data

The empirical analysis in this paper uses administrative data from the CSLP and the PSIS. We briefly discuss each of these sources and the sample we use for our analysis.

[10] Note that our data do not include any information about a loan after default, so we do not know whether any specific loan has been rehabilitated or if any collections have been made after the default.

Canada Student Loans Program (CSLP) Administrative Data

The CSLP administrative data contain recipient-level longitudinal information from the 2003-2004 loan year to the 2015-2016 loan year.[11] The CSLP provides data on the loan application process, loan disbursement, and annual repayment amounts for each individual. Using the repayment records in these data, we calculate the rate of return R_i^e for each borrower based on equation (2.2). The CSLP also provides information on borrowers' institutions, fields of study, and basic demographics. It, therefore, provides a unique opportunity to study the distribution of realized returns conditional on the observable characteristics of borrowers to get *ex ante* expected returns for different individuals.

Postsecondary Student Information System (PSIS)

In addition to the CSLP, we also use administrative data from the PSIS, which covers all students who have enrolled in or graduated from any public college or university in Canada for each reporting year.[12] The available PSIS data cover from 2009-2010 to 2016-2017 for all provinces and territories, and as far back as 2005-2006 for the four Atlantic provinces (Newfoundland and Labrador, New Brunswick, Nova Scotia, and Price Edward Island). Since the reporting year in the PSIS largely overlaps with the loan year for the CSLP, these data can be linked to provide more information on individual characteristics, including immigration status and a richer classification for field of study.

Sample Selection

Since we are interested in the rate of return on student loans at the undergraduate level, we limit the sample to individuals who have received student loans for undergraduate studies. To ensure a relatively long period of observed payments, we focus on an early cohort that received its last undergraduate loan disbursement in the 2005 loan year, which we refer to as the 2005 cohort.

[11] A loan year starts in August 1st and ends in July 31th of the following calendar year.

[12] A reporting year starts from the day after the end of the institution's previous winter term, which is usually a date in April, May or June, and ends in one year from this start date. The PSIS covers students registered at the institution at any time during the reporting year.

The longest repayment period we observe is 11 years (from 2005-2006 to 2015-2016), which is one year longer than the duration of standard debt-based repayment in Canada. Some borrowers are observed for shorter repayment periods, because they were enrolled in undergraduate studies for another year or two after their last undergraduate loan disbursement or continued on to other levels of study (e.g., non-degree or master/PhD), borrowing more. The final loan consolidation year is later than 2005 for these borrowers, resulting in a shorter observed repayment period.

Because we want to link to the PSIS records for some of our analysis, we limit our sample to borrowers from the Atlantic provinces who attended public post-secondary institutions in those provinces.[13] We further limit our sample to individuals aged 18-30 when they received their last undergraduate loan.

In the CSLP repayment data, the accounting identity that the amount repaid plus the amount outstanding must equal the loan consolidation amount is not respected for many borrowers. In many cases, there is an easy explanation (confirmed through conversations with CSLP officials), which we address as appropriate.[14] In other cases, there is no easy explanation. Discrepancies can result from coding errors or changes in the reporting system over time and across service providers. Among those discrepancies that are not easily understood, we drop those for which the absolute discrepancy is at least 5% of the consolidation loan amount, excluding about 520 individuals. We include the other approximately 920 individuals with a positive absolute discrepancy of less than 5% to maintain sample representativeness and size. The final sample consists of around 5,690 individuals.[15] About 90% of our sample can be linked to PSIS records.

[13] The PSIS only covers public institutions and is only available for Atlantic provinces prior to 2009. We only restrict the sample to those that can be linked to the PSIS when necessary.

[14] For example, sometimes a borrower may have multiple loan consolidations, because he leaves and then returns to school. In principle, the loan consolidation amount should be cumulative, reflecting the total outstanding balance at the time. However, there are cases where the later loan consolidation amount clearly does not include outstanding amounts from previous consolidations. Yet, actual repayment is based on the total outstanding loan amount, which leads to a discrepancy between the repayment record and the reported last consolidation amount. In this case, we correct the inconsistency using the actual repayment record to calculate the loan amount.

[15] Numbers of observations are rounded to 10 as required by Statistics Canada due to confidentiality concerns.

Descriptive Statistics

Table 2.1 reports descriptive statistics for the main variables in our sample. Our goal is to predict student loan returns from the lender's perspective at the time of last undergraduate loan disbursement. Therefore, the variables discussed below reflect the state of the borrower at last undergraduate loan disbursement, unless otherwise specified.

Nearly two-thirds of the sample is female. The average age when they received their last undergraduate loan was 22 years-old, and most borrowers were either single independent (56%) or dependent (38%), while only 4% were married/common law and 2% were single parents. Permanent residents make up only 1% of our sample of borrowers, while the rest were Canadian citizens.

In terms of geographic distribution, about 30% were from New Brunswick, 30% from Newfoundland and Labrador, 40% from Nova Scotia, and 5% from Prince Edward Island. The location of the institution attended has a similar distribution. 17% of our sample went to other levels of study after undergraduate — 12% went to non-degree studies and 5% went to master/PhD studies (this is unknown to the lender at last undergraduate loan disbursement).

Only 4% of borrowers in our sample had enrolled in a college (i.e. non-baccalaureate granting institutions), while the rest attended universities for undergraduate studies. The vast majority of our sample (93%) attended a university ranked among the top 49 Canadian institutions by their reputation.[16]

The average amount of undergraduate loans was $15,900.[17] About one-third borrowed less than $10,000 at the undergraduate level. Another one-third borrowed more than $10,000 but less than $20,000. About 30% borrowed more than $20,000 but less than $30,000. Only about 6% borrowed more than $30,000. Since borrowers could attend other levels of study and continue to borrow after their undergraduate studies, the amount of their last consolidated loan could exceed the amount borrowed as an undergraduate. It is also possible that borrowers have repaid some of their loans in between studies. The average amount of the last consolidated loan is $17,600, roughly 10% more than the average undergraduate loan amount.

[16]See `https://www.macleans.ca/education/canadas-top-school-by-reputation-2020/` for the 2020 ranking list.

[17]Dollar values are rounded to 100 if greater than 1,000 or 10 if less than 1,000 as required by Statistics Canada.

Table 2.1: Descriptive Statistics

Variables	Mean	Standard deviation
Gender		
Female	0.64	
Male	0.36	
Immigration status		
Canadian citizen	0.99	
Permanent resident	0.01	
Dependency category		
Married/Common law	0.04	
Single parent	0.02	
Single independent	0.56	
Dependent	0.38	
Issue province/Home province		
New Brunswick	0.28	
Newfoundland and Labrador	0.28	
Nova Scotia	0.39	
Prince Edward Island	0.05	
Study province		
New Brunswick	0.29	
Newfoundland and Labrador	0.26	
Nova Scotia	0.41	
Prince Edward Island	0.04	
Last study level		
Non-degree	0.12	
Undergraduate	0.83	
Master/Phd	0.05	
Institution type		
College	0.04	
Ranked University	0.93	
Unranked University	0.03	
Age at last undergraduate loan disbursement	22.4	2.6
Undergraduate loan amount	15,900	9,200
< $10,000	0.33	
$10,000 - $19,999	0.32	
$20,000 - $29,999	0.29	
≥ $30,000	0.06	
Last consolidated loan amount	17,600	9,900

Both the CSLP and the PSIS provide information on borrowers' field of study, using different classifications. Table 2.2 lists the detailed categories and the distribution of borrowers across each classification.[18] As one can see, the classification system in PSIS is more detailed than that of the CSLP; although, there is considerable overlap. In CSLP, nearly half of all borrowers studied in Arts/Science, a very broad category. The PSIS uses the Classification of Instructional Programs (CIP) primary grouping, which distinguishes better between arts and sciences majors.[19] Based on this classification, we see that about 17% of borrowers had studied in humanities, 14% in social and behavioural sciences, and another 8% in physical and life sciences; only 2% had majored in visual and performing arts and communication technologies. Both classifications show that roughly 15% of borrowers had majored in each of business-related fields, education/community service fields, and health-related fields.

2.5 Returns on Student Loans in Canada

We calculate the rate of return for each borrower described in equation (2.2), using the discount rate of $d = 2.1\%$ to discount payments to the period when the last undergraduate loan is disbursed, 2005.[20] Our data suggest that the average individual-level rate of return is -4.9% across all borrowers. This does not distinguish between borrowers who borrowed \$1,000 and those who borrowed \$30,000; however, this distinction is important for the CSLP in calculating its total return on its loan portfolio. To assess the average return to the CSLP on the total amount lent to this cohort, we weight each observed return by the undergraduate loan amount. This weighted average return is -5.0%, indicating that, on average, the CSLP has lost 5 cents for every dollar it lent out to this cohort of students (from the Atlantic provinces).

We next discuss rates of return based on the last repayment status. We then discuss average returns for different subgroups of students based on their demographic characteristics, field of study, and post-secondary institution. Finally, we use multivariate regression methods

[18] Details on the areas of study included in each category can be found in Appendix A.1.

[19] In CIP, we separate the primary classification of law and social and behavioural sciences into (i) law and (ii) social and behavioral sciences.

[20] The Government's cost of borrowing is estimated to be 2.1% for the 2011-2012 loan year (Office of the Chief Actuary, 2012). It has not changed much over the last few years (see, for example, Office of the Chief Actuary (2019)).

Table 2.2: Distribution of Field of Study

Field of Study	Mean
CSLP categories	
Administration/Business	0.13
Agriculture	0.01
Arts/Science	0.49
Community Services/Education	0.16
Dentistry	0.01
Engineering/Technology	0.05
Health Sciences	0.13
Law	0.01
Medicine	0.01
PSIS categories	
Education	0.13
Visual and Performing Arts, and Communications Technologies	0.02
Humanities	0.17
Social and Behavioural Sciences	0.14
Business, Management and Public Administration	0.14
Physical and Life Sciences, and Technologies	0.08
Mathematics, Computer and Information Sciences	0.02
Architecture, Engineering and Related Technologies	0.05
Agriculture, Natural Resources and Conservation	0.02
Health and Related Fields	0.16
Law	0.01
Other	0.07

to estimate the importance of individual and institutional factors in determining the returns. These estimates are used to calculate predicted returns based on information observed by the government at the time student loans were disbursed. These predicted returns are critical for evaluating concerns about adverse selection and potential cream-skimming by private lenders.

2.5.1 Returns by Last Repayment Status

As discussed in Section 2.3, there are five terminal repayment statuses. Table 2.3 provides information about borrowing and repayment by last repayment status.

Table 2.3: Borrowing and Repayment by Last Observed Repayment Status

	Fraction	Avg. Amt. Borrowed	Avg. Amt. Owed	Std. Dev. Amt. owed	$\frac{L_{i,T_i}}{(1+d)^{T_i} L_{i,0}}$	Avg. P_{T_i}	Avg. Months in RAP
Paid in full	52%	15,300					
In repayment	24%	21,100	10,500	8,000	38%	2,400	
In RAP	6%	23,900	16,500	9,000	54%	870	
IR/RAP stage 1	3%	23,700	17,500	10,100	58%	990	39
RAP stage 2	3%	23,800	16,000	7,600	53%	770	91
In default	16%	17,200	17,700	10,800	93%	110	
Bankruptcy	2%	19,700	16,600	11,600	71%	540	

By the end of the sample period, 52% of borrowers had fully repaid their loans, 24% were in standard debt-based repayment, 6% were enrolled in RAP, (with half in stage 2), 16% had defaulted, and another 2% had declared bankruptcy. On average, those in RAP borrowed the most, while those who paid in full had borrowed the least. Those still in repayment owed, on average, about \$10,500 at the end of our sample period, while the average amount still owed among those enrolled in RAP, in default, or in bankruptcy ranged from \$16,500-17,700.

The share of debt still owed (in present value) at the end of our sample period, $(1+d)^{-T_i} L_{i,T_i}/L_{i,0}$, varies considerably across loan statuses and is informative about the potential losses in different cases. For example, those in default still owe 93% of their original loan amount. Since only 30% of this remaining balance is typically repaid, average losses for those who default are quite high. Losses are even greater for those declaring bankruptcy, since these borrowers still owed 71% when they entered bankruptcy and additional payments are the exception. Of course, bankruptcy itself is quite rare, so these losses (and our assumption of zero collections from those entering bankruptcy) have little impact on average returns for the overall population of borrowers.

Regarding the annual payment during the last sample year, P_{T_i}, those still in (standard) repayment were paying, on average, \$2,400. At that amount, it would take about five years to repay the average amount owed (\$10,500) among these borrowers. As noted earlier, we assume that those enrolled in RAP will continue to pay P_{T_i} until their time on RAP expires or their loan is fully repaid, whichever is earlier. Given the sizeable amounts still owed (\$16,500 on average), relatively low payment amounts (\$870, on average), and number of months already on RAP, the typical borrower on RAP 10–11 years after leaving school will see much of his student debt forgiven.

Table 2.4 reports the mean and standard deviation of returns by last repayment status (under our assumptions). The average return from those who paid in full is about 10%, much higher than the overall average, as one might expect. Under our assumptions, those still in repayment at the end of our sample period (11 years of repayment in most cases) generate the highest return, because they pay interest accrued at a higher rate (5.5%) than the discount rate (2.1%) for a longer period of time. Due to uncollected amounts, those in RAP, default, or bankruptcy have very negative returns, on average. While most borrowers who enter default do so early (as evidenced by the high share of debt still owed), there is much greater variation in the timing at which borrowers enter bankruptcy, resulting in considerable variation in the returns from those borrowers.[21] Variation in the timing of entry into RAP and income-based payment amounts lead to considerable variation in the returns for borrowers enrolled in RAP at the end of the sample period.

[21] Given that the average share of debt still owed (in present value) is 93% for defaulters, the average return would be about 30% lower for defaulters if we assume nothing is collected after default, compared to our baseline assumption.

Table 2.4: Return by Last Repayment Status

	Mean	Standard deviation
Paid in full	0.101	0.089
In repayment	0.166	0.111
In RAP	-0.474	0.365
In default	-0.597	0.154
Bankruptcy	-0.652	0.318

2.5.2 Average Returns by Borrower and Institutional Characteristics

As noted earlier, average returns for our full sample are negative at -4.9%. Table 2.5 reports average returns (and their standard errors) by demographic characteristics, field of study, and type of institution attended. On average, men provide a lower return than women, which is interesting in light of the much-discussed disadvantage women face in the labor market (Fortin, 2019). Borrowers who studied in New Brunswick and Prince Edward Island have higher returns, on average, than those who studied in Newfoundland and Labrador and Nova Scotia.

We observe large differences in returns depending on the year in study when borrowers received their last undergraduate loan. Those whose last undergraduate loan was disbursed in the first year of post-secondary school have a significantly lower average return (about 5 percentage points lower) than those who last borrowed in years 2 or 3, and 11 percentage points lower than those who last borrowed in years 4 and above. This is not surprising given that those who last borrowed in years 3 and lower are more likely to have dropped out without a degree.

Those who studied in colleges have a slightly higher return than those who studied in universities, but the difference is not statistically significant. As we see below, this does not simply reflect lower loan burdens. Borrowers who studied in unranked universities have a significantly lower return than those who studied in colleges or ranked universities.

Average returns differ substantially across fields of study. Focusing on the more detailed PSIS classification, we see that education majors (6.8%) and health majors (5.8%) generate modest positive returns, while those majoring in arts (-23.4%) and humanities (-19.0%) are

Table 2.5: Average Return by Individual Characteristics, Major, and Type of Institution Attended

Characteristics	Mean	Standard Error
Gender		
Female	-0.043	0.006
Male	-0.060	0.007
Study province		
New Brunswick	-0.030	0.008
Newfoundland and Labrador	-0.055	0.009
Nova Scotia	-0.062	0.007
Prince Edward Island	-0.012	0.021
Year in study		
1	-0.118	0.010
2-3	-0.073	0.008
4-6	-0.009	0.006
Institution type		
College	-0.035	0.019
University	-0.049	0.005
Ranked University	-0.047	0.005
Unranked University	-0.132	0.027
CSLP field of study		
Administration/Business	-0.005	0.011
Arts/Science	-0.137	0.007
Community Services/Education	0.050	0.009
Engineering/Technology	0.037	0.015
Health Sciences	0.065	0.009
PSIS field of study		
Education	0.068	0.010
Visual and Performing Arts, and Communications Technologies	-0.234	0.040
Humanities	-0.190	0.013
Social and Behavioural Sciences	-0.122	0.014
Business, Management and Public Administration	-0.027	0.012
Physical and Life Sciences, and Technologies	-0.061	0.016
Mathematics, Computer and Information Sciences	0.018	0.027
Architecture, Engineering and Related Technologies	0.041	0.015
Agriculture, Natural Resources and Conservation	-0.020	0.034
Health and Related Fields	0.058	0.009
Other	-0.131	0.018

25–30 percentage points lower, yielding sizeable losses. Conclusions are similar based on the CSLP field of study classification.

2.5.3 Predicting Returns on Student Loans

Table 2.5 shows that rates of return on student loans vary considerably across many dimensions. We now use standard multivariate regression models to identify which factors are most important determinants of those returns and to estimate predicted returns based on factors observable to the government at the time loans are disbursed.

Table 2.6 shows estimates from ordinary least squares (OLS) regression of realized returns on (i) a linear spline function of undergraduate loan amount with knots at \$10,000, \$20,000, and \$30,000, (ii) individual characteristics, including gender, dependency status, age, year in undergraduate study, and loan issue/home province, (iii) CSLP field of study indicators, and (iv) institutional indicators for ranked universities as well as an indicator for attendance at (two-year) colleges. (Schools are ordered by Maclean's 2020 ranking in the table.) Column (1) includes institution and college indicators only, while column (2) includes field of study indicators only. Column (3) controls for loan amount, individual characteristics, and field of study. Column (4) includes all variables. Since these specifications do not use any data from the PSIS, this table is based on the full sample of borrowers.

Since column (1) only includes institution indicators, the regression coefficients simply reflect differences in average returns between each ranked university, as well as all (two-year) colleges, and all unranked universities. By themselves, institution indicators explain about 2.5% of the variation in returns (see the R^2 statistic at the bottom of the table). As noted in Table 2.5, returns are generally higher for students attending ranked universities and (two-year) colleges compared to unranked universities. Students from only 4/12 ranked universities produce returns that are not significantly higher than unranked universities (at 5% significance level); however, students attending 7/12 ranked universities generate lower returns than students attending colleges (though not all differences are significant).

By comparing the estimates in column (1) with those of column (4), which includes student loan amounts, all demographic factors, and field of study indicators, we can examine whether

Table 2.6: Regression of Rate of Return on Characteristics (Full Sample)

	(1)	(2)	(3)	(4)
Loan amt (in \$10,000)			-0.055**	-0.053**
(Loan amt - 10,000) × $\mathbb{1}$(amt > \$10,000)			-0.002	-0.006
(Loan amt - 20,000) × $\mathbb{1}$(amt > \$20,000)			-0.011	-0.005
(Loan amt - 30,000) × $\mathbb{1}$(amt > \$30,000)			0.053	0.051
Gender				
Female			(base)	(base)
Male			-0.010	-0.012
Dependency category				
Married/Common law			-0.035	-0.034
Single parent			-0.219***	-0.211***
Single independent			-0.006	-0.005
Dependent			(base)	(base)
Age				
18			(base)	(base)
19			-0.025	-0.021
20			-0.038	-0.035
21			-0.013	-0.012
22			0.019	0.020
23			0.016	0.016
24			0.003	0.002
25			0.006	0.003
26			-0.008	-0.012
27			0.039	0.036
28			-0.016	-0.017
29			0.003	0.005
30			0.021	0.014
Year in study				
1			(base)	(base)
2			0.040**	0.036**
3			0.058***	0.058***
4			0.131***	0.128***
5			0.120***	0.120***
6			0.163***	0.172***
Issue province				
New Brunswick			(base)	(base)
Newfoundland and Labrador			-0.016	-0.001
Nova Scotia			-0.017	-0.032**

	(1)	(2)	(3)	(4)
Prince Edward Island			0.035*	0.002
Field of study (CSLP categories)				
Administration/Business		0.132***	0.112***	0.115***
Agriculture		0.069	0.054	0.065
Arts/Science		(base)	(base)	(base)
Community Services/Education		0.187***	0.207***	0.212***
Dentistry		0.226***	0.251***	0.215***
Engineering/Technology		0.174***	0.161***	0.153***
Health Sciences		0.202***	0.207***	0.192***
Law		0.231***	0.272***	0.249***
Medicine		0.164***	0.179***	0.161***
Institution				
Dalhousie University	0.150***			0.094***
Memorial University of Newfoundland	0.076***			0.015
Mount Allison University	0.076*			0.097**
St. Francis Xavier University	0.122***			0.095***
Acadia University	0.079**			0.047
University of New Brunswick	0.112***			0.041
Saint Mary's University	0.008			0.036
University of Prince Edward Island	0.113***			0.081**
Mount St. Vincent University	0.074**			0.046
Université de Moncton	0.138***			0.018
Cape Breton University	-0.069**			-0.028
St. Thomas University	0.021			0.018
Colleges	0.096***			0.112***
Unranked Universities	(base)			(base)
Constant	-0.132***	-0.137***	-0.106***	-0.143***
Number of observations	5,690	5,690	5,690	5,690
R^2	0.0253	0.0718	0.1120	0.1214

***: significant at 1%
**: significant at 5%
*: significant at 10%

differences in the rate of return across institutions are confounded by other systematic differences in the characteristics and choices of individuals attending these institutions. Comparing the estimated institution coefficients across these columns, we see that many of the institution differences become smaller (and statistically insignificant) in column (4), suggesting that some of the institutional differences in rates of return are due to differential selection into institutions: higher-return students tend to attend institutions associated with higher returns. Highlighting the importance of student selection, the variance of institution effects across all borrowers drops by 55% from column (1) to column (4).[22] After controlling for other characteristics, students attending a few of the highest ranked institutions from the Atlantic provinces (Dalhousie University, Mount Allison University, St. Francis Xavier University) still generate returns that are 9–10% higher than the typical unranked university. Interestingly, these returns are still slightly lower than the returns generated by two-year college students (11.2%). Because column (4) includes student debt levels, the high rates of return at colleges relative to most universities is not simply explained by their lower debt loads. College students generate a high return when compared with other similar university students who borrow the same amount.

The specification in column (2) includes field of study indicators only, showing significant differences in average returns across fields (relative to Arts/Science majors). While these estimates provide no new information relative to the average returns reported in Table 2.5, this specification facilitates a comparison with the estimates in column (4), which simultaneously controls for all other factors. Additionally, the R^2 in column (2) immediately reveals the importance of college major in predicting rates of return: differences across majors explain 7.2% of all variation in realized rates of return, nearly three times as much variation across institutions. Unlike with the institution indicators, estimated effects of college majors are quite stable after controlling for loan amounts and borrower characteristics (column (3)), as well as institution dummies (column (4)). The variance of college major effects declines by only 5% from column (1) to column (4), so differences in returns across majors are not driven by differences in the selection of students. Focusing on column (4), we see that nearly every major (except Agriculture) provides a significantly higher return than Arts/Science, with a few differences (law, dentistry, and community/education) greater than 20 percentage points. The high return majors are not particularly surprising, since these fields are associated with high post-school earnings in Canada (Frenette and Frank, 2016). It is notable, however, that the differences in returns are much greater than those across institutions, which are almost all less than 10 percentage points.

Columns (3) and (4) of Table 2.6 show that other borrower characteristics are also important determinants of returns on student loans. Not surprisingly, the amount of student debt has

[22] Specifically, when giving each borrower their institution effect in column (1) (i.e. their average institution return), the variance of these effects is 0.0029. Using the institution effects in column (4), the variance of these effects drops to 0.0013.

a negative effect on the estimated rate of return, all else equal. For every additional \$10,000 in loan amount, the return is expected to decrease by 5.3 percentage points. The insignificance of the additional spline coefficients indicates that additional loan amounts have a roughly linear effect on rates of return. Other individual characteristics are also important, including dependency status and year in undergraduate study. Students that are single parents have significantly lower returns than dependent students (by about 20 percentage points). Returns are similar across dependent, single independent, and married students. Year in study has important effects, with higher rates of return for those who have been in school longer. Those who receive their last undergraduate loan in the fourth year or higher generate returns that are 12-16% higher than those in the first year of post-secondary schooling. These differences likely reflect the fact that those further along in school when they receive their last student loan are more likely to graduate and receive their undergraduate degree. Returns differ little by gender, age, or home province. Altogether, student debt, demographic characteristics, field of study, and institution effects explain 12.1% of the variation in rates of return.

Table 2.7 reports similar regression results for our PSIS-linked sample, where we now include immigration status and the PSIS field of study grouping. Since the effects of student debt, demographic characteristics, and institutions are quite similar to those of Table 2.6, we do not discuss them again. One advantage of the PSIS is our ability to estimate returns for immigrants separately from citizens. We find 10% lower rates of returns for the very small fraction of immigrants in our sample. A more important advantage is the more detailed PSIS fields of study categorization, especially in regards to Arts/Science. Focusing on column (4), we see that most majors generate higher returns than humanities (the reference group). The highest return majors include law, education, health-related fields, and engineering-related fields, while the lowest include humanities, visual and performing arts, and communications technologies. The difference between majors is substantial with the gap between the highest (law) and lowest (visual arts) return majors more than 35 percentage points. As in Table 2.6, these differences are much greater than the differences across institutions. Choices about major are a much more important determinant of student loan repayment than the choice of institution. Altogether, the improved measure of major choice (and measure of immigration) in the PSIS means that returns are better predicted in Table 2.7, with the R^2 increasing to 0.139.

Table 2.7: Regression of Rate of Return on Characteristics (Linked Sample)

	(1)	(2)	(3)	(4)
Loan amt (in $10,000)			-0.075***	-0.074***
(Loan amt - 10,000) × $\mathbb{1}$(amt > $10,000)			0.025	0.022
(Loan amt - 20,000) × $\mathbb{1}$(amt > $20,000)			-0.021	-0.016
(Loan amt - 30,000) × $\mathbb{1}$(amt > $30,000)			0.038	0.034
Gender				
Female			(base)	(base)
Male			-0.021**	-0.021**
Immigration status				
Canadian citizen			(base)	(base)
Permanent resident			-0.100**	-0.102**
Dependency category				
Married/Common law			-0.050*	-0.050*
Single parent			-0.221***	-0.205***
Single independent			-0.004	-0.003
Dependent			(base)	(base)
Age				
18			(base)	(base)
19			-0.011	-0.009
20			-0.017	-0.014
21			0.000	0.001
22			0.023	0.025
23			0.026	0.026
24			0.016	0.015
25			0.025	0.021
26			0.006	0.004
27			0.052	0.048
28			-0.001	-0.004
29			0.016	0.018
30			0.050	0.043
Year in study				
1			(base)	(base)
2			0.038**	0.034*
3			0.056***	0.054***
4			0.136***	0.130***
5			0.121***	0.118***
6			0.168***	0.175***
Issue province				

New Brunswick			(base)	(base)
Newfoundland and Labrador			-0.023*	0.001
Nova Scotia			-0.016	-0.027*
Prince Edward Island			0.045**	0.039
Field of study (PSIS categories)				
Education		0.258***	0.281***	0.284***
Visual and Performing Arts, and Communications		-0.044	-0.058*	-0.053
Humanities		(base)	(base)	(base)
Social and Behavioural Sciences		0.068***	0.047***	0.043***
Business, Management and Public Administration		0.163***	0.149***	0.157***
Physical and Life Sciences, and Technologies		0.130***	0.127***	0.120***
Mathematics, Computer and Information Sciences		0.208***	0.195***	0.193***
Architecture, Engineering and Related Technologies		0.231***	0.216***	0.203***
Agriculture, Natural Resources and Conservation		0.170***	0.155***	0.172***
Health and Related Fields		0.248***	0.245***	0.235***
Law		0.293***	0.332***	0.308***
Other		0.059***	0.076***	0.067***
Institution				
Dalhousie University	0.146***			0.079***
Memorial University of Newfoundland	0.064**			-0.004
Mount Allison University	0.066			0.106**
St. Francis Xavier University	0.112***			0.071**
Acadia University	0.053			0.018
University of New Brunswick	0.101***			0.022
Saint Mary's University	-0.011			0.031
University of Prince Edward Island	0.107***			0.031
Mount St. Vincent University	0.057*			0.062*
Université de Moncton	0.130***			0.018
Cape Breton University	-0.087**			-0.063*
St. Thomas University	0.001			0.022
Colleges	0.077**			0.090**
Unranked Universities	(base)			(base)
Constant	-0.121***	-0.190***	-0.149***	-0.173***
Number of observations	5,130	5,130	5,130	5,130
R^2	0.0285	0.0863	0.1296	0.1394

***: significant at 1%

**: significant at 5%

*: significant at 10%

To further investigate the roles played by fields of study and institutions, we calculate the fraction of the total variance of predicted returns that can be explained by each. Table 2.8 reports the results from this decomposition exercise. Column (1) is based on the estimates from column (4) of Table 2.6, while column (2) is based on the analogous column (4) specification in Table 2.7. Looking at column (1) of Table 2.8, which uses the CSLP field of study categories, the total variance of predicted returns is 0.0137. Column (2) shows that the improved PSIS field of study categories yields a higher variance in predicted returns of 0.0157. In both cases, the field of study indicators explain about two-thirds of the variance in predicted returns.[23] By contrast, institution differences explain less than 10% of the variance. To further examine whether there is any difference between college and university students, we conduct a similar exercise for university students only. The results are shown in columns (3) and (4). The basic patterns remain the same with institution differences explaining slightly less of the variation in returns across borrowers.

Table 2.8: Variance of Predicted Return

	(1) CSLP Full	(2) PSIS Full	(3) CSLP University	(4) PSIS University
Total variance of predicted returns	0.0137	0.0157	0.0144	0.0164
Percent explained by field	61.3%	67.5%	61.1%	67.1%
Percent explained by institution	9.4%	8.9%	7.6%	6.7%

We now turn to a comparison of predicted and realized rates of return. Figure 2.1 plots the kernel density of realized and predicted returns from the most generous specification using the full sample (i.e. column (4) in Table 2.6).[24] Not surprisingly, realized returns are more disbursed than predicted returns. Both returns are bimodal; however, realized returns are more negatively skewed with a sizeable mode at returns around -70%, reflecting significant losses

[23]To calculate the variance of predicted returns explained by field (institution), we assign each borrower their estimated field (institution) effect from column (4) of Table 2.6 or 2.7 (or the analogous specification for university students only), then calculate the variance of these effects.

[24]We use the Epanechnikov kernel and the “optimal” bandwidth, which is the width that would minimize the mean integrated squared error.

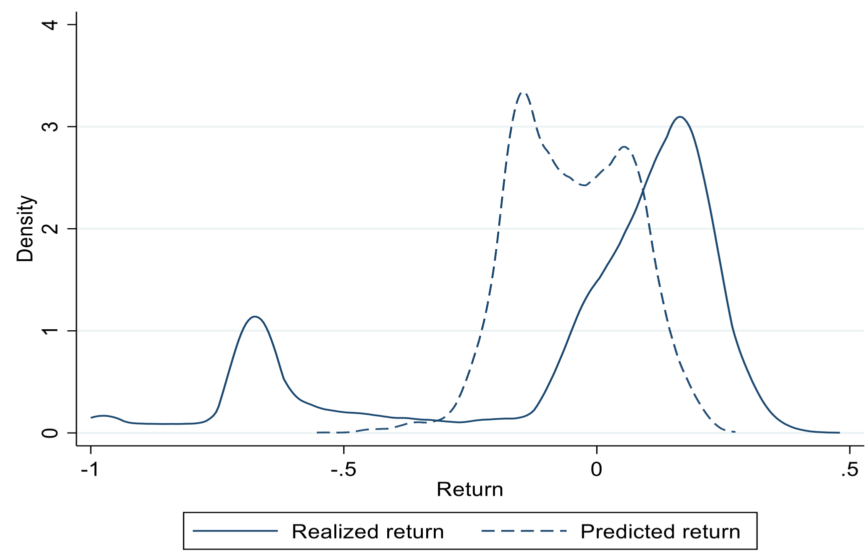


Figure 2.1: Realized vs. Predicted Return, Kernel Density

associated with bankruptcy, default, and (in some cases) RAP. The vast majority of borrowers generate a positive return, however. Interestingly, very few borrowers appear to generate the average return of -5%, highlighting that heterogeneity in returns is a central feature of the CSLP. While the very low return outcomes are difficult to predict, we do observe both a positive and negative mode for predicted returns with the negative mode peaking around -20%. Figure 2.2 plots the kernel density for college and university students separately. The general shapes for these densities are similar to their overall counterparts; however, the negative mode for predicted returns is much more pronounced for (two-year) college students despite a much less pronounced mode around -70% for their realized returns. Due to the much stronger negative skewness of realized returns, median realized returns are positive (9.0% for university students vs. 7.1% for college students), while median predicted returns are negative (-5.2% for university students vs. -6.1% for college students). Recall from Table 2.5 that average returns are also negative for both types of students, but less negative for college students (-4.9% for university students vs. -3.5% for college students).

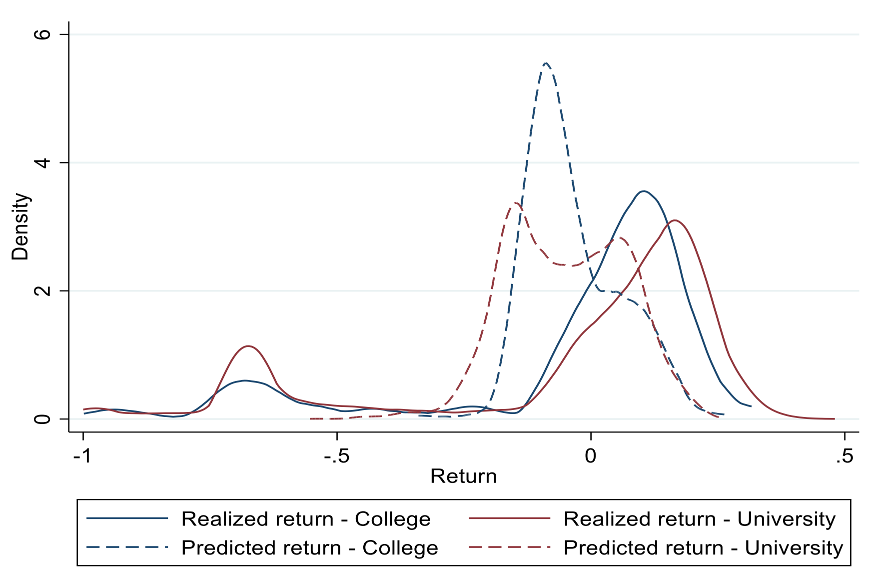


Figure 2.2: Realized vs. Predicted Return by Institution Type, Kernel Density

2.5.4 Internal Rate of Return to CSLP

Thus far, our analysis assumes a CSLP discount rate of d =2.1% in calculating rates of return, showing that the government loses about 5% on its loans (to Atlantic province undergraduates from the 2005 cohort). While our assumed discount rate is based on the officially reported cost of funds for CSLP, the true cost of funds is difficult to determine due to the aggregate risk inherent in loan portfolios for any given cohort (Lucas and Moore, 2010a). We can alternatively calculate an internal rate of return (i.e. the discount rate that would imply a zero average return) based on loan amounts and repayments for all borrowers in our sample. Doing so, we obtain an IRR for the 2005 cohort of Atlantic province undergraduates of 1.343%.[25] This implies that if the borrowing cost for the CSLP is less than 1.343%, the program would earn a positive return. For a higher cost of funds, CSLP would experience a loss.

2.5.5 Cross-Subsidization and Adverse Selection

The substantial heterogeneity in predicted returns — expected returns conditional on factors observable to the government and other potential lenders — suggests considerable *ex ante* cross-subsidization. Borrowers with high predicted returns effectively subsidize those with

[25] In calculating the IRR for the full cohort, we weight individuals by the undergraduate loan amount.

low predicted returns. This raises adverse selection concerns that private lenders could draw away the pool of borrowers with high returns, leaving the government with an even more negatively selected pool of borrowers and lower returns. We explore this issue by calculating how the average (weighted) return and IRR to the CSLP (for the 2005 cohort of Atlantic province undergraduates) would be impacted by cream-skimming by private lenders. In particular, we calculate the average weighted return and IRR when borrowers with predicted returns above different thresholds are excluded from the CSLP portfolio. Table 2.9 shows that the average weighted return would fall to -6.4% if borrowers with expected returns higher than 10% are excluded; the IRR would fall to 1.14%. If private lenders were able to siphon away all borrowers with predicted return above 1%, the average weighted return would drop considerably to -12.5%, and the IRR would drop to 0.26%. The potential for cream-skimming would seem to be a serious concern for the viability of the CSLP.

Table 2.9: Average Return and IRR If Excluding High Expected Return Borrowers

	Average weighted return	Internal rate of return
Exclude predicted return > 1%	-0.125	0.26%
Exclude predicted return > 3%	-0.112	0.45%
Exclude predicted return > 5%	-0.094	0.71%
Exclude predicted return > 10%	-0.064	1.14%

2.6 Conclusion

This paper studies rates of return on government student loans in Canada using novel administrative data covering 11 years of loan repayments for most borrowers. Using a discount rate of 2.1% (based on official cost of fund estimates), we show that the CSLP earns a -5% rate of return on the 2005 cohort of undergraduates from Atlantic provinces. Alternatively, CSLP earns an IRR of 1.34% on this portfolio of loans, suggesting any cost of funds above that rate would yield losses.

We exploit rich information on personal characteristics, loan amounts, field of study, and institution of attendance to explain differences in the rate of return across different types of borrowers. Our estimates imply substantial *ex ante* heterogeneity in the rate of return on student loans. While demographic characteristics like dependency status, immigration status, and year of study play some role in determining rates of return, most of the variation across borrowers is driven by differences in field of study, which explains roughly two-thirds of the variation in predicted returns. Differences in returns across students attending different institutions are also important but explain less than 10% of the variation in predicted returns. Altogether, the factors we (and the government) observe can explain as much as 14% of the variation in realized returns.

The differences in *ex ante* predicted returns can be sizeable with students in some majors generating a predicted return that is more than 30% higher than others, while students at some institutions possess returns that are 10% higher than other institutions. Using all available *ex ante* information, we find that predicted returns range from -25% to 25% for most borrowers, suggesting considerable cross-subsidization in an *ex ante* sense. High-return borrowers effectively subsidize low-return borrowers.

This pooling of high- and low-risk borrowers raises serious adverse selection concerns related to potential cream-skimming from private lenders. We show that if private lenders were to siphon away all borrowers with predicted returns above 10% (a fairly high profitability margin), the rate of return on the remaining portfolio would drop to -6.4% while the IRR would drop to 1.14%.

The CSLP and PSIS are rich sources of information on the experience and student loan borrowing of Canadian postsecondary students. There are many interesting and important questions that can be answered with these data. In this paper, we focus on only one early cohort, but a similar analysis can be applied to other cohorts to examine whether repayment patterns have changed over time, given documented recent trends in education costs, borrowing, and repayment behavior.[26] Given that debt is increasing, more borrowers are on RAP, and default rate is declining, it would be interesting to know how the return has changed for recent cohorts.

Given our interest in predicting *ex ante* returns and the potential for private cream-skimming, it would also be worthwhile exploring modern "big data" analysis techniques in search of better predictive models.

[26] According to Statistics Canada (https://doi.org/10.25318/3710015001-eng; https://doi.org/10.25318/3710000301-eng), average undergraduate tuition for full-time studies rose from $4,200 in 2005–2006 to $6,500 in 2019–2020. Average student debt owed at undergraduate graduation increased from $19,600 in 2005 to $23,000 in 2015 (Statistics Canada: https://doi.org/10.25318/3710003601-eng). The number of borrowers on RAP has almost doubled from its introduction in 2010–2011 to 2016–2017 (Human Resources and Skills Development Canada, 2012; Employment and Social Development Canada, 2018), while the three-year default rate has declined gradually from 17% in 2005–2006 to 9% in 2015–2016 (Employment and Social Development Canada, 2016, 2017, 2018).

Chapter 3

College Enrollment, Parental Transfers, and Student Loans

3.1 Introduction

Borrowing for post-secondary education in the U.S. has been increasing in the past few decades. According to Bricker et al. (2015), the total outstanding student loan balance was $1.27 trillion by the end of 2015, surpassing auto loans and credit cards to become the largest form of consumer debt outside of mortgages. At the same time, the student loan non-repayment rate remains high. The national three-year cohort default rate was 10.8% in 2015.[1]

Besides student loans, family transfers are also an important source of funding for higher education. There is evidence that parents not only pay for their children's tuition but also sometimes provide support when children have difficulty repaying student loans (Lochner et al., 2018). Many parents make post-college transfers to their children (Brown et al., 2012), especially when their post-school earnings are low. Parental transfers, including the provision of co-residency, can be a valuable insurance channel against labor market risk (Kaplan, 2012; McGarry, 2016).

Another recent economic trend that has important implications for higher education financing is the increasing labor market uncertainty in the U.S. over the past few decades (Moffitt and Gottschalk, 2012; Lochner and Shin, 2014). Negative shocks to college graduates over the first

[1] The three-year national cohort default rate is the percentage of federal student loan borrowers who enter repayment within the cohort fiscal year (begins on October 1st of a year and ends on September 30th of the following year) and default within the three-year period that begins on October 1st of the same fiscal year. Source: Department of Education, `https://www2.ed.gov/offices/OSFAP/defaultmanagement/cdr.html`.

few years in the labor market can lead to difficulties in repaying student loans. Increasing labor market risk, along with high debt levels, high default rates, and the need for parental assistance, indicates that youth may be inadequately insured through the current repayment scheme. This may affect their educational choices in the first place, especially those with limited parental resources.

Given these economic trends, some researchers and policy makers suggested that student loan repayment schemes should provide more insurance against labor market risk through more accessible income-based repayment (IBR) plans. Although some forms of assistance are currently available to borrowers, the participation rate is quite low. Under the current system, borrowers can choose between traditional repayment plans and IBR plans.[2] Borrowers may apply for loan deferment or forbearance that allows them to temporarily stop making payments if they experience financial hardship.[3] If borrowers do not apply for deferment or forbearance and fail to make payments, they are considered to be in default.[4] Though borrowers can suspend loan repayment during hardship through deferment or forbearance, many eligible borrowers in delinquency do not (Cunningham and Kienzl, 2011). Also, among Direct Loan borrowers in active repayment in 2014, only about 20% were enrolled in an IBR plan (Government Accountability Office, 2015). Dynarski and Kreisman (2013) point out that costs such as detailed paperwork associated with applying for forbearance, deferment, or IBR may prevent borrowers, especially those in distress, from using repayment assistance. Based on evidence from a randomized field experiment, Mueller and Yannelis (2019) find that the IBR plan take-up rate increased significantly when non-monetary costs were reduced. The take-up rate increased by 34 percentage points when borrowers received pre-populated applications for electronic signature. The current low take-up rate of repayment assistance has led to policy proposals that IBR schemes should be more accessible to provide insurance against labor market risk.

[2]The traditional repayment plans include the Standard Repayment Plan, Graduated Repayment Plan, and Extended Repayment Plan. Under traditional repayment plans, payments are based on debt, that is, the larger the debt is, the higher the payments are in each period. The IBR plans include the Revised Pay As You Earn Repayment Plan, Pay As You Earn Repayment Plan, Income-Based Repayment Plan, and Income-Contingent Repayment Plan.

[3]The main difference between deferment and forbearance is that borrowers may not be responsible for paying the interest that accrues on certain types of loans during the deferment period.

[4]For Federal Stafford Loans, borrowers are considered to be in default if they do not make their scheduled student loan payments for a period of at least 270 days.

There are, however, several concerns regarding the implementation of a more accessible IBR, including the effect on government revenue. Lochner et al. (2018) find that eliminating the non-monetary costs of applying for IBR assistance (i.e., moving to an automatic IBR scheme) could lead to a sizable revenue loss for the government. On the one hand, borrowers who are currently in forbearance or default may begin to repay under IBR, which could increase government revenue. However, given their low earnings, the repayment amounts would be small or even zero. On the other hand, borrowers with low post-college earnings who currently make full payments with parental help or personal savings may pay less under IBR, leading to revenue losses. The analysis from Lochner et al. (2018) suggests that the loss of revenue exceeds the gain, at least for the first five years of repayment.

Given the revenue loss under IBR, the government may increase the interest rates of student loans to balance the budget, raising the costs of post-secondary education for many who currently repay their loans in full. This may discourage some youth from attending college and result in undesirable distributional effects on welfare. The impact of IBR on education is unclear, however, given that more insurance could also lead to higher enrollment. Another concern about IBR is the moral hazard issue — IBR is essentially an income tax, which may discourage labor supply.

To study the effects of IBR and address these concerns, a structural approach is needed. In this paper, I develop a dynamic life-cycle model that allows endogenous transfers from parents to children, together with children's education, borrowing, repayment, and labor supply decisions. Labor market risks, including the uncertainty of getting a job and shocks to wages, are considered in the model to examine the effects on labor supply when more insurance is provided by IBR.

The model is estimated using the National Longitudinal Survey of Youth 1997 (NLSY97), which has detailed information about youth's family backgrounds, education, college financing, employment, and parental transfers. My estimated model suggests that students from low-income families borrow more and are more likely to default when IBR is not available. Also, low-income families are more likely to pay the sign-up cost and to enroll in IBR in counterfactual simulations, indicating that they need the insurance provided by IBR the most. IBR crowds out savings and parental transfers as it provides more insurance to borrowers. The response of the labor supply to IBR is negligible, so the problem of moral hazard is less likely to be a matter of concern. Under IBR with sign-up costs, students borrow more and repay less so that the interest rates of student loans must increase to keep the program self-sustaining. Yet even after interest rates increase, college enrollment is still higher than the baseline. When removing the sign-up cost while keeping the budget fixed, college enrollment increases further. This cost-free and budget-neutral IBR is welfare-improving, and welfare increases more for relatively poor and high-ability families.

In related literature, some studies conclude that optimal student loan repayment policy should be income-contingent (Gary-Bobo and Trannoy, 2014; Findeisen and Sachs, 2016; Lochner and Monge-Naranjo, 2016). A few papers model detailed repayment plans and quantify the effects of alternative student loan policies on college enrollment, borrowing behavior and default rates (Ionescu, 2009, 2011; Ji, 2018). Unfortunately, none of these studies examine the role of student loan repayment policies when considering endogenous parental transfers. Abbott et al. (2019) examine the equilibrium effects of alternative financial aid policies with endogenous parental transfers, but they do not model detailed student loan repayment plans, nor do they consider changes to the current debt-based repayment structure.

Lochner et al. (2018) find that parental support and personal savings substantially reduce student loan repayment problems in Canada. The authors also study the effects of removing the non-monetary costs of IBR when considering endogenous parental transfers in a theoretical context. However, they do not incorporate endogenous labor supply, college enrollment, or borrowing and they are unable to calculate the welfare implications of switching to universal IBR. Therefore, given the importance of parental support and increasing labor market risks, my paper contributes to the literature by incorporating endogenous parental transfers and labor market risks to empirically examine the effects of IBR.

This chapter proceeds as follows. Sections 3.2 and 3.3 provide details of the model. Section 3.4 discusses the data and descriptive statistics. Section 3.5 outlines the estimation strategy and results. Section 3.6 examines the effects of implementing IBR schemes. Section 3.7 concludes.

3.2 Model

3.2.1 Overview

The basic unit in the model is a family, which consists of a parent (p) and a youth (y). They make joint decisions with full commitment and one-sided altruism — only parents are altruistic towards their children. Since there is no strategic behavior by assumption, the problem is equivalent to the parent making all family decisions. The decision period begins when the youth graduates from high school. Every period non-negative transfers from the parent to the youth are endogenously decided within the family. Student loan borrowing and repayment decisions are modeled if the youth goes to college. Labor market risks faced by the youth, including the risk of getting a job offer and permanent and transitory wage shocks, are also incorporated.

Time is finite and discrete ($t = 0, 1, 2, ..., T$). A model period corresponds to one year. At $t = 0$, after knowing their idiosyncratic taste shocks for college, the family decides whether the youth goes to college and which type of college to attend (2-year or 4-year). If the youth goes to a 2-year or 4-year college, he is committed to 2 or 4 years of schooling before entering the post-school labor market, so dropout is not allowed. Students can finance education through government grants and loans, parental transfers, and working during college. After schooling is completed, shocks related to job offers, wages, and labor supply preferences are realized at the beginning of each period. The family makes decisions on youth labor supply, student loan repayment, parental transfers, and other assets. All shocks are assumed to be distributed independently over time and from each other.

3.2.2 Preferences

Parents have time-separable preferences, and they are potentially altruistic toward their children. The per-period utility of a parent derived from her own consumption is:

$$u_t^p = f^p(c_t^p) = \frac{(c_t^p)^{1-\gamma}}{1-\gamma},$$

where c_t^p is the parent's consumption. Parents' labor supply is assumed to be inelastic: they work full-time until the last period of the model.

Youth's preferences are also time-separable and are defined over consumption, labor supply, college enrollment, and student loan borrowing and repayment status. The per-period utility of a youth is:

$$u_t^y = \frac{(c_t^y)^{1-\gamma}}{1-\gamma} + v_t(h_t, s_t, \boldsymbol{\varepsilon}_t^v) + z_t(\boldsymbol{X}_t, \boldsymbol{\varepsilon}_0^z) + f^b(d_t, s_t) + f^d(R_t, D_t),$$

where c_t^y is the youth's consumption, $v_t(h_t, s_t, \boldsymbol{\varepsilon}_t^v)$ is the utility of working, $z_t(\boldsymbol{X}_t, \boldsymbol{\varepsilon}_0^z)$ is the non-pecuniary taste for attending college, $f^b(d_t, s_t)$ is the utility cost of borrowing from student loans during college, and $f^d(R_t, D_t)$ is the utility cost of defaulting.

Youth's labor supply is discrete: $h_t \in \{0, 0.5, 1\}$. $s_t \in \{NC, C, U\}$ represents youth's enrollment status at t: not in college, in 2-year college, or in 4-year college. Because schooling choice is a one-time choice at $t = 0$, the enrollment status s_t only depends on the initial schooling choice s_0 and time t:

$$s_t = \begin{cases} C, & \text{if} \quad s_0 = C \quad \text{and} \quad t \leq 2; \\ U, & \text{if} \quad s_0 = U \quad \text{and} \quad t \leq 4; \\ NC, & \text{otherwise.} \end{cases}$$

$v_t(h_t, s_t, \boldsymbol{\varepsilon}_t^v)$ depends on youth's work intensity, college enrollment, and stochastic shocks:

$$\begin{aligned} v_t(h_t, s_t, \boldsymbol{\varepsilon}_t^v) = {} & (\alpha_{v1} + \varepsilon_t^{v1}) \cdot \mathbb{1}(h_t = 0.5) + (\alpha_{v2} + \varepsilon_t^{v2}) \cdot \mathbb{1}(h_t = 1) \\ & + \alpha_{v3} \cdot \mathbb{1}(h_t = 0.5) \cdot \mathbb{1}(s_t = C) + \alpha_{v4} \cdot \mathbb{1}(h_t = 1) \cdot \mathbb{1}(s_t = C) \\ & + \alpha_{v5} \cdot \mathbb{1}(h_t = 0.5) \cdot \mathbb{1}(s_t = U) + \alpha_{v6} \cdot \mathbb{1}(h_t = 1) \cdot \mathbb{1}(s_t = U), \end{aligned} \tag{3.1}$$

where $\mathbb{1}(\cdot)$ is an indicator function, which equals 1 if the condition holds or 0 otherwise. Youth can work part-time ($h_t = 0.5$), full-time ($h_t = 1$), or stay unemployed ($h_t = 0$). Labor supply shocks ε_t^{vi} are assumed to be distributed normally: $\varepsilon_t^{vi} \sim N(0, \sigma_{vi}^2)$ for $i \in \{1, 2\}$.

The non-pecuniary taste for attending college $z_t(\boldsymbol{X}_t, \boldsymbol{\varepsilon}_0^z)$ depends on a vector of deterministic state variables including s_t, youth's AFQT which is a test score that measures cognitive ability, and parent's income I^p, i.e., $\boldsymbol{X}_t = (s_t, AFQT, I^p)$, and idiosyncratic shocks:

$$\begin{aligned} z_t(\boldsymbol{X}_t, \boldsymbol{\varepsilon}_0^z) = {} & (\alpha_{z1} + \alpha_{z2} AFQT + \alpha_{z3} \log(I^p) + \varepsilon_0^{z1}) \cdot \mathbb{1}(s_t = C) \\ & + (\alpha_{z4} + \alpha_{z5} AFQT + \alpha_{z6} \log(I^p) + \varepsilon_0^{z2}) \cdot \mathbb{1}(s_t = U). \end{aligned} \tag{3.2}$$

The preference shocks ε_0^{zi} allow for heterogeneity in college attendance among youth with similar observed characteristics. They are assumed to be distributed normally: $\varepsilon_0^{zi} \sim N(0, \sigma_{zi}^2)$ for $i \in \{1, 2\}$. It is crucial to have the non-pecuniary cost to capture the college entry patterns observed in the data.

The utility cost of borrowing from student loans $f^b(d_t, s_t)$ reflects the cost associated with applying for student loans and debt aversion. This cost depends on whether youth are borrowing from student loans and which type of college they are in:

$$f^b(d_t, s_t) = \alpha_{b1} \cdot \mathbb{1}(d_t > 0) \cdot \mathbb{1}(s_t = C) + \alpha_{b2} \cdot \mathbb{1}(d_t > 0) \cdot \mathbb{1}(s_t = U), \tag{3.3}$$

where d_t is the youth's student loan borrowing amount during period t.

The utility cost of defaulting $f^d(R_t, D_t)$ has the following functional form:

$$f^d(R_t, D_t) = \alpha_d \cdot \mathbb{1}(R_t = 0) \cdot \mathbb{1}(D_t > 0), \tag{3.4}$$

where R_t is the student loan repayment at t and D_t is the amount owed on student loan at t.

The family's expected life-time utility at $t = 0$ is given by:

$$\mathbb{E}_0 \Big\{ \sum_{t=0}^{T} \beta^t (u_t^p + \eta u_t^y) + \beta^{T+1} (V_{T+1}^p + \eta V_{T+1}^y) \Big\},$$

where $\eta > 0$ is the altruism factor. V_{T+1}^p is the parent's terminal value function (derived from the parent's consumption only), and V_{T+1}^y is the youth's terminal value function.

3.2.3 Budget Constraints

The family's budget constraints can be divided into two groups: one for the parent and one for the youth. For the youth, budget constraints are slightly different during school and after school.

During college, the youth's consumption satisfies:

$$c_t^y + x_{t+1}^y + (COA - gr) = (1 + r_f)x_t^y + w_t^c(h_t) + T_t + d_t. \tag{3.5}$$

Youth can work and receive wage income $w_t^c(h_t)$ during college, get parental transfers T_t, borrow from government student loans, and borrow or save on the private market at rate r_f. They have to pay a net price of $COA - gr$ for college, which is the cost of attending college (COA) minus the amount of grants and scholarships (gr).

There is an upper limit on how much can be borrowed from government student loans:

$$0 \leq d_t \leq \bar{d}_t. \tag{3.6}$$

Student loan borrowing limits $\bar{d}_t$ depend on which year of college they are in, as shown in Table 3.1, and are set by the government. The balance of the loan D_t evolves according to:

$$D_{t+1} = D_t + (1 + q \cdot r_d)d_t, \tag{3.7}$$

with $D_0 = 0$, where I assume there is no repayment during school. If parental income is less than the sample median ($q = 0$), students are eligible for subsidized loans, and interest does not accumulate during college. Otherwise ($q = 1$), interest starts to accumulate in college at rate r_d.

Parental transfers have to be non-negative:

$$T_t \geq 0. \tag{3.8}$$

The endogenous parental transfers make the model capable of studying the response of parental transfers to policy changes.

x_t^y reflects assets other than government student loans, and r_f is the interest rate on those assets. Youth can borrow from the private market, but there is a borrowing limit:

$$x_{t+1}^{y} \geq -\bar{x}_{t}^{y}. \tag{3.9}$$

For periods not in college, the youth's consumption satisfies:

$$c_{t}^{y} + x_{t+1}^{y} = (1 + r_f)x_{t}^{y} + w_t(h_t) + T_t - R_t. \tag{3.10}$$

The student loan repayment amount R_t is equal to 0 if youth do not go to college, go to college without taking student loans, or have already paid off their student loans. Debt balance evolves according to:

$$D_{t+1} = (1 + r_d)D_t - R_t. \tag{3.11}$$

Wage income $w_t(h_t)$ depends on labor supply choices. Details on repayment plans, borrowing limits, and wage functions are discussed in the following subsections.

For parents, the budget constraint remains the same in all periods:

$$c_{t}^{p} + T_t = I^P. \tag{3.12}$$

Parental income I^p is assumed to be the same over time, due to data availability. It is also assumed that parents cannot borrow or save. This assumption is not innocuous but necessary because otherwise, an additional continuous state variable would be added and increase the computation complexity. The key implication of this assumption is that it affects the timing of when parents make transfers. Parents are assumed to have a constant income every period, and they do not face any labor market uncertainties themselves. If $\beta(1 + r_f)$ is equal to 1, the only incentive for them to save is to insure against youth's labor market shocks and make transfers when youth are borrowing constrained (Altonji et al., 1997). If parents are assumed not to be able to save, they make transfers to youth even when youth are not borrowing constrained such that youth can save and insure against labor market risks by themselves. As a result, this assumption makes transfers more costly in terms of lifetime utility compared to the case in which parents can save. The alternative would be assuming that parents can save but youth cannot, which would overestimate the value of any alternative insurance channel (e.g., student loan repayment plan) since youth would be unable to self-insure through savings (Kaplan, 2012). As one of this paper's main goals is to evaluate the insurance provided by alternative repayment plans, it is crucial to assume that youth can save.

3.2.4 College Costs

The cost of attending college COA includes tuition, fees, and room and board. This cost depends on the state of residence and type of college (2-year or 4-year). Grants gr depend on AFQT and parental income for 4-year college students:

$$gr^u = \alpha^u_{g1} + \alpha^u_{g2} \cdot AFQT + \alpha^u_{g3} I^p. \tag{3.13}$$

Because grants are not highly correlated with AFQT for 2-year college students in my data, I assume that grants for 2-year college students depend only on parental income:

$$gr^c = \alpha^c_{g1} + \alpha^c_{g2} I^p. \tag{3.14}$$

3.2.5 Labor Market

Because labor market risks are crucial to consider when evaluating the insurance mechanism, the model considers the labor market risks faced by youth, including wage shocks and uncertainty on getting a job. When youth are in college, it is assumed that jobs are always available. However, the wage is stochastic:

$$w^c_t(h_t) = \exp(\bar{w}^c + \varepsilon^c_t)\left[\mathbb{1}(h_t = 1) + \alpha^{col}_p \cdot \mathbb{1}(h_t = 0.5)\right], \tag{3.15}$$

where $\varepsilon^c_t \sim N(0, \sigma^2_c)$. If youth work part-time, their wages will be a fraction (α^{col}_p) of full-time wages.

For a high school graduate or college graduate, in each period, the youth receives a job offer with some probability, which depends on his human capital and whether he worked in the last period:

$$\Pr(j_t = 1) = \Phi(\alpha_{j1} + \alpha_{j2}\Psi_t + \alpha_{j3}\mathbb{1}(h_{t-1} = 0)), \tag{3.16}$$

where Φ is the standard normal cumulative distribution function. j_t is a dummy variable indicating whether the youth gets a job offer or not: $j_t \in \{0, 1\}$. If the youth receives no offer, he becomes unemployed and $h_t = 0$. If he receives a job offer, the wage is also known to him, and he can either accept it ($h_t > 0$) or reject it and remain unemployed ($h_t = 0$).[5] The wage function is:

$$\begin{aligned} w_t(h_t) = w_t^f\Big[&\mathbb{1}(h_t = 1) + \mathbb{1}(h_t = 0.5) \cdot \Big(\alpha_p^{nc} \cdot \mathbb{1}(s_0 = NC) \\ &+ \alpha_p^c \cdot \mathbb{1}(s_0 = C) + \alpha_p^u \cdot \mathbb{1}(s_0 = U)\Big)\Big] + b \cdot \mathbb{1}(h_t = 0). \end{aligned}$$

w_t^f is the full-time wage rate. If youth work part-time, their wages will be a proportion of full-time wages. The ratio is α_p^{nc}, α_p^c, and α_p^u for high school graduates, 2-year college graduates, and 4-year college graduates, respectively. If youth do not work, they get unemployment benefit b.

The full-time wage rate w_t^f is a function of human capital Ψ_t and stochastic shocks:

$$w_t^f = \exp(\Psi_t + \nu_t + \varepsilon_t^w). \tag{3.17}$$

ν_t is the permanent shock which evolves according to $\nu_{t+1} = \nu_t + \xi_t$ where ξ_t is distributed normally: $\xi_t \sim N(0, \sigma_\xi^2)$. ε_t^w is the transitory wage shock and is also distributed normally: $\varepsilon_t^w \sim N(0, \sigma_w^2)$.

Human capital, Ψ_t, is a function of youth's AFQT, education, and work experience:

$$\Psi_t = \alpha_{\psi1} + \alpha_{\psi2} \cdot AFQT + \alpha_{\psi3} \cdot \mathbb{1}(s_0 = C) + \alpha_{\psi4} \cdot \mathbb{1}(s_0 = U) + \alpha_{\psi5} \cdot H_t, \tag{3.18}$$

where work experience at period t, H_t, sums over the experience in all previous working years: $H_t = H_{t-1} + h_{t-1}$ with $H_0 = 0$.

[5] All youth are assumed to be not working at $t = 0$, i.e., $h_0 = 0$.

3.2.6 Student Loans

Student Loan Borrowing Limits

There are limits on the amount of student loans that youth can borrow in college, which are set according to the Federal Family Education Loan (FFEL) Program. There are two main types of loans under the FFEL program for students: subsidized and unsubsidized Stafford loans.[6] The government covers the interest on subsidized loans when borrowers are in college, but interest on unsubsidized loans accumulates when students are in college. The loan limits for the Stafford loan program were constant in nominal terms from 1993 to 2007 (Table 3.1).[7]

Table 3.1: Stafford Loans Limits for Dependent Students

	Subsidized	Unsubsidized	Total
First year	2,625	2,625	2,625
Second year	3,500	3,500	3,500
Third year	5,500	5,500	5,500
Fourth year	5,500	5,500	5,500

Universities use a formula to calculate each student's expected family contribution (EFC) and use that to determine whether the student is eligible for subsidized loans. Students who have a cost of schooling greater than their EFC are eligible for subsidized loans. The EFC is calculated from parental income and assets, student's income and assets, and the number of other children from the family attending college. For simplicity, I assume that students whose parents' income is below the median are eligible for subsidized loans.

[6] Another type of loan under the FFEL program is the Parent Loan for Undergraduate Students (PLUS), but it is mainly for parents.

[7] The limits are different for dependent and independent students. Most traditional college students are dependent students, even if they are paying their own way through college or no longer have a relationship with their parents. To simplify the problem, I assume all youth are dependent students and eligible for the limits for dependent students as shown in Table 3.1.

Student Loan Repayment

In the baseline model, borrowers have the option of paying the standard debt-based payment amount or defaulting: $R_t \in \{R_t^d, R_t^{def}\}$. The debt-based payment amount R_t^d is equal to the debt balance D_t amortized over the remaining repayment period (T_t^d) at rate r_d:

$$R_t^d = D_t \left[\frac{r_d(1+r_d)^{T_t^d}}{(1+r_d)^{T_t^d} - 1} \right].$$

If a borrower chooses to default, he does not make any payment during that period ($R_t^{def} = 0$), but there is a utility cost ($f^d(R_t, D_t)$ defined in Equation (3.4)) to represent the punishments associated with default such as wage garnishment and exclusion from borrowing in the private market. The remaining repayment period T_t^d evolves according to: $T_{t+1}^d = \max\{0, T_t^d - 1\}$, regardless of which repayment plan is chosen at t.

Given the high default rate, it is crucial to allow borrowers to default in the model. Default also offers borrowers the option of delaying payments when they have difficulty repaying, although it may be costly due to the utility cost. Theoretically, those with a higher debt-to-income ratio and limited resources, including parental transfers and savings, are more likely to default because they have no other better options. The introduction of IBR could be particularly beneficial to this group of borrowers, as IBR could be a better insurance option with lower costs. As a result, the default rate may be lower after IBR is introduced. To examine the empirical effects, in Section 3.6, I introduce a stylized IBR plan as the third option for borrowers with more details provided in that section.

3.2.7 Private Financial Market

Youth can also borrow and save on the private financial market. They face the following borrowing constraint: $x_{t+1}^y \geq -\bar{x}_t^y$. The borrowing limit $\bar{x}_t^y$ is a function of human capital and age at t:

$$\bar{x}_t^y = \exp\{\alpha_{x1}^y + \alpha_{x2}^y \Psi_t + \alpha_{x3}^y age_t\}. \tag{3.19}$$

3.2.8 Terminal Value Function

Because the model focuses on college entrance, early labor market outcomes, and student loan repayment, I do not extend the model horizon until the end of the youth's life cycle. Instead, I assume T is earlier and use terminal value functions to represent utility afterwards. Assumptions for the terminal value functions are similar to those in Kaplan (2012). I assume that (i) the parent's life ends at $t = T$, and (ii) after $t = T$ all youth work full-time with no uncertainty about their wages. Given these assumptions, the terminal value functions have a closed-form solution.[8]

Because parents' life ends at T, $V^p_{T+1} = 0$. I assume that youth finish high school and are 18 years old at $t = 1$. I solve the model for $T = 25$ periods, which is long enough to cover college enrollment and student loan repayment. Youth work full-time after T until age 68.[9] For youth, the state variables at $T+1$ are human capital Ψ_{T+1} and assets x^y_{T+1}. From $T+1$ to the end of the life cycle, youth receive a constant stream of income every year: $w_{T+1} = \exp(\Psi_{T+1})$. Assuming that they can smooth consumption, the terminal value function for youth is given by:

$$V^y_{T+1} = \frac{(\alpha^y_{T1} x^y_{T+1} + \alpha^y_{T2} w_{T+1})^{(1-\gamma)}}{1-\gamma},$$

where $\alpha^y_{T1} = \left(\frac{1-\beta^{T_y}}{1-\beta}\right)^{\frac{1}{1-\gamma}} \cdot \frac{r_f}{1-(\frac{1}{1+r_f})^{T_y}}$ and $\alpha^y_{T2} = \left(\frac{1-\beta^{T_y}}{1-\beta}\right)^{\frac{1}{1-\gamma}}$. The number of years between the last period of the model and the end of life is $T_y = 26$.

3.3 Model in Recursive Form

This section describes the problem in the form of Bellman equations. Each subsection lists the state variables and choices of the family at each stage.

[8] An alternative approach would be to parameterize the terminal value function as a function of the state variables and estimate the parameters along with other structural parameters. However, to identify these parameters, it is necessary to have data on state variables in later periods (especially youth's assets). The NLSY97 does not have this information due to the relatively short life span it has covered.

[9] I assume that life ends at age 68. Because the average age that mothers gave birth is 26 in the NLSY97, the model covers parents from ages 44 to 68. This assumption makes the length of life the same for youth and parents.

3.3.1 College Choices

The family's expected value function for state $\Omega_0 = (AFQT, State, I^p, x_0^y)$ at $t = 0$ is:

$$\mathcal{V}_0(\Omega_0) = \mathbb{E}[V_0(\Omega_0, \boldsymbol{\varepsilon}_0^z)]. \tag{3.20}$$

The expectation is taken over the taste shocks for college $\boldsymbol{\varepsilon}_0^z$. After the taste shocks are realized, the family makes the college attendance choice s_0. Value function $V_0(\Omega_0, \boldsymbol{\varepsilon}_0^z)$ is the maximum of the expected value functions of the three educational choices:

$$V_0(\Omega_0, \boldsymbol{\varepsilon}_0^z) = \max\{\mathbb{E}[V_1(\Omega_1|s_0 = NC)], \mathbb{E}[V_1(\Omega_1|s_0 = C)], \mathbb{E}[V_1(\Omega_1|s_0 = U)]\},$$

where $\mathbb{E}[V_1(\Omega_1|s_0 = NC)]$, $\mathbb{E}[V_1(\Omega_1|s_0 = C)]$, and $\mathbb{E}[V_1(\Omega_1|s_0 = U)]$ are the expected value of not attending college, attending 2-year college, and attending 4-year college, respectively. Details about the state variables, choices, and shocks in each case are discussed below.

3.3.2 In College

If the youth is attending college ($s_0 \in \{C, U\}$), the family's value function is:

$$V_t(\Omega_t) = \max_{\theta_t} \{u_t^p + \eta u_t^y + \beta \mathbb{E}\left[V_{t+1}(\Omega_{t+1})|\Omega_t, \theta_t\right]\}, \tag{3.21}$$

subject to (3.5), (3.6), (3.7), (3.8), (3.9), and (3.12).

The state variables at this stage are: $\Omega_t = (I^p, \Psi_t, x_t^y, D_t, COA{-}gr, s_0, \boldsymbol{\varepsilon}_0^z, \boldsymbol{\varepsilon}_t^v, \varepsilon_t^c)$. Choices include parental transfers, labor supply, assets, and student loan borrowing, i.e., $\theta_t = (T_t, h_t, x_{t+1}^y, d_t)$. The expectation in equation (3.21) is taken over labor supply preference shocks $\boldsymbol{\varepsilon}_t^v$ and wage shocks in college ε_t^c, except in the last period of school where V_{t+1} is the value function after school, which is defined in the next subsection.

3.3.3 Post-College/No College

After the youth graduates from college with student loans, the family's value function is:

$$V_t(\Omega_t) = \max_{\theta_t} \left\{ u_t^p + \eta u_t^y + \beta \mathbb{E}\left[V_{t+1}(\Omega_{t+1}) | \Omega_t, \theta_t\right] \right\}, \tag{3.22}$$

subject to (3.8), (3.9), (3.10), (3.11), (3.12), and $h_t = 0$ if $j_t = 0$.

The state variables at this stage are: $\Omega_t = (I^p, \Psi_t, x_t^y, h_{t-1}, D_t, T_t^d, s_0, \nu_{t-1}, j_t, \boldsymbol{\varepsilon}_t)$. The vector of shocks $\boldsymbol{\varepsilon}_t$ includes labor supply preference shocks and wage shocks (permanent and transitory): $\boldsymbol{\varepsilon}_t = (\boldsymbol{\varepsilon}_t^y, \xi_t, \varepsilon_t^w)$. Choices include parental transfers, labor supply, assets, and student loan repayment, i.e., $\theta_t = (T_t, h_t, x_{t+1}^y, R_t)$. If the youth does not get a job offer ($j_t = 0$), there is no labor supply choice: $h_t = 0$. The expectation in equation (3.22) is taken over labor supply preference shocks, wage shocks, and the job offer realization.

If the child does not go to college ($s_0 = NC$), goes to college but does not borrow from student loans, or pays off student loans, $D_t = 0$ and $T_t^d = 0$, so $R_t = 0$ is chosen.

3.4 Data and Summary Statistics

The data used in this paper are from the National Longitudinal Survey of Youth 1997 (NLSY97), which is a longitudinal survey that follows a nationally representative sample of 8,984 youth who were 12 to 16 years old as of December 31, 1996. The cohort was first interviewed in 1997 and was followed annually until 2011. Since 2011, the interviews have been conducted biennially. The sample I use covers survey years 1997–2011 and 2013. The NLSY97 has detailed information about youth's family backgrounds, education, college financing, employment, and parental transfers.

For the analysis in this paper, I use the cross-sectional sample, which is representative of people living in the U.S. during the initial survey. I restrict the sample only to males whose highest education is high school, 2-year college, or 4-year college (the definition of education groups is explained in detail below). Males who have ever served in the military are excluded, as are those with missing data on key variables such as the Armed Forces Qualification Test (AFQT), parental income, or state of residence at high school graduation. The final sample size is 1,069. Table 3.2 reports the summary statistics for my sample. Monetary values, including assets, income, and transfers, are all adjusted by the Consumer Price Index (CPI) using 2004 as the base year. I use a simulated minimum-distance method in estimating the model. Key moments include statistics on education distribution, labor supply, student loans, wages, parental transfers, asset distribution, and default rates. Because NLSY97 does not cover certain aspects, such as student loan repayment and defaults, some moments come from external data sources. Details of the sample and the construction of key variables and moments are explained as follows.

Table 3.2: Selected Statistics by Education

	All	HS	2-year	4-year
Sample size	1,069	400	229	440
Average AFQT	52.76	34.33	48.57	71.70
Average parental income*	64,620	46,384	59,396	83,916
Average youth assets at 18*	5,444	4,870	4,100	6,666
% with parental monetary transfers at 18	48.74	37.25	49.34	58.86
Average parental monetary transfers at 18 if >0*	2,856	952	1,888	4,373
% live with parents at 18	92.80	89.26	92.58	96.10
% with parental monetary transfers at 25	22.08	12.75	25.33	28.86
Average parental monetary transfers at 25 if >0*	1,502	977	1,080	1,906
% live with parents at 25	28.06	28.25	36.24	23.64
% with college loans			37.55	64.32
Average annual college loans if >0*			1,517	2,134
Average annual college grants if >0*			2,687	6,285

* Measured in 2004 dollar

3.4.1 College Enrollment

There are three education categories in the model: high school graduates, 2-year college graduates, and 4-year college graduates. In my sample, high school graduates are those whose highest degree is high school and never enrolled in college. 2-year college graduates are those whose highest degree is an associate's degree. Also included in the 2-year graduate category are high school graduates who enrolled in 2-year college for no less than 24 months.[10] 4-year college graduates are those whose highest degree is a bachelor's degree.[11] As shown in Table 3.2, high school, 2-year college, and 4-year college graduates account for 37%, 21%, and 41% of the sample, respectively. The average cognitive ability measured by AFQT is increasing in education.

Two issues that are not considered in the model are delayed post-secondary entry and taking additional years to graduate from a two- or four-year college. I do not model the exact timing that youth enter college. As long as youth enter college at some point in time, they are included when calculating the relevant moments. In my sample, only around 10% of college graduates delay by one year or more and 5% delay by two years or more.

Some two- or four-year college graduates stay enrolled in college for longer than two or four years; however, I do not model extended stays in college. Moments calculated by year after school (e.g., labor supply, log wages, and parental transfers) refer to the year after graduation from (or last year of) college. The extended stay in college also affects the calculation of moments related to student loans and grants, which is explained in detail in the next subsection.

3.4.2 College Loans and Grants

The NLSY97 asks questions about each college and term in which a youth was enrolled. Students were asked about the total amount of loans from the government or other sources for each school attended and each term since the date of the last interview. I sum the number of loans borrowed when they were in college, then divide it by the months they spent in college and

[10] Some 2-year college graduates attended 4-year college. Those who enrolled in 4-year college for more than 24 months are excluded from the 2-year graduate category.

[11] Those who enrolled in 4-year college for more than 24 months but did not get a bachelor's degree are dropped because they are different from a 2-year college or 4-year college graduate defined here in terms of cognitive ability, parental income, and labor market outcomes. This excludes about 160 individuals (15% of the sample).

multiply by 12 to get an average annual student loan amount.[12] This calculation is done separately for 2-year and 4-year colleges, if students attended both. Since this amount may include private student loans, which are not modeled, I calculate the average annual amount of loans (if positive) that are capped by the annual government limits. According to Table 3.2, 38% of the 2-year college graduates take out student loans when enrolled in a 2-year college, and the average annual amount borrowed, if positive, is $1,517. 64% of the 4-year college graduates take out student loans, and the average annual amount borrowed, if positive, is $2,134.

Youth were also asked about the total amount of grants for each school attended and each term since the date of the last interview. I calculate annual grants using the same approach (i.e., I sum the grants when they were in college, then divide it by the months they spent in college and multiply by 12), separately for 2-year and 4-year colleges. The statistics on grants are reported in Table 3.2. The average annual grants, if positive, are $2,687 for 2-year college graduates and $6,285 for 4-year college graduates.

3.4.3 Parental Income and Parental Transfers

In the survey year 1997, parents reported gross household income in the past year if the youth was not independent.[13] This gross household income variable is used to construct parental income in 1996. In addition, parents reported their own income and their spouse's or partner's income in the past year from survey years 1998 to 2001. I sum the income of the parent and his or her spouse or partner and use that as the parental income of a youth. Therefore, there are at most five years (1996–2000) of observations of parents' incomes. I take the average income of the parents over the years available and use it as the parental income of the youth. Table 3.2 shows that average parental income is increasing in children's education.

There are two forms of parental transfers in the data: monetary and co-residency. Monetary transfers from parents are reported in both the income and college sections of the data. In the income section, each youth was asked whether he or she received money from parents and

[12]The NLSY97 provides monthly enrollment status. For a typical college student who enrolled in the academic year but took a break during the summer, the enrollment status would still be "enrolled" for the summer months. Therefore, the number of months he spent at college would include those summer months.

[13]NLSY97 youth were considered independent if they have had a child, were enrolled in a 4-year college, were no longer enrolled in school, were not living with any parents or parent-figures, or had ever been married or were in a marriage-like relationship at the time of the survey.

how much was received during the previous year for survey years 1997–2003. Since survey year 2004, the exact transfer amount has not been asked, and only a categorical amount is available.[14] The midpoint for each category is taken as the amount of transfers.[15]

In the college section, each youth was asked about whether he got aid from family or friends and the total amount of gifts or loans from parents and other relatives for each school and each term attended since the date of the last interview. I sum the total gifts or loans from parents to construct yearly parental transfers in college. I use the sum of the transfers reported in the income section and in the college section as the annual monetary transfers from parents.

Another form of parental transfer is co-residency, which is an important form of insurance and support that parents provide to children (Kaplan, 2012). Youth were asked about the relationship of people who live in the household to the youth as of the survey date. If students were living with parent-figures (including biological, step, adoptive, and foster parents), they are defined as co-resident with parents. Co-residency is not modeled explicitly in the model but is assigned a monetary value and included in total parental transfers.[16]

Table 3.2 shows that 49% of the sample receive monetary transfers from parents at age 18, which declines to 22% at age 25. Parents also provide room and board for their children. At age 18, 93% of the youth live with their parents, while 28% of them live with parents at age 25.

3.4.4 Youth Assets

In the first three survey rounds, if students were age 18 or if they met one of the other independence criteria, they were asked questions about various kinds of assets. Using this information, I construct the net worth of youth at age 18. Net worth includes housing and property values, automobiles, checking and savings accounts, bonds, stocks, life insurance, pension value, business wealth, student loan debt, and categories for other assets and debts.[17]

From round 4 onward, in the first interview after they turned ages 20, 25, 30, and 35,

[14]The transfer categories are: 1–500; 501–1,000; 1,001–2,500; 2,501–5,000; 5,001–7,500; 7,501–10,000; and more than 10,000.

[15]For the category that "more than 10,000", the transfer amount is assumed to be 10,001.

[16]The value of co-residency is assumed to be $7,800 per year ($650 per month), as reflected in Johnson (2013) and Kaplan (2012).

[17]The measure of assets is the same as in Johnson (2013).

respondents were again asked questions about their assets. NLSY97 creates the net worth of the respondents at ages 20, 25, and 30, which I use as the moments to be matched.

3.4.5 Youth Labor Supply

To construct moments related to the labor supply, I use the weekly employment data to determine whether a respondent was working and whether the employment was part-time or full-time during the year. NLSY97 reports the total number of hours worked by a respondent at any job during a given week, starting from the week when the respondent turned 14. I calculate the average weekly working hours by summing the working hours each year and dividing it by the number of weeks covered in that year. If the average weekly hours are fewer than 10, the respondent is categorized as not working. If the average weekly hours are no less than 10 but less than 30 hours, the respondent is categorized as working part-time. If the average weekly hours are equal to or more than 30 hours, the respondent is categorized as working full-time.

3.4.6 Geographic Information

In the model, tuition varies by the state of residence of the youth, which is also one of the initial conditions. NLSY97 has geographic information on which state the respondent lived in as of the survey year. I use the state respondents lived in when they received their high school degrees. If that is missing, I use the state they were in at age 18. I follow Johnson (2013) in grouping states into four categories by their average tuition levels. I use data from the Digest of Education Statistics on average public and private tuition by state between 1998 and 2004, as 95% of the sample received their high school degrees during that period. I sort the states based on average tuition weighted by enrollment and then divide the states into four categories so each category has approximately the same total population.[18] Next, I calculate the average tuition for 2- and 4-year colleges within each category. Details on the grouping of states and the average tuition are in Appendix B.1.

[18]The population of each state is from the 2001 intercensal data.

3.4.7 External Moments

Two sets of external moments are used: tuition elasticities and the default rate. To identify the variances of college preference shocks, I fit four external moments, which are the education elasticities of tuition — changes in 2- or 4-year college enrollment when 2- or 4-year tuition increases by \$1,000, including both self- and cross-elasticities. The elasticity moments are from Kane (1995), with tuition changes adjusted by CPI. The numbers are close to the estimates found in previous literature (Deming and Dynarski, 2010).

In that NLSY97 does not have information on student loan repayment, I fit one additional external moment, which is the average two-year cohort default rate for undergraduate borrowers between 1997 and 2011 from Yannelis (2015).[19]

3.5 Estimation

3.5.1 Estimation Strategy

Some parameters are fixed exogenously, while the remaining parameters are estimated using simulated minimum-distance.

Externally Calibrated Parameters

Table 3.3 shows the values of parameters that are exogenously determined. The risk aversion parameter γ is set to 2, and the annual discount factor β is 0.97, which follow standard values in the literature. The interest rate in the private market is 5.9% for borrowing and 0.9% for saving.[20] For subsidized student loans, the interest rate is zero when borrowers are enrolled in school. For unsubsidized loans, the interest rate is 2.29% while borrowers are in school. The

[19]The two-year default rate is the fraction of graduated borrowers who default within two years from the date of entering repayment.

[20]Source: Johnson (2013). The borrowing rate is the average prime rate from 2001–2007 minus inflation plus a 2% risk premium. The savings rate is the average real interest rate on one-year U.S. government bonds from 2001–2007.

interest rate is 3.7% for both subsidized and unsubsidized loans after borrowers leave college.[21] The student loan repayment parameters are based on current student loan policy. The standard debt-based repayment period is 10 years.

Table 3.3: Exogenously Given Parameters

Description	Symbol	Value	Source
Risk aversion parameter	γ	2	Standard from literature
Discount factor	β	0.97	Standard from literature
Private market interest rates	r_f	borrow 5.9%, save 0.9%	Johnson (2013)
Student loans interest rates	r_s	2.29% in school; 3.7% post school	Average from 1998/99-2005/06
Unemployment benefit	b	\$6000 per year	Kaplan (2012)
Baseline amortized periods	T^d	10	
Cutoff for income-based	y_{min}	13,699	
Income based repayment rate	ρ_{inc}	10%	Policy
Income based repayment period	T_{inc}	20	
2-year college tuition by states	COA		Digest of Education Statistics
State group 1		1,396	
State group 2		963	
State group 3		2,006	
State group 4		2,741	
4-year college tuition by states			
State group 1		4,900	
State group 2		7,508	
State group 3		8,718	
State group 4		13,359	
Room and board 2-year college		4,539	
Room and board 4-year college		6,532	
Terminal value function			Own calculation
Youth assets coefficient	α^y_{T1}	0.00418	
Youth wage coefficient	α^y_{T2}	0.05484	

[21] Before 2006–2007, interest rates on Stafford Loans (now known as Direct Loans) were variable, with different rates, depending on whether the borrower was in school, in the 6-month grace period after leaving school, or in the repayment period. The interest rates are the same in school and in the grace period but 0.6% higher during the repayment period. The average real interest rate on unsubsidized loans while in school is 2.29%, and the average real interest rate while in repayment is 2.89% from academic years 1998–1999 to 2005–2006 (Source: Federal Student Aid; `https://studentaid.ed.gov/sa/types/loans/interest-rates#older-rates`). In addition, most federal student loans have loan fees that are a percentage of the total loan amount. The loan fee is deducted proportionately from each loan disbursement. The money the students receive will be less than the amount they actually borrow, yet they are responsible for repaying the entire amount they borrowed. The loan fee before 2005–2006 was 4%. Loan fees are essentially a form of up-front interest. If the loan has a 10-year repayment term, a 4% fee will make the interest rate increase from 2.89% to 3.7%.

While not working, youth can receive unemployment benefits, which are set to \$6,000 per year (Kaplan, 2012). The terminal value function parameters are also fixed. According to the formula in Section 3.2.8, if everyone's life ends at age 68: $\alpha^y_{T1} = 0.00418$, and $\alpha^y_{T2} = 0.05484$.

Internally Estimated Parameters

The remaining 44 parameters are estimated using 128 moments. These moments are sufficient to identify all the parameters (identification is discussed in Section 3.5.2). The full list of the moments can be found in Appendix B.2. Key moments include:

- Fraction going to 2-year college and 4-year college for the full sample and by AFQT quartile, parental income quartile, and state group.
- Tuition elasticity of education: change in the fraction of 2- or 4-year college graduates when the 2- or 4-year college tuition increases by 1,000 dollars.
- Labor supply by education and year after school.
- Mean of log wages of youth working full-time by education level and year after school.
- Average parental transfers across all education levels and years.
- Average youth assets at different ages.
- Average annual amount of student loans by college type.
- 2-year cohort default rate among college graduates.

The estimated parameters are displayed in Table B1. Computational details are discussed in Appendix B.4.

3.5.2 Discussion of Identification

In this subsection, I briefly discuss the identification of internally estimated parameters. The identification of certain parameters is more straightforward, because they are arguments of the functions that relate observable variables to observable outcomes. These parameters include the grant function parameters in (3.13) and (3.14) and the log wage function parameters in (3.18). NLSY97 provides data on grants and log wages, so the identification comes from the observed data pattern. In addition, other parameters related to wages, including the mean and variance of log wages during college ($\bar{w}^c$ and σ_c^2), the variance of permanent and transitory shocks (σ_ξ^2 and σ_w^2), and the ratio of part-time wages to full-time wages, can be identified from the relevant wage statistics.

For the remaining parameters, I provide intuitive arguments below on how some of the moments can help to identify them. The altruism factor η can be identified from the average level of parental transfers. Conditional on parents' income and youth's income and assets, the optimal transfer is directly influenced by the weight that parents place on their children's utility.

The parameters related to preference for college (Equation (3.2)) can be identified with data on college enrollment. Conditional on labor market outcomes and monetary costs of college, the heterogeneity in the education distribution across groups by observables, such as AFQT, parental income, and state of residence, is helpful in identifying the coefficients of observables in (3.2). The external tuition elasticities help identify the variances of college preference shocks (σ_{z1}^2 and σ_{z2}^2). The variance of preference shocks determines how responsive college enrollments are to the changes in tuition costs. Therefore, the elasticity moments are useful sources of identification for the variances of the preference shocks.

Identification of parameters related to labor supply preference (Equation (3.1)) is similar. Data on labor supply by education can help identify the coefficients in (3.1). The variances of preference shocks (σ_{v1}^2 and σ_{v2}^2) affect the degree of change in the labor supply as a response to wage changes. Therefore, preference shock variances can be identified from data on labor supply by year as wages generally change over time. Employment transitions, i.e., the probability of working conditional on not working in the previous period and the probability of not working conditional on working in the previous period, help identify the job offer probability parameters in (3.16).

The argument for the identification of the remaining parameters is as follows. The fraction of borrowing and the average amount of borrowing are the keys to identifying the utility cost of borrowing in college (α_{b1} and α_{b2} in (3.3)). The default rate can be used to identify the utility cost of default (α_d in (3.4)). The key moments for identifying the borrowing limit parameters in (3.19) are the asset accumulation patterns of youth by age.[22]

3.5.3 Parameter Estimates and Model Fit

The parameter estimates and standard errors are reported in Table B1. Overall, the parameter estimates are reasonable. The estimated value for the altruism factor η is 0.05.[23] Results from the estimation show that the preference for college attendance is increasing both in AFQT and parental income. Grants are decreasing with family income at 2-year colleges, while in 4-year colleges, grants are increasing in AFQT and decreasing with family income. Wages increase with AFQT, education, and experience. The probability of receiving a job offer is rising in the human capital of the youth, but is much lower if the youth did not work in the previous period. Borrowing limits increase in human capital and age. Signs of the estimated utility cost of borrowing and default are as expected.

Overall, the model fits the data patterns well. The actual data and simulated data on 2- and 4-year college enrollment of the full sample and by subgroup are shown in Table 3.4. Simulated data closely match the pattern that college enrollment increases with ability and parental income. Table 3.5 shows the fit of tuition elasticities. The model generates the expected signs of elasticities, and the magnitudes are close.

The fit of average full-time log wages by education is shown in Figure 3.1, which indicates average full-time log wages by year after school for high school graduates, 2-year college graduates, and 4-year college graduates, respectively. Due to the length of NLSY97, I only match 10 years after school for high school graduates, 8 years after school for 2-year college graduates, and 6 years after school for 4-year college graduates. Beyond that, the sample sizes decrease, so I do not use moments from later years. The model fits the pattern of wages well.

Statistics on actual and simulated labor supply and employment transitions by education are shown in Table 3.6. The table summarizes the average fraction not working, working part-time, and working full-time over the years after school for different education groups. Simulated moments are quite closely aligned with actual labor supply and employment transitions.

[22] Similar to Johnson (2013).

[23] This is close to the estimate in Kaplan (2012), which is 0.04.

Table 3.4: Model Fit - Education

	HS grad		2-year college		4-year college	
	Data	Model	Data	Model	Data	Model
All	0.37	0.37	0.21	0.22	0.41	0.41
AFQT quartile 1	0.68	0.68	0.23	0.24	0.09	0.08
AFQT quartile 2	0.47	0.45	0.26	0.28	0.27	0.27
AFQT quartile 3	0.27	0.24	0.22	0.22	0.50	0.54
AFQT quartile 4	0.08	0.10	0.14	0.13	0.78	0.76
Family income quartile 1	0.57	0.64	0.25	0.16	0.18	0.20
Family income quartile 2	0.41	0.40	0.24	0.24	0.35	0.36
Family income quartile 3	0.37	0.26	0.20	0.25	0.42	0.49
Family income quartile 4	0.13	0.18	0.17	0.22	0.69	0.60
Tuition state group 1	0.39	0.38	0.26	0.21	0.35	0.41
Tuition state group 2	0.35	0.34	0.22	0.24	0.44	0.42
Tuition state group 3	0.38	0.35	0.14	0.22	0.48	0.43
Tuition state group 4	0.38	0.41	0.21	0.21	0.41	0.39

Table 3.5: Model Fit - Tuition Elasticity of Education (Change in Enrollment if Tuition Increases by $1,000)

	Data	Model
2-year enrollment to 2-year tuition	-0.034	-0.020
4-year enrollment to 2-year tuition	0.013	0.010
2-year enrollment to 4-year tuition	0.004	0.002
4-year enrollment to 4-year tuition	-0.009	-0.006

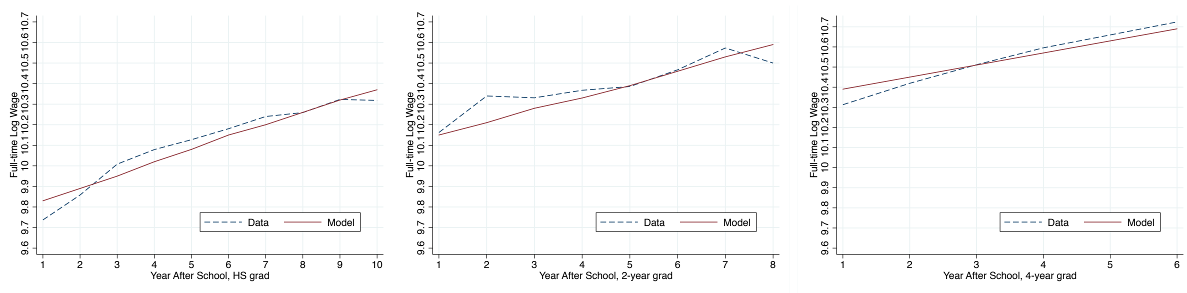


Figure 3.1: Full-time Log Wage

Table 3.6: Model Fit - Labor Supply and Transition

	HS grad		2-year grad		4-year grad	
	Data	Model	Data	Model	Data	Model
Fraction not work	0.07	0.08	0.05	0.05	0.04	0.05
Fraction part-time work	0.17	0.13	0.12	0.15	0.11	0.14
Fraction full-time work	0.76	0.80	0.83	0.81	0.84	0.81
Prob. not work to work	0.68	0.74	0.73	0.75	0.79	0.78
Prob. work to not work	0.05	0.04	0.04	0.03	0.03	0.03

Table 3.7 shows the fit of parental transfers. The targeted moment is the average parental transfers across all periods. The rest of the table contains moments that are not targeted in the estimation, including the percentage of youth receiving parental transfers and the average amount of parental transfers by youth education. The model does well in matching the average transfer levels, except that it underestimates the transfers for high school graduates. This is probably due to the model assumption of a homogeneous altruism factor. With a homogeneous altruism factor, high-income parents tend to make more transfers to their children, but in the data, transfers are relatively flat with parental income, indicating that there may be heterogeneity in parental altruism, as discussed in Kaplan (2012) and Park (2016). As parental income is highly correlated with children's education levels, the model with a homogeneous altruism factor tends to underestimate the average transfers of high school graduates and to overestimate

the average transfers of 4-year college graduates.[24] Another possible explanation is that the value of co-residency is assumed to be the same for all groups, but this value could be lower for youth with lower-income parents. Thus, the data moment overstates the value of parental transfers for these youth. However, without further evidence, it cannot be identified whether the data pattern is driven by the former (heterogeneous altruism factor), the latter (heterogeneous value of co-residency), or both.

Table 3.7: Model Fit - Parental Transfers

	Data	Model
Targeted Moment		
Average transfers	3,862	3,796
Untargeted Moments		
Fraction get transfers by education		
HS grad	0.14	0.11
2-year student	0.52	0.78
4-year student	0.73	0.84
2-year grad	0.14	0.13
4-year grad	0.22	0.14
Average transfers by education		
HS grad	2,752	1,104
2-year student	6,988	7,676
4-year student	10,779	13,955
2-year grad	2,186	1,521
4-year grad	1,553	1,958

The model also fits well along other dimensions. Table 3.8 shows the fit of the model for student loan borrowing and default. The average two-year cohort default rate for undergraduate borrowers is 9.26% between 1997 and 2011 (Yannelis, 2015). The model generates a slightly

[24] Kaplan (2012) discusses some ways to fix this problem, including allowing for a different curvature of utility for the two generations, adding an iceberg cost of transferring resources from the parent to the youth, or allowing for a negative correlation between altruism and parental income.

higher two-year default rate than the data. Table 3.9 shows the simulated average assets at different ages, which match the actual data closely.

Table 3.8: Model Fit - College

Panel A: In College

	2-year college		4-year college	
	Data	Model	Data	Model
Fraction not work	0.12	0.11	0.29	0.30
Fraction part-time work	0.43	0.42	0.50	0.50
Fraction full-time work	0.45	0.47	0.20	0.20
Avg. annual loans	867	913	2,005	2,020

Panel B: Student Loans Repayment

	Data	Model
Two-year default rate	9.26	11.07

Table 3.9: Model Fit - Assets

	Data	Model
Average assets at age 20	15,400	10,394
Average assets at age 25	30,725	33,699
Average assets at age 30	67,100	73,231

3.6 Introducing Income-based Repayment Plans

In this section, I use the estimated model to examine the effects of an IBR plan on various outcomes, including repayment, parental transfers, labor supply, college enrollment, and welfare. The IBR plan that I study in this section is similar to the current Pay As You Earn (PAYE) plan, which was not available when my sample started repaying their loans.[25] Annual income-contingent repayment is 10% of the discretionary income, but not more than the amount under the 10-year standard debt-based repayment plan. Discretionary income is equal to the borrower's income minus $13,699, which is 150% of the poverty guideline for a single-person family.[26] If discretionary income is negative, the borrower pays nothing. The payment period is 20 years. After this period, any remaining debt is forgiven. In practice, PAYE enrollment is not automatic: borrowers must apply for enrollment and report their income and family size every year to recalculate payments. Costs associated with understanding the regulations and filling out detailed paperwork may prevent distressed borrowers from using these forms of repayment assistance (Dynarski and Kreisman, 2013). Partly due to these non-monetary costs, the IBR enrollment rate was only 20% in 2014 (Government Accountability Office, 2015).

I conduct three policy experiments. First, I consider introducing IBR unexpectedly after youth graduate from college. This allows me to examine the effects of IBR, conditional on college enrollment and student loan borrowing, on post-college decisions and outcomes such as repayment, default, parental transfers, and labor supply. Second, I study the equilibrium effects of implementing a budget-neutral IBR with sign-up costs to match the current IBR sign-up rate. This can reveal which group of borrowers are most likely to be enrolled in IBR when sign-up costs are incurred and what the impacts are on college enrollment, student loan borrowing, and default rates. Finally, I remove the non-monetary sign-up cost while keeping the budget fixed. This is an ideal situation where everyone is automatically enrolled in IBR. It allows me to further examine the effects of eliminating sign-up costs on college enrollment and welfare.

3.6.1 Unexpected IBR

The first policy experiment is to introduce the IBR plan unexpectedly after youth graduate from college, so they do not know the existence of the plan before making decisions on college

[25] Besides PAYE, there are other IBR plans that are currently available. They differ in the calculation of payments and the time frames. See https://studentaid.ed.gov/sa/repay-loans/understand/plans for details on repayment plans.

[26] The poverty guidelines for the 48 contiguous states and the District of Columbia for a single-person family was $12,140 in 2018, and 150% of that amount is $18,210, which is equivalent to around $13,699 in 2004 dollars.

enrollment and student loan borrowing. Borrowers are free to choose IBR at the first year of repayment and, once they have chosen IBR, they have to commit to it for the remaining repayment periods. If they do not choose IBR in the first year, they can only choose between debt-based repayment and default for the rest of the repayment periods. This policy experiment allows me to examine the effects of IBR on post-college decisions and outcomes, conditional on college enrollment and student loan borrowing.

One interesting question is what percentage of borrowers choose IBR if it is introduced without any sign-up costs. My simulation shows that under this policy, the sign-up rates for IBR are 99.14% for 2-year college borrowers and 97.06% for 4-year college borrowers (97.50% for all borrowers), which are significantly higher than the recent 20% sign-up rate. This means that for most borrowers, IBR is better than debt-based repayment and default because of the insurance and potential debt write-off it provides. In the baseline, there are only two options for repayment: debt-based payment or default. Table 3.10 summarizes the characteristics of borrowers in the first year of repayment by their repayment status and college type. The table shows that those with high debt, low incomes, low levels of parental transfers, and little savings are more likely to default in the baseline. After the introduction of IBR, those who default in the baseline switch to IBR because they pay very little due to low incomes, benefit from the additional insurance and debt relief under IBR, and no longer suffer the cost of default. Now, compare IBR to the debt-based repayment plan. If the non-monetary cost of IBR is zero, IBR is at least as good as debt-based repayment in a static setting. What makes IBR more attractive is that it provides potentially reduced payments, insurance, and debt write-off. One drawback of IBR is its longer repayment period, during which more interest accrues compared to the standard repayment case. But the high sign-up rate indicates that the cost of accumulated interest is outweighed by the benefits. Hence, if the non-monetary cost of IBR is zero or negligible, almost everyone chooses IBR and no one defaults.

The post-college decisions and outcomes that I focus on include repayment, default rates, parental transfers, and labor supply. To examine whether there is heterogeneity in response to the policy across borrowers from different families, I report the effects in the first year after graduation by parental income quartile in Table 3.11. In the baseline, youth from richer families tend to borrow less on average, except that youth from the highest parental income quartile

Table 3.10: Characteristics of Borrowers in First Year of Repayment (Baseline)

	2-year graduate		4-year graduate		All	
	Non-defaulter	Defaulter	Non-defaulter	Defaulter	Non-defaulter	Defaulter
Fraction	94.8%	5.2%	88.9%	11.1%	90.2%	9.8%
Average debt	4,018	5,503	8,843	13,632	7,765	12,713
Average wage	26,028	9,997	33,060	13,230	31,489	12,864
Average debt-based/wage	2.4%	8.1%	4.5%	15.7%	4.0%	14.8%
Average asset	10,132	-556	19,796	-439	17,637	-452
Average parental income	49,735	27,114	82,481	44,302	75,164	42,358
Average parental transfers	1,089	723	3,483	2,273	2,948	2,098

borrow more on average than youth from the third parental income quartile. That is because a higher fraction of quartile four borrowers attend a 4-year college compared to quartile three borrowers.

Moving from the baseline to the unexpected IBR, the average repayment during the first year does not change much for the lowest parental income quartile borrowers, but changes the most for the highest quartile borrowers — their average repayment decreases by \$301 during the first year. Looking at 2-year graduates and 4-year graduates separately (Table 3.12), I find that for 4-year graduates from low-income families, their average repayment actually increases by \$70. This is because, under the debt-based repayment plan, a sizable fraction of borrowers from low-income families default at an early stage. After IBR is introduced, they start to repay small amounts. On the contrary, borrowers from high-income families receive the largest reduction in repayment, on average, because they rarely default in the baseline while often paying a lower income-based amount under the IBR. As shown in Table 3.11, 23% of the borrowers from the lowest income quartile default, while only 1% of the borrowers from the highest quartile default in the baseline. After IBR is introduced, the default rate is reduced to zero for all quartiles.

Table 3.11 also shows that introducing IBR crowds out parental transfers, especially for high-income families. On average, parental transfers fall by \$582 during the first year for the highest income group. For low-income families, average parental transfers do not change since transfers are negligible prior to the policy change.

Table 3.11: Effects of Unexpected IBR on College Graduates with Student Loans in First Year of Repayment by Parental Income Quartile

		Parental Income Quartile			
		Q1	Q2	Q3	Q4
Average debt		11,334	8,736	6,342	7,651
Average repayment	Baseline	963	902	667	904
	Unexpected IBR	961	767	557	603
	Change	-2	-136	-109	-301
Default rate	Baseline	23%	11%	8%	1%
	Unexpected IBR	0%	0%	0%	0%
	Change	-23%	-11%	-8%	-1%
Average parental transfers	Baseline	9	196	972	8,432
	Unexpected IBR	9	195	939	7,849
	Change	0	-1	-33	-582
Average youth wage	Baseline	30,515	31,288	29,540	27,960
	Unexpected IBR	30,382	31,176	29,490	27,907
	Change	-133	-112	-50	-53
Fraction work full-time	Baseline	87.37%	84.94%	80.93%	74.39%
	Unexpected IBR	86.73%	84.34%	80.58%	74.07%
	Change	-0.64%	-0.60%	-0.35%	-0.32%
Average youth asset	Baseline	14,014	11,678	6,794	28,404
	Unexpected IBR	13,486	11,387	6,644	27,724
	Change	-528	-292	-150	-680
Average youth consumption	Baseline	17,834	19,833	22,277	29,073
	Unexpected IBR	18,231	20,146	22,453	29,418
	Change	397	314	176	345
Average parent consumption	Baseline	24,096	44,862	68,059	117,081
	Unexpected IBR	24,096	44,863	68,092	117,664
	Change	0	1	33	582
Fraction that has debt write-off	Unexpected IBR	10.01%	6.88%	7.00%	7.13%
Average debt write-off if positive	Unexpected IBR	6,041	4,606	3,464	5,076

Table 3.12: Effects of Unexpected IBR on College Graduates with Student Loans by College Type and Parental Income Quartile

		Parental Income Quartile			
		Q1	Q2	Q3	Q4
2-year college graduates					
Average debt		5,290	3,466	3,375	3,234
Average first year repayment	Baseline	571	394	406	393
	Unexpected IBR	469	312	286	276
	Change	-102	-82	-120	-117
Average fraction has debt write-off	Unexpected IBR	9.54%	6.06%	7.88%	6.94%
Average debt write-off if positive	Unexpected IBR	2,666	1,603	1,868	1,664
4-year college graduates					
Average debt		15,645	10,486	7,068	7,988
Average first year repayment	Baseline	1,242	1,071	731	943
	Unexpected IBR	1,312	918	624	628
	Change	70	-153	-107	-315
Average fraction has debt write-off	Unexpected IBR	10.36%	7.15%	6.78%	7.14%
Average debt write-off if positive	Unexpected IBR	8,258	5,452	3,918	5,329

One concern about IBR is moral hazard — income-based repayment is essentially an income tax which may discourage labor supply if the borrower does not expect to repay the loan fully. My simulation shows that borrowers from low-income families reduce their labor supply the most, mainly switching from full-time to part-time work. However, the percentage that changes the labor supply is minimal — less than 1%. Therefore, the moral hazard effect is trivial.

Results from the simulation show that borrowers across all quartiles reduce their savings, as IBR provides more insurance and reduces the need for self-insurance through savings. Youth consumption increases in all groups, but the sources are different. For youth from low-income families, increased consumption mainly comes from the reduction in assets. For high-income family borrowers, the decline in both student loan repayments and assets leads to higher youth consumption. However, the decrease in parental transfers offsets part of the increase. Consumption of low-income parents does not change, while high-income parents consume more because parental transfers decline as a result of the crowding-out effects of IBR. Regarding total household consumption, a higher-income family benefits more. Although only the results for the first repayment year are reported here, similar patterns exist for other repayment years.

IBR has two opposing effects on the government budget. On the one hand, given that the average repayment is lower, the government may expect a revenue loss compared to the baseline. In addition, the government writes off the remaining debt under IBR at the end of 20 years. My simulation shows that, on average, 7.58% of borrowers receive a debt write-off and the average write-off amount, if positive, is $4,813. The last two rows in Table 3.11 show that borrowers from low-income families are more likely to have debt write-offs and higher write-off amounts. On the other hand, it is also possible for the government to collect more, given that the repayment period is longer — 20 years compared to 10 years in the baseline. Therefore, the impact on the government budget is ambiguous without further quantitative evidence.

To quantify the effects on the government budget, I calculate the government's internal rate of return (IRR), which is the interest rate at which the present value of the total student loan borrowing amounts equals the present value of the total repayment amounts.[27] The IRR is 3.38% in the baseline, and it is reduced to 3.20% after IBR is introduced unexpectedly, which means a loss of revenue for the government. Thus, the IBR raises the cost of the student loan program, and the government would have to raise the interest rate or subsidize the program to keep the budget constant.

3.6.2 IBR with Sign-up Cost

Next, I consider the case when agents know that IBR is available before making college enrollment decisions. As shown in Section 3.6.1, the IBR sign-up rates are very high — nearly 100% when introduced at no cost. However, the IBR plan enrollment rate was only 20% in 2014 (Government Accountability Office, 2015), indicating that there are nontrivial costs associated with enrollment in IBR plans. Therefore, in this subsection, I apply a non-monetary cost

[27]Another option for comparison is to discount the cash flows from and to the government over the life of student loans using a certain rate, such as the Treasury rate. However, Lucas and Moore (2010b) point out that using Treasury rates without risk adjustment for discounting tends to underestimate the cost of student loan programs compared to discounting at market rates. Therefore, instead of using Treasury rates, I calculate the IRR.

to match the 20% IBR enrollment rate. As the previous subsection demonstrated, IBR results in revenue losses to the government, so I increase interest rates to keep the IRR the same as the baseline. In sum, the policy experiment examined in this subsection is a budget-neutral IBR with sign-up costs, known to the agents prior to college entrance.

The IBR sign-up cost is modeled as a one-time utility cost to be paid by the borrower at the first repayment period after enrolling in IBR. Table 3.13 summarizes the characteristics of borrowers by their repayment plan in the first year of payment. Compared to those paying the debt-based amount, those paying the reduced amount (IBR or default) have a higher debt-to-wage ratio and lower parental transfers. This finding is consistent with the prediction in Lochner et al. (2018), which uses a theoretical model where the non-monetary costs are moderate. While 4% of borrowers are still choosing default, it is much lower than the default rate in the baseline. Compared to the defaulters, those in IBR have significantly lower parental transfers. This is mainly because those with little parental support benefit more from the reduced payments and insurance provided by IBR than other borrowers.

Table 3.13: Features of Borrowers under IBR with Utility Cost by Repayment Status

	Debt-based	IBR	Default
Fraction	76%	20%	4%
Average debt	6,750	16,226	9,466
Average parental income	85,416	32,734	58,674
Average parental transfers	4,316	171	4,093
Average wage	29,589	32,367	8,913
Average debt-based repayment to wage ratio	4.0%	9.0%	14.8%

Intuitively, IBR provides more insurance against labor market shocks. Thus, it encourages youth to go to college and borrow more from student loans. In the meantime, however, the interest rate on student loans must increase to keep the budget constant, which increases the cost of post-secondary education and discourages some youth from attending college and borrowing. Column (2) of Table 3.14 summarizes the aggregate effects on college enrollment and student loan borrowing. To equate the IRR, the interest rate needs to increase by 0.11 percent-

Table 3.14: Effects of IBR on Education and Borrowing

	(1) Baseline	(2) IBR + sign-up cost + balance budget	(3) IBR + balance budget
Internal Rate of Return	3.38%	3.38%	3.38%
Student Loans Interest Rate			
In School	2.29%	2.40%	2.53%
In Repayment	3.70%	3.81%	3.94%
Education			
Frac. HS grad	36.9%	36.1%	35.1%
Frac. 2-year grad	21.9%	22.1%	22.8%
Frac. 4-year grad	41.2%	41.8%	42.1%
Student loans			
Avg. loans, 2-year	913	1,174	1,594
Avg. loans, 4-year	2,020	2,322	2,642

age points, which means that the interest rates of student loans increase from 2.29% to 2.40% in school and from 3.70% to 3.81% in repayment. My simulation also shows that college enrollment increases — 2-year enrollment increases by 0.2 percentage points and 4-year enrollment increases by 0.6 percentage points. Average borrowing also increases for both 2-year and 4-year college students. In practice, IBR with utility costs might be more feasible compared to the case where sign-up costs are eliminated entirely (discussed in the next subsection) given that the costs may reflect a stigma associated with underpaying, which cannot be completely eliminated.

3.6.3 Equilibrium Effects of IBR

Now consider removing the non-monetary cost of IBR while keeping the budget constant. Since IBR without any sign-up costs encourages more borrowers to enroll (as shown in Section 3.6.1), the interest rates must increase to keep the student loan program financially self-sustaining. As shown in column (3) of Table 3.14, compared to the baseline, the interest rate needs to increase by 0.24 percentage points to equalize the IRR.

A comparison of columns (2) and (3) shows that the elimination of IBR sign-up costs further encourages youth to attend college. Two-year college enrollment increases by 0.7 percentage points and 4-year college enrollment increases by 0.3 percentage points. Taken together, moving from the baseline to this cost-free, budget-neutral IBR increases both 2- and 4-year enrollment by 0.9 percentage points. Removing non-monetary costs also encourages college students to borrow more from student loans.

To examine the distributional effects, I report changes in college enrollment by parental income and AFQT quartile in Table 3.15. The enrollment of the low parental income group increases the most. For the lowest parental income group, 2-year college enrollment increases by 2.7 percentage points and 4-year college enrollment increases by 1.3 percentage points. The low AFQT group gains more in 2-year college enrollment, while changes in 4-year college enrollment are similar across different AFQT quartiles. In terms of overall college enrollment (2- and 4-year), low parental income and low AFQT youth benefit the most.

Table 3.15: Effects of IBR on Education by College Type

	2-year Enrollment			4-year Enrollment		
	Baseline	IBR + balance budget	Change	Baseline	IBR + balance budget	Change
Full Sample	21.9%	22.8%	0.9%	41.2%	42.1%	0.9%
By Parental Income Quartile						
Income Q1	16.0%	18.7%	2.7%	20.3%	21.6%	1.3%
Income Q2	24.3%	25.1%	0.8%	35.9%	37.5%	1.6%
Income Q3	25.0%	25.2%	0.2%	49.4%	49.6%	0.2%
Income Q4	22.4%	22.2%	-0.1%	59.7%	59.9%	0.3%
By AFQT Quartile						
AFQT Q1	24.0%	26.2%	2.2%	8.1%	8.8%	0.7%
AFQT Q2	28.1%	29.0%	0.9%	27.3%	28.2%	0.9%
AFQT Q3	22.1%	22.5%	0.4%	53.6%	54.6%	1.0%
AFQT Q4	13.3%	13.4%	0.0%	76.4%	77.1%	0.7%

To examine the impact of different IBR policies on welfare, I calculate changes in welfare overall, by parental income, and by AFQT quartile. The results are shown in Table 3.16. The welfare measurement is the fraction of consumption that households are willing to give up to live under the new policy so that the expected lifetime utility is the same as in the baseline. When IBR is introduced with sign-up costs (column (1)), households are willing to give up 0.04% of their consumption, which means that the policy is welfare-improving. All parental income groups benefit from the policy, but the quartile two families benefit the most. As shown in Section 3.6.2, borrowers from low-income families are more likely to pay the non-monetary cost and enroll in IBR. In addition, the college enrollment rate increases the most for low-income families. When the sign-up cost is removed (column (2)), welfare increases further — average households are willing to give up 0.08% of their consumption. Compared to column (1), the elimination of the sign-up cost increases the welfare of families across all quartiles. In both policy experiments, the magnitude of welfare improvement is increasing with youth's ability.

In sum, my policy experiments show that IBR reduces average payments and crowds out parental transfers. The response of the labor supply to IBR is minimal therefore the problem of moral hazard is less likely to be a matter of concern. When IBR is implemented with sign-up costs, those with fewer parental resources are more likely to participate in IBR. Under IBR with sign-up costs, students borrow more and repay less so that the interest rates of student loans have to increase to keep the program self-sustaining. Yet, even after interest rates increase, college enrollment is still higher than the baseline. When removing the sign-up costs while keeping the budget fixed, college enrollment increases further. This cost-free and budget-neutral IBR is welfare-improving, and welfare increases more for relatively poor and high-ability families. For different families, the source of welfare improvement is different. Low-income families benefit from higher college enrollment, less default, more debt write-offs, higher youth consumption, and more insurance, while high-income families benefit from reduced payments and increased total household consumption when their post-school earnings are low.

Table 3.16: Effects of IBR on Welfare

	Change of Consumption	
	(1) IBR + sign-up cost + balance budget	(2) IBR + balance budget
Aggregate	-0.04%	-0.08%
By Parental Income Quartile		
Income Q1	-0.03%	-0.04%
Income Q2	-0.13%	-0.25%
Income Q3	-0.04%	-0.10%
Income Q4	-0.07%	-0.10%
By AFQT Quartile		
AFQT Q1	-0.02%	-0.04%
AFQT Q2	-0.03%	-0.07%
AFQT Q3	-0.07%	-0.13%
AFQT Q4	-0.17%	-0.21%

Note: The welfare measurement is the fraction of consumption that households are willing to give up to live under the new policy so that the expected lifetime utility is the same as in the baseline.

3.7 Conclusion

This paper develops a dynamic life-cycle model to study the effects of introducing an IBR plan with endogenous parents-to-children transfers, together with children's education, borrowing, repayment, and labor supply decisions. My estimated model suggests that students from low-income families borrow more and are more likely to default. They are also more likely to pay the sign-up cost to enroll in an IBR in counterfactual simulations, indicating that they need the insurance provided by IBR the most. IBR crowds out savings and parental transfers as it provides more insurance to borrowers. Regarding the moral hazard issue, my simulation shows that the labor supply response is trivial after IBR is introduced.

IBR encourages youth to borrow more and repay less, leading to government revenue losses. Therefore, the interest rates of student loans must increase to keep the program self-sufficient. My simulation shows that a budget-neutral IBR without sign-up costs increases college enrollment rates and improves welfare with relatively poor families benefiting more.

As college tuition and student borrowing have increased dramatically in recent decades, along with increasing labor market risk (Lochner and Monge-Naranjo, 2016), the repayment problem has become more prominent. With continually rising tuition, resources such as savings and parental transfers play more important roles. This framework is promising in studying optimal student loan policies when considering the insurance roles played by these resources. In addition, countries such as Canada have also introduced more generous repayment plans (e.g., Repayment Assistance Plan), so this framework is useful in studying the effects of those policies in an international context. One possible extension of the model is to add dropout risk during college (Chatterjee and Ionescu, 2012), since a large fraction of students do not get a degree but have student loans when they leave college. This additional risk may amplify the role of parental resources and affect the educational choices of those with little parental assistance. Introducing IBR provides insurance that would, therefore, benefit poor families more. However, IBR could also lead to a larger government budget deficit, given that dropouts may have more difficulties in repaying their loans.

Chapter 4

Returns to Skill and the Evolution of Skills for Older Men

4.1 Introduction

Despite decades of research on the topic, there remains considerable interest in better understanding growing inequality in the U.S. and many other developed countries. An important focus of much of this research has been on the extent to which growing wage inequality is the result of rising returns to skill (often attributed to skill-biased technological change) vs. growth in the variance of skills across workers. Due to a lack of direct measures of skill, researchers are typically forced to make (strong) assumptions about the evolution of skill distributions or returns, with little external validation or evidence regarding those assumptions.

This paper establishes nonparametric identification of the returns to skills and cross-sectional distribution of skills over time given the availability of repeated cross-section data on wages and at least two skill measurements every period, with at least one continuous skill measure repeated each period. With longitudinal data, we show that it is also possible to identify the dynamics of skills (i.e. the distribution of skills in period t conditional on skills in period $t-1$). Our constructive identification strategy suggests a multi-stage estimation approach, which simplifies considerably if one of the repeated measurements is known to be linear in skills, something that is straightforward to verify. We use these methods and longitudinal data from the Health and Retirement Study (HRS) to estimate the evolution of skill returns and distributions, as well as the dynamics of skills, for men in the U.S. from 1996–2016.

Several distinct literatures in economics aim to distinguish interpersonal differences in skills from the market-level returns to those skills. For example, the primary objective of many empirical studies on discrimination is to determine the extent to which race or gender differences in wages, as well as the evolution of those gaps over time, can be explained by group differences in skill levels.[1] Similarly, researchers often attempt to decompose differences in the wage returns to schooling across countries (Leuven et al., 2004; Hanushek and Zhang, 2009) or over time (Heckman et al., 1998; Bowlus and Robinson, 2012) into differences in worker skill levels (deriving from, e.g., heterogeneous school quality or home environments) and in the wage returns to those skills. Related research has framed the rapid rise in residual wage inequality (i.e., inequality within narrowly defined demographic groups) over the past several decades as a combination of changes in the distribution of unmeasured skills and in their labor market returns (e.g., Juhn et al., 1989; Katz and Murphy, 1992; Lemieux, 2006; Autor et al., 2008; Lochner et al., 2020).

Skill measurement is a critical challenge in all of these literatures. In most cases (e.g. studies using Census data or data from the Current Population Surveys (CPS)), only a crude proxy or correlate of skill is available (e.g., educational attainment, per-pupil spending when young, age), especially when researchers are interested in studying inequality across long time periods. In these cases, skills are often equated with available measures like education or labor market experience. Other studies explicitly aim to estimate the role of unmeasured skills. To this end, Juhn et al. (1989) assume that the distribution of these skills remained constant over the period they study, attributing all growth in the variance of log wage residuals to an increase in the return to unobserved skill. Lemieux (2006) instead assumes that the variance of skills within narrowly defined observable groups (e.g., within age-education-race categories) remained unchanged over time, allowing for changes in the distribution of skills through changes in the composition of the workforce (by age, education, and race). He finds that a sizeable fraction of the growth in residual inequality can be traced to changes in the distribution of skills caused by the aging and growing education of the population. Using longitudinal data on wages from the Panel Study of Income Dynamics (PSID), Lochner et al. (2020) relax the assumption of invariant within-group distributions, assuming instead that variation in skill growth among older

[1]There are vast literatures on race and gender wage differentials surveyed in Altonji et al. (2012). Among the most closely related studies on race, see Card and Lemieux (1996); Neal and Johnson (1996), and Chay and Lee (2000). See Blau and Kahn (1997, 2017) for closely related studies on gender wage gaps.

workers is idiosyncratic. Their estimates suggest declining returns to unobserved skill over the late 1980s and 1990s, while the growth in residual wage inequality is instead explained by growth in the variance of skills (due to growing variation in skill growth). Clearly, assumptions about the evolution of skill distributions have important implications for the conclusions one draws about driving forces underlying rising wage inequality.[2]

In some cases, researchers have used more specialized data sets like the National Longitudinal Surveys of Youth (NLSY), which contain cognitive test scores as direct measures of skill. Using the 1979 Cohort of the NLSY, Neal and Johnson (1996) demonstrate that differences in adolescent cognitive achievement (as measured by the Armed Forces Qualifying Test, AFQT) can, by themselves, explain the sizeable differences in wages between young black and white men in the U.S.[3] Several studies have also used the NLSY cohorts to study the evolution of inequality since the 1980s. For example, Herrnstein and Murray (1994) argue that the U.S. has become more meritocratic based on sharply increasing wage returns to AFQT. Others have used AFQT measures in an effort to disentangle whether the growing differences in earnings by educational attainment reflect rising returns to schooling or rising returns to cognitive ability (Heckman and Vytlacil, 2001; Taber, 2001; Castex and Dechter, 2014).[4] Deming (2017) exploits other non-cognitive measures in the NLSY, estimating that the returns to social skills have risen since the 1980s.

It is noteworthy that commonly used data sources with direct skill measures do not typically contain the same measures over time for the same individuals. For example, cohorts of the NLSY contain AFQT scores measured only once, during adolescence. Thus, studies using the NLSY estimate the effects of adolescent cognitive achievement, rather than contemporaneous skills, on wages later in life. Given the practical challenge of sorting out age and time effects

[2] In related research, Card and Lemieux (1996) and Chay and Lee (2000) use CPS data to study the extent to which changes in skill gaps and the returns to skill can explain the evolution of black – white wage differentials. Card and Lemieux (1996) consider a single skill model (composed of both observed and unobserved components) with restrictions on the evolution of skills over time, while Chay and Lee (2000) consider a model with differently priced observed and unobserved skills, placing restrictions on the distribution of unobserved skills within observable groups over time.

[3] To study differential returns to cognitive achievement across countries at a point in time, Leuven et al. (2004) and Hanushek and Zhang (2009) use international data from the International Adult Literacy Survey (IALS) while Hanushek et al. (2015) exploits data from an expanded set of countries from the Programme for International Assessment of Adult Competencies (PIAAC).

[4] In related work, Altonji et al. (2012) study changes in the distribution of skills (overall and by race and gender) for two NLSY cohorts using AFQT scores, education, and other individual and family characteristics.

from only a few birth-year cohorts, studies following individuals over time from one of the NLSY cohorts cannot determine whether growing wage inequality is driven by differential lifecycle growth in skills by AFQT or rising returns to skill (Heckman and Vytlacil, 2001).

Grogger and Eide (1995) and Murnane et al. (1995) address this issue by exploiting data on cognitive achievement and wages from two separate cohorts (National Longitudinal Study of the High School Class of 1972 (NELS72) and High School and Beyond (HSB)). Comparing the earnings of individuals at the same ages (roughly age 24), their estimates suggest that both the returns to schooling and cognitive skill rose between 1978 and 1986. However, a limitation of both studies is that the cognitive tests (taken during the last year of high school), while similar, were not the same across the two surveys. Thus, any differences in their mapping to true cognitive skills would be reflected in the estimated returns to skill over time. Castex and Dechter (2014) improve upon these studies by comparing the wages of men over ages 18-28 from the 1979 and 1997 Cohorts of the NLSY (born roughly 20 years apart), who took the same AFQT test during adolescence. Their estimates suggest that the returns to adolescent cognitive achievement declined while the returns to schooling increased between the 1980s and 2000s.[5] They also find that log wage differences by AFQT were quite similar upon labor market entry across the two cohorts, with log wage gaps by AFQT increasing with experience for the 1979 Cohort but not the 1997 Cohort. Thus, the difference in estimated returns to AFQT across cohorts only appeared after individuals had spent several years in the labor market.

While this cross-cohort approach to estimating changes in the returns to skills over time requires weaker assumptions than studies following a single cohort over time, it is not assumption-free. Because most respondents took the AFQT during adolescence, one can only interpret changes in wage returns to cognitive skills at older ages as changes in the returns to skill over time if the evolution of skills between the age of the test and the age at which wages are com-

[5] In related work, Deming (2017) uses the 1979 and 1997 Cohorts of the NLSY to estimate changes in the wage returns to social skills over time. Because the same measures of social skills are not available for both NLSY cohorts, Deming (2017) works with normalized measures of social skills (measured during adolescence), effectively assuming identical distributions (and measurement quality) across the cohorts. Thus, the rising returns to social skills he estimates could also reflect greater variation in social skills (or more precise measurements) for the more recent cohort. Edin et al. (2017) exploit administrative data in Sweden that contain consistent cognitive and non-cognitive measures collected for men entering the military at ages 18-19 for cohorts born between 1951 and 1975. They estimate that the return to non-cognitive skill roughly doubled from 1992 to 2013, while the return to cognitive skill rose in the 1990s but fell in the 2000s.

pared was the same across cohorts.[6] This would be violated if, for example, variation in early lifecycle skill growth changed. Indeed, the finding that log wage gaps by AFQT evolved differently with work experience for the two NLSY cohorts suggests that early skill growth may have differed between them. Alternatively, the returns to cognitive skill may have evolved differently in the 1980s vs. 2000s, which cannot be easily distinguished from differential lifecycle skill growth patterns for the two cohorts.

Despite any limitations, these cohort comparison studies provide some of the most convincing evidence on changes in the returns to cognitive skills over time. Yet, they only offer, at best, a few snapshots for the U.S.: Grogger and Eide (1995) and Murnane et al. (1995) find that the returns to cognitive ability rose between 1978 and 1986, while Castex and Dechter (2014) estimate that the returns declined between the 1980s and 2000s. These studies do not tell us anything about the 1960s and early 1970s or the most recent decade. Nor do they inform us as to when the sizeable decline in estimated returns occurred between the 1980s and 2000s. Yet, these are periods of considerable debate in the literature on log wage residual inequality and the evolution of returns to skill (Juhn et al., 1989; Katz and Murphy, 1992; Lemieux, 2006; Autor et al., 2008; Lochner et al., 2020).

Although we cannot comment on earlier time periods, we use biennial data from 1996-2016 HRS to estimate the evolution of returns to skill and skill distributions over this more recent period. Unlike previous studies, our data contain measures of wages and scores from the same cognitive tests repeated every other year. As a result, our approach requires no assumptions on skill distributions nor their dynamics. Instead, our key identification assumption is that the measurement function for at least one continuous repeated test (taken at ages 52+) is identical over time (i.e., the mapping between skills and expected individual test scores is time invariant). We further show that our use of panel data enables us to identify the lifecycle dynamics of skills under very general conditions.

Our estimates suggest that the returns to (cognitive) skill were relatively stable over the late

[6] The same caveat applies to studies by Grogger and Eide (1995) and Murnane et al. (1995), who use data from NELS72 and HSB, and to Edin et al. (2017), who use administrative data from Sweden. Additionally, the fact that NLSY respondents took the AFQT test at different ages across the cohorts requires adjustments based on implicit assumptions regarding rank stability of cognitive scores across testing ages. Of course, test scores measured at age 22 for the oldest in the 1979 Cohort are likely to be much more strongly related to cognitive skills at ages 18-28 than are test scores measured at age 12 for the youngest in the 1997 Cohort.

1990s and early 2000s but rose significantly after the Great Recession. While these estimates are noisy, they are roughly consistent with the patterns estimated by Lochner et al. (2020) using data from the PSID.

Our results also intersect with a broader literature studying cognitive skills late in life; although, this literature typically estimates latent skills (or factors) derived only from cognitive tests without linking (or anchoring) those skills to wages. By incorporating wages in our analysis, we are able to describe how skills, measured in log wage units, evolve for individuals ages 52-70. We document considerable variation in skills among individuals in their early- to mid-50s across cohorts (with much higher skills among those from earlier birth cohorts), but these differences are largely dissipated by the time individuals reach their early 60s as earlier cohorts experienced much faster declines in skill than did later cohorts. Indeed, there is little evidence of skill depreciation prior to age 63 (when last observed) among the cohort born 1954-1959.

Consistent with several studies using the HRS and similar data in other countries, we show that late-life cognitive performance differs significantly across education groups (Cagney and Lauderdale, 2002; Mazzonna and Peracchi, 2012) and racial groups (Zsembik and Peek, 2001; Karlamangla et al., 2009; Castora-Binkley et al., 2015). Our estimates suggest that, among older workers, cognitive skill differences are quite similar across ages and explain roughly one-third of the education gaps and nearly half of the race gaps in log wages.[7]

Prior research has also shown that retirement is negatively correlated with cognition (Adam et al., 2007); although, self-selection into retirement has posed challenges in estimating the causal relationship between retirement and cognition.[8] We make no effort to attribute causality; instead, we simply show lifecycle skill profiles from ages 52-70 for workers choosing to retire at different ages. Our results suggest that age 55 skill levels are increasing in retirement age, with those retiring at age 65 or older possessing roughly 15% more skills than those retiring

[7] See Card (1999) and Heckman et al. (2006) for comprehensive surveys of wage differences by education. Neal and Johnson (1996) show that adolescent skill gaps can explain much of the early-career differences in wages by race.

[8] Several studies use exogenous policy variation, such as eligible retirement ages and pension policies, as instruments to identify the causal effects, producing mixed findings. Some of these studies estimate significant negative effects of retirement on cognition (Rohwedder and Willis, 2010; Bonsang et al., 2012; Mazzonna and Peracchi, 2012, 2017), while others find no causal effect (Coe and Zamarro, 2011; Coe et al., 2012). Still others estimate heterogeneous effects across different occupations (Mazzonna and Peracchi, 2017) or across individuals retiring early vs. at the statutory age (Celidoni et al., 2017).

prior to age 55. We see no evidence of sharp declines in cognitive skills surrounding retirement ages, with skills relatively constant through age 60 for those retiring prior to age 55 and skills declining almost linearly from ages 52 to 70 for those retiring over ages 55-64. To the extent that retirement does lead to cognitive decline, our results suggest that its impacts are relatively small or largely offset by other lifecycle forces.

Finally, we show that skills are quite persistent, with individual fixed effects accounting for more than a third of all skill variation at age 60. Year-to-year skill innovations are also persistent with an autocorrelation of 0.93.

This chapter proceeds as follows. Section 4.2 describes our model of skill dynamics, the relationship between skills and wages, and other skill measurement functions. We discuss identification and estimation of skill returns, measurement functions, skill distributions, and the dynamics of skills. Section 4.3 describes the HRS data we use in estimation, while Section 4.4 presents our estimation results. We offer concluding thoughts in Section 4.5.

4.2 Methodology

In this section, we provide a general model of the evolution of skills and log wages. Since we focus on older workers (ages 50+), we assume skills evolve exogenously, reflecting growth and/or depreciation. With panel data on both log wages and test-based measures of skills, we describe identification and estimation of the distribution of skills and their dynamics, the evolution of skill return functions over time, and the mapping between skills and their test-based measurements.

4.2.1 Model

Let $\ln W_{i,a,t}$ be the log wage for individual i at age a in year t and $T_{i,j,a,t}$ reflect skill test score measure $j = 1, \ldots, J$. We consider the following specification for log wages and skill measures:

$$
\begin{aligned}
\ln W_{i,a,t} &= \gamma_t + \lambda_t \theta_{i,a,t} + \varepsilon_{i,a,t}, \\
T_{i,1,a,t} &= \tau_1(\theta_{i,a,t}) + \eta_{i,1,a,t}, \\
T_{i,j,a,t} &= G_j(\tau_j(\theta_{i,a,t}) + \eta_{i,j,a,t}), \text{ for } j = 2, \ldots, J,
\end{aligned}
\tag{4.1}
$$

where $\theta_{i,a,t}$ denotes unobserved skill, λ_t denotes "return" to skill in period t, $\varepsilon_{i,a,t}$ and $\eta_{i,j,a,t}$ are idiosyncratic non-skill shocks to wages and test measurement errors, respectively, and $\tau_j(\cdot)$ is a strictly increasing, age- and time-invariant measurement function that maps unobserved skill to cognitive measure j combined with a weakly increasing function $G_j(\cdot)$ for $j = 2, \ldots, J$. Notice that the model allows ordered discrete measures for $j = 2, \ldots, J$ if $T_{i,1,a,t}$ is continuous. Observations are i.i.d. over individual i for any (j, a, t). For any individual i, we assume that $(\theta_{i,a,t}, \varepsilon_{i,a,t}, \eta_{i,j,a,t})$ are mutually independent for all test measures j and that each of these variables is also independent of past and future realizations of the other two variables. The idiosyncratic measurement error $\eta_{i,j,a,t}$ is independent over j and t. Since a and t move together for each individual, $\eta_{i,j,a,t}$ is also independent over a, but it need not be identically distributed over ages or time. We normalize $\lambda_{t^*} = 1$ for some year t^*, which effectively measures skills in year t^* log wage units. We also normalize $E(\varepsilon_{i,a,t}) = E(\eta_{i,j,a,t}) = 0$ for all (j, a, t). Identification requires no assumptions about the serial dependence structure for log wage shocks $\varepsilon_{i,a,t}$.

Let $\alpha_{a,t} \equiv E(\theta_{i,a,t})$ and $\bar{\theta}_{i,a,t} := \theta_{i,a,t} - \alpha_{a,t}$ be the de-meaned skill value. Then, we can rewrite log wages as follows:

$$
\begin{aligned}
\ln W_{i,a,t} &= \gamma_t + \lambda_t(\theta_{i,a,t} - \alpha_{a,t} + \alpha_{a,t}) + \varepsilon_{i,a,t} \\
&= \gamma_t + \lambda_t \alpha_{a,t} + \lambda_t \bar{\theta}_{i,a,t} + \varepsilon_{i,a,t} \\
&\equiv \widetilde{\gamma}_{a,t} + \lambda_t \bar{\theta}_{i,a,t} + \varepsilon_{i,a,t},
\end{aligned}
$$

where $\widetilde{\gamma}_{a,t} := \gamma_t + \lambda_t \alpha_{a,t}$. Regressing log wages on interactions of age and time dummies yields consistent estimates of $\tilde{\gamma}_{a,t}$ and log wage residuals $w_{i,a,t} := \lambda_t \bar{\theta}_{i,a,t} + \varepsilon_{i,a,t}$. We work with these residuals below to discuss identification and estimation of returns to skill, λ_t, and the evolution of skill distributions.

4.2.2 Identification

Since the continuous measurement function $\tau_1(\cdot)$ is assumed to be time invariant, observing this same measurement along with at least one other measurement and log wage residuals in multiple periods, we can identify the returns to skill each period, nonparametric (age/cohort-specific) distributions of skills each period, and the dynamics of skills. (We also obtain nonparametric identification for the measurement function $\tau_1(\cdot)$ and the corresponding error distributions $f_{\eta_{1,a,t}}$ for all t.) The identification of skill dynamics is considered both under a general Markov structure and under the AR(1) structure with a fixed effect term. Notice that the latter requires identification of each component separately while the former focuses only on the conditional density of $\theta_{i,a,t}$ given $\theta_{i,a-1,t-1}$. We are careful to note which features of our model can be identified with repeated cross-section data alone and which require panel data.

Returns to Skill, Cross-Sectional Skill Distributions, and Measurements

We begin by discussing identification of returns to skill, cross-sectional skill distributions, test measurement functions, and the distribution of test score measurement errors.

Consider a normalized age and time pair (a^*, t^*), where $\lambda_{t^*} = 1$ and $\alpha_{a^*,t^*} = 0$. Let $c^* = t^* - a^*$ be the corresponding cohort. Applying Theorem 1 in Hu and Schennach (2008), we can identify the distributions $F_{\theta_{a^*,t^*}}(\cdot)$, $F_{\varepsilon_{a^*,t^*}}(\cdot)$, and $F_{\eta_{1,a^*,t^*}}(\cdot)$, as well as the measurement function $\tau_1(\cdot)$ from the joint density of $(w_{i,a^*,t^*}, T_{i,1,a^*,t^*}, ..., T_{i,J,a^*,t^*})$ for $J \geq 2$.[9] Appendix C.1 provides details of the regularity conditions and the identification result.

Now, we consider an arbitrary age and time pair $(a, t) \neq (a^*, t^*)$, rewriting the model as

$$
\begin{aligned}
w_{i,a,t} &= \tilde{\theta}_{i,a,t} + \varepsilon_{i,a,t}, \\
T_{i,1,a,t} &= \tilde{\tau}_{1,a,t}(\tilde{\theta}_{i,a,t}) + \eta_{i,1,a,t} \\
T_{i,j,a,t} &= G_j\left(\tilde{\tau}_{j,a,t}(\tilde{\theta}_{i,a,t}) + \eta_{i,j,a,t}\right) \text{ for } j = 2, \ldots, J,
\end{aligned}
\tag{4.2}
$$

where $\tilde{\theta}_{i,a,t} := \lambda_t \bar{\theta}_{i,a,t} = \lambda_t(\theta_{i,a,t} - \alpha_{a,t})$ and $\tilde{\tau}_{j,a,t}(x) := \tau_j(x/\lambda_t + \alpha_{a,t})$. Notice that $E(\tilde{\theta}_{i,a,t}) = 0$. Using the same arguments as above, we can identify $F_{\tilde{\theta}_{i,a,t}}(\cdot)$, $F_{\varepsilon_{i,a,t}}(\cdot)$, $F_{\eta_{i,1,a,t}}(\cdot)$, and $\tilde{\tau}_{1,a,t}(\cdot)$ from

[9] We make no effort to separately identify $\tau_j(\cdot)$ from $G_j(\cdot)$ for $j = 2, ..., J$, which may require additional assumptions for discrete measures and is not necessary given a single repeated continuous measure $T_{i,1,a,t}$.

the joint density of $(w_{i,a,t}, T_{i,1,a,t}, ..., T_{i,J,a,t})$.

Knowledge of $\tau_1(\cdot)$ from (a^*, t^*) and $\tilde{\tau}_{1,a,t}(\cdot)$ from any other (a, t) identifies λ_t and $\alpha_{a,t}$.[10] To see this, consider two points θ_1 and θ_2 on the support of θ such that $\theta_1 < \theta_2$. Since $\tau_1(\cdot)$ is strictly increasing, $\theta_1 < \theta_2$ implies that $\tilde{\tau}_{1,a,t}(\theta_1) < \tilde{\tau}_{1,a,t}(\theta_2)$. By definition, $\tilde{\tau}_{1,a,t}(\theta_1) = \tau_1(\theta_1/\lambda_t + \alpha_{a,t})$ and $\tilde{\tau}_{1,a,t}(\theta_2) = \tau_1(\theta_2/\lambda_t + \alpha_{a,t})$. Solving this system of equations identifies

$$\begin{aligned}
\lambda_t &= \frac{\theta_1 - \theta_2}{\tau_1^{-1}\left(\tilde{\tau}_{1,a,t}(\theta_1)\right) - \tau_1^{-1}\left(\tilde{\tau}_{1,a,t}(\theta_2)\right)} \\
\alpha_{a,t} &= \frac{\theta_2 \tau_1^{-1}\left(\tilde{\tau}_{1,a,t}(\theta_1)\right) - \theta_1 \tau_1^{-1}\left(\tilde{\tau}_{1,a,t}(\theta_2)\right)}{\theta_2 - \theta_1}.
\end{aligned}$$

Having identified λ_t, $\alpha_{a,t}$, and $F_{\tilde{\theta}_{i,a,t}}(\cdot)$, we can then identify $F_{\theta_{i,a,t}}(\theta) = F_{\tilde{\theta}_{i,a,t}}(\lambda_t(\theta - \alpha_{a,t}))$.

We emphasize that none of the identification results thus far require panel data. Identification of the returns to skills and cross-sectional distributions of skills over time can be achieved with repeated cross-section data. While we have explicitly considered the case with $J \geq 2$ repeated measures each period, it is clear that only a single continuous measurement must be repeated every period. Other measurements can differ from period to period.[11]

Skill Processes: General Approach

Without any additional assumptions on $\theta_{i,a,t}$ and $\varepsilon_{i,a,t}$, we can identify their serial dependence structure with panel data on our continuous measure $T_{i,1,a,t}$ (given what we have already identified in Section 4.2.2). To see this, consider two time periods at (a^*, t^*) and (a, t). For example, $a = a^* + 1$ and $t = t^* + 1$. Let $\varphi_{\eta_{i,1,a^*,t^*}, \eta_{i,1,a,t}}(x_1, x_2) := E[\exp(-i(x_1\eta_{i,1,a^*,t^*} + x_2\eta_{i,1,a,t}))]$ be a characteristic function of $(\eta_{i,1,a^*,t^*}, \eta_{i,1,a,t})$ and define characteristic functions for other random

[10] Note that identification of λ_t and $\alpha_{a,t}$ further identifies time effects in log wage equations, γ_t, from $\tilde{\gamma}_{a,t}$.

[11] Even more generally, as long as each period of data contains at least two independent measures, identification can be achieved with overlapping periods that contain a repeated continuous measurement (e.g., one continuous measurement over periods 1 and 2 with a different continuous measurement over periods 2 and 3, etc.). The same continuous measurement need not be available over the entire sample period.

variables similarly. Note that

$$
\begin{aligned}
\varphi_{T_{i,1,a^*,t^*},T_{i,1,a,t}}(x_1,x_2) &= \varphi_{\tau_1(\theta_{i,a^*,t^*}),\tau_1(\theta_{i,a,t})}(x_1,x_2)\cdot\varphi_{\eta_{i,1,a^*,t^*},\eta_{i,1,a,t}}(x_1,x_2)\\
\varphi_{\tau_1(\theta_{i,a^*,t^*}),\tau_1(\theta_{i,a,t})}(x_1,x_2) &= \frac{\varphi_{T_{i,1,a^*,t^*},T_{i,1,a,t}}(x_1,x_2)}{\varphi_{\eta_{i,1,a^*,t^*},\eta_{i,1,a,t}}(x_1,x_2)}\\
&= \frac{\varphi_{T_{i,1,a^*,t^*},T_{i,1,a,t}}(x_1,x_2)}{\varphi_{\eta_{i,1,a^*,t^*}}(x_1)\varphi_{\eta_{i,1,a,t}}(x_2)},
\end{aligned}
$$

where the last equality holds by the time-independence of $\eta_{i,1,a,t}$ distributions. Since we know all distributions on the right hand side, we can identify the joint distribution of $(\tau_1(\theta_{i,a^*,t^*}),\tau_1(\theta_{i,a,t}))$. Then, the joint distribution of $(\theta_{i,a^*,t^*},\theta_{i,a,t})$ is identified by

$$
\begin{aligned}
F_{\theta_{i,a^*,t^*},\theta_{i,a,t}}(x_1,x_2) &= P(\theta_{i,a^*,t^*}\le x_1,\theta_{i,a,t}\le x_2)\\
&= P(\tau_1(\theta_{i,a^*,t^*})\le\tau_1(x_1),\tau_1(\theta_{i,a,t})\le\tau_1(x_2)),
\end{aligned}
$$

where the final expression is identified since we know both $\tau_1(\cdot)$ and $\varphi_{\tau_1(\theta_{i,a^*,t^*}),\tau_1(\theta_{i,a,t})}(\cdot,\cdot)$. Therefore, we can construct the conditional density $f_{\theta_{i,a,t}|\theta_{i,a^*,t^*}}$ from the joint distribution and identify the serial dependence of $\theta_{i,a,t}$. Similarly, we can identify the joint distribution of $(\varepsilon_{i,a^*,t^*},\varepsilon_{i,a,t})$ by noting that

$$
\begin{aligned}
\varphi_{\varepsilon_{i,a^*,t^*},\varepsilon_{i,a,t}}(x_1,x_2) &= \frac{\varphi_{w_{i,a^*,t^*},w_{i,a,t}}(x_1,x_2)}{\varphi_{\theta_{i,a^*,t^*},\widetilde{\theta}_{i,a,t}}(x_1,x_2)}\\
&= \frac{\varphi_{w_{i,a^*,t^*},w_{i,a,t}}(x_1,x_2)}{\varphi_{\theta_{i,a^*,t^*},\theta_{i,a,t}}(x_1,\lambda_t x_2)\times\exp(-i\alpha_{a,t}\lambda_t x_2)},
\end{aligned}
$$

where we already know $\alpha_{a,t}$, λ_t, and the two joint distributions on the right hand side. The serial dependence of $(\varepsilon_{i,a,^*,t^*},\varepsilon_{i,a,t})$ follows immediately.

Skill Process: AR(1) and Fixed Effect

We now consider the identification problem for the skill process when skills are decomposed into the following three components: (i) a systematic lifecycle skill growth component, which can differ freely across cohorts, $\alpha_{a,t}$; (ii) an individual fixed effect ψ_i; and (iii) an AR(1) component $\phi_{i,a,t}$. Thus, the skill process can be written as follows:

$$\begin{aligned} \theta_{i,a,t} &= \alpha_{a,t} + \psi_i + \phi_{i,a,t}, \\ \phi_{i,a,t} &= \rho\phi_{i,a-1,t-1} + \nu_{i,a,t}, \end{aligned} \tag{4.3}$$

where $\nu_{i,a,t}$ is independent over t. We further assume that ψ_i is independent of $\phi_{i,a,t}$ and $\nu_{i,a,t}$ for all (a,t). We normalize $E(\psi_i) = E(\nu_{i,a,t}) = E(\phi_{i,a,t}) = 0$ for all (a,t), which implies $\alpha_{a,t} = E(\theta_{i,a,t})$ as before. We normalize $\alpha_{a^*,t^*} = 0$ for some (a^*, t^*).[12]

Notice that we have already identified the returns to skills and age/cohort- and time-specific skill distributions using only repeated cross-sections of log wages and skill measures (see Section 4.2.2). To identify the basic components in equation (4.3), we need panel data from (at least) three periods, t, $t+1$, and $t+2$.

First, we can use similar arguments as in Section 4.2.2 to identify the joint distribution of $(\theta_{i,a,t}, \theta_{i,a+1,t+1}, \theta_{i,a+2,t+2})$. Then, we can construct the following moment conditions from equation (4.3):

$$\begin{aligned} Var(\theta_{i,a,t}) &= Var(\psi_i) + Var(\phi_{i,a,t}) \\ Cov(\theta_{i,a,t}, \theta_{i,a+1,t+1}) &= Var(\psi_i) + \rho Var(\phi_{i,a,t}) \\ Cov(\theta_{i,a,t}, \theta_{i,a+2,t+2}) &= Var(\psi_i) + \rho^2 Var(\phi_{i,a,t}). \end{aligned}$$

Solving this system of equations, we can identify ρ as

$$\rho = 1 - \frac{Var(\theta_{i,a,t}) - Cov(\theta_{i,a,t}, \theta_{i,a+2,t+2})}{Var(\theta_{i,a,t}) - Cov(\theta_{i,a,t}, \theta_{i,a+1,t+1})},$$

where moments on the right hand side are identified from the joint density of $(\theta_{i,a,t}, \theta_{i,a+1,t+1}, \theta_{i,a+2,t+2})$.

Next, we can identify cohort- and time-specific distributions for skill shocks and cohort-specific distributions for the fixed effects. We fix a cohort and let t_0 be the first year it is observed. First, since we already know $\alpha_{a,t}$ and the joint distribution $F_{\theta_{i,a+1,t+1},\theta_{i,a,t}}$ for all (a,t),

[12] It is not necessary to use the same t^* here as used for normalizing $\lambda_{t^*} = 1$; however, we do so in this section to simplify the exposition.

we identify the left hand side of the following two equations:

$$\theta_{i,a,t_0} - \alpha_{a,t_0} = \phi_{i,a,t_0} + \psi_i$$
$$\frac{(\theta_{i,a+1,t_0+1} - \alpha_{a+1,t_0+1}) - (\theta_{i,a,t_0} - \alpha_{a,t_0})}{\rho - 1} = \phi_{i,a,t_0} + \frac{\nu_{i,a+1,t_0+1}}{\rho - 1}.$$

Since ϕ_{i,a,t_0}, ψ_i, and $\nu_{i,a+1,t_0+1}$ are mutually independent, we can identify their distributions by applying Kotlarski's Lemma (Kotlarski, 1967). Second, we identify the distribution of $\phi_{i,a,t}$ sequentially for all $t \geq t_0 + 1$ from $\phi_{i,a+1,t+1} - \phi_{i,a,t} = (\theta_{i,a+1,t+1} - \alpha_{a+1,t+1}) - (\theta_{i,a,t} - \alpha_{a,t})$ by applying standard deconvolution arguments, since we know the distribution of the right hand side and that of ϕ_{i,a,t_0}. Finally, we identify the distribution of $\nu_{i,a,t}$ for $t \geq t_0 + 1$ from $\phi_{i,a,t} = \rho\phi_{i,a-1,t-1} + \nu_{i,a,t}$ by applying the deconvolution arguments again.

4.2.3 Estimation

We can estimate the full model nonparametrically, e.g. the sieve maximum likelihood estimator as in Hu and Schennach (2008). However, it is quite challenging in practice as the objective function involves multiple integration over many unobservables.

We mitigate this computational difficulty by developing a three-step estimation procedure based on the identification strategy outlined earlier. First, one can estimate the measurement functions $\tau_j(\cdot)$ and the distribution of skills θ_{i,a,t^*} from the cross-sectional observations of $\{w_{i,a,t^*}, T_{i,1,a,t^*}, \ldots, T_{i,J,a,t^*}\}$ for all ages a at time t^*. Second, repeated cross-sectional observations $\{w_{i,a,t}, T_{i,1,a,t}, \ldots, T_{i,J,a,t}\}$ at $(a,t) \neq (a^*, t^*)$ can be used along with the estimated $\widehat{\tau}_1(\cdot)$ (from Step 1) to estimate the skill return λ_t and skill distribution of $\theta_{i,a,t}$ for all (a,t). Finally, panel data (if available), can be used to estimate the dynamics of skill distributions.

The estimation of skill dynamics (and other features of the model) becomes much simpler when any continous measurement function $\tau_j(\cdot)$ estimated in Step 1 turns out to be linear. We discuss this simpler estimation approach at the end of this section.

Step 1: Estimating $\tau_j(\cdot)$ and $f_{\theta_{a,t^*}}(\cdot)$

We can estimate $\tau_j(\cdot)$ functions and the skill distributions $f_{\theta_{a,t^*}}(\cdot)$ for all ages using cross-sectional data at time t^*. Normalizing $\lambda_{t^*} = 1$ and $\alpha_{a^*,t^*} \equiv E(\theta_{i,a^*,t^*}) = 0$, we consider a non-parametric maximum likelihood estimation (NPMLE) approach by using flexible functional form and distributional assumptions (e.g. polynomials for measurement functions, mixtures of normal distributions for densities, or sieve estimation using Hermite polynomials). The complexity of the underlying parameter space can be adjusted depending on the model structure and the sample size.

Let $f_{w_{a^*,t^*},T_{1,,a^*,t^*},\ldots,T_{J,a^*,t^*}}$ be the joint density function of $(w_{a^*,t^*,}, T_{1,a^*,t^*}, \ldots, T_{J,a^*,t^*})$. Since observations are i.i.d. over individuals, we drop the subscript i unless it causes any confusion. For simplicity, we assume that $G(\cdot)$ is an identity function and all T_j are continuous.[13] The independence assumption among $(\theta_{a^*,t^*}, \varepsilon_{a^*,t^*}, \eta_{1,a^*,t^*}, \ldots, \eta_{J,a^*,t^*})$ implies that

$$\begin{aligned}
&f_{w_{a^*,t^*},T_{1,a^*,t^*},\ldots,T_{J,a^*,t^*}}(w, T_1, \ldots, T_J) \\
&= \int_\Theta f_{\varepsilon_{a^*,t^*}}\left(w - \theta; \boldsymbol{\beta}_{\varepsilon_{a^*,t^*}}\right) \\
&\quad \times f_{\eta_{1,a^*,t^*}}\left(T_1 - \tau_1(\theta; \boldsymbol{\beta}_{\tau_1}); \boldsymbol{\beta}_{\eta_{1,a^*,t^*}}\right) \times \cdots \times f_{\eta_{J,a^*,t^*}}\left(T_J - \tau_J(\theta; \boldsymbol{\beta}_{\tau_J}); \boldsymbol{\beta}_{\eta_{J,a^*,t^*}}\right) \\
&\quad \times f_{\theta_{a^*,t^*}}\left(\theta; \boldsymbol{\beta}_{\theta_{a^*,t^*}}\right) d\theta,
\end{aligned} \tag{4.4}$$

where $\boldsymbol{\beta}_x$ for a generic x denotes a vector of the parameters or the polynomial coefficients for the unknown distribution or function. Recall that all density functions in (4.4) should satisfy the mean zero restriction. The above density function can be used to form the log-likelihood function. Let $\boldsymbol{\beta}_{a^*,t^*} \equiv (\{\boldsymbol{\beta}_{\tau_j}, \boldsymbol{\beta}_{\eta_{j,a^*,t^*}}\}_{j=1}^J, \boldsymbol{\beta}_{\varepsilon_{a^*,t^*}}, \boldsymbol{\beta}_{\theta_{a^*,t^*}})$ be the stacked vector of all unknown parameters. It can be estimated by

$$\widehat{\boldsymbol{\beta}}_{a^*,t^*} = \underset{\beta_{a^*,t^*} \in \mathcal{B}}{\arg\max} \frac{1}{|\mathcal{I}_{c^*}|} \sum_{i \in \mathcal{I}_{c^*}} \log f_{w_{a^*,t^*},T_{1,a^*,t^*},\ldots,T_{J,a^*,t^*}}(w_{i,a^*,t^*,}, T_{i,1,a^*,t^*}, \ldots, T_{i,J,a^*,t^*}; \boldsymbol{\beta}_{a^*,t^*}), \tag{4.5}$$

where $\mathcal{I}_{c^*}$ is the subset of individuals who belong to cohort $c^* = t^* - a^*$ and $|\mathcal{I}|$ is the number of elements in set $\mathcal{I}$. Of particular interest to us are the estimates $\{\widehat{\boldsymbol{\beta}}_{\tau_j}\}_{j=1}^J$ and $\boldsymbol{\beta}_{\theta_{a^*,t^*}}$, which give

[13] When T_j for $j \geq 2$ include a discrete measure, we can replace $f_{\eta_{j,a^*,t^*}}$ with a proper discrete probability mass function. For example, see Appendix C.2

us the estimated measurement functions $\{\widehat{\tau}_j(\cdot)\}_{j=1}^J$ and skill distributions $\widehat{f_{\theta_{a^*,t^*}}}(\cdot)$, respectively.[14] We can repeat the estimation procedure in (4.5) for each cohort (a, t^*) and estimate the skill distribution $f_{\theta_{a,t^*}}$ of cohort $c = t^* - a$ at time t^*.

We can also increase the estimation efficiency of $\widehat{\tau}_j(\cdot)$ by including all cohorts in a single optimization procedure. For any cohort $c = a - t^*$, define the cohort specific objective function:

$$Q(\boldsymbol{\beta}_{a,t^*}) \equiv \frac{1}{|\mathcal{I}_c|} \sum_{i \in \mathcal{I}_c} \log f_{w_{a,t^*},T_{1,a,t^*},\ldots,T_{J,a,t^*}}(w_{i,a,t^*}, , T_{i,1,a,t^*}, \ldots, T_{i,J,a,t^*}; \boldsymbol{\beta}_{a,t^*}). \tag{4.6}$$

Let $\boldsymbol{\beta}_{t^*} \equiv \{\boldsymbol{\beta}_{a,t^*}\}_{a \in \mathcal{A}}$ be the stacked parameter vector, where $\mathcal{A}$ is an index set of all different ages (cohorts) at time t^*. Then, the parameter of interest as well as some nuisance parameters can be estimated by

$$\widehat{\boldsymbol{\beta}}_{t^*} = \arg\max_{\boldsymbol{\beta}_{t^*} \in \mathcal{B}^{|\mathcal{A}|}} \sum_{a \in \mathcal{A}} Q(\boldsymbol{\beta}_{a,t^*})$$

Notice that we normalize the mean of the skill distribution only in cohort c^*, so the density function $f_{\theta_{a,t^*}}$ for $a \neq a^*$ is allowed to have a non-zero mean. This estimation approach will be more efficient if the measurement errors $\{\varepsilon_{a,t^*}\}_{a \in \mathcal{A}}$ and $\{\eta_{a,t^*}\}_{a \in \mathcal{A}}$ have identical distributions across different cohorts.

Finally, we note that if one is simply interested in determining whether any of the $\tau_j(\cdot)$ functions is linear (or only the density of skills in period t^* across all ages/cohorts is desired), then the likelihood in equation (4.4) can alternatively be written in terms of $f_{\theta_{t^*}}$ (with parameters $\boldsymbol{\beta}_{\theta_{t^*}}$), normalizing this density to be mean zero. Indeed, this is the approach we take below in determining that one of our measures is linear in skills.

Step 2: Estimating λ_t and $f_{\theta_{a,t}}(\cdot)$

Now, we discuss estimation of skill returns, λ_t, and the skill distributions, $f_{\theta_{a,t}}(\cdot)$, with additional cross-sectional data at time $t \neq t^*$. Embedding the estimated measurement function $\widehat{\tau}_1(\cdot)$ and

[14] Appendix C.2 provides expressions for likelihoods assuming mixtures of normal distributions for errors and discusses the case of both discrete and continuous measurements, T_{j,a^*,t^*}. Details on estimation of standard errors for $\widehat{\boldsymbol{\beta}}_{a^*,t^*}$ are also provided in Appendix C.2.

the unknown skill return λ_t, we can write the density function at time t as

$$
\begin{aligned}
& f_{w_{a,t},T_{1,a,t},\ldots,T_{J,a,t}}(w, T_1, \ldots, T_J) \\
& = \int_\Theta f_{\varepsilon_{a,t}}\left(w - \lambda_t\theta; \boldsymbol{\beta}_{\varepsilon_{a,t}}\right) \\
& \times f_{\eta_{1,a,t}}\left(T_1 - \widehat{\tau}_1(\theta); \boldsymbol{\beta}_{\eta_{1,a,t}}\right) \times f_{\eta_{2,a,t}}\left(T_2 - \tau_2(\theta; \boldsymbol{\beta}_{\tau_2}); \boldsymbol{\beta}_{\eta_{2,a,t}}\right) \times \cdots \times f_{\eta_{J,a,t}}\left(T_J - \tau_J(\theta; \boldsymbol{\beta}_{\tau_J}); \boldsymbol{\beta}_{\eta_{J,a,t}}\right) \\
& \times f_{\theta_{a,t}}\left(\theta; \boldsymbol{\beta}_{\theta_{a,t}}\right) d\theta.
\end{aligned}
\tag{4.7}
$$

Define $Q(\lambda_t, \tilde{\boldsymbol{\beta}}_{a,t})$ as in (4.6) by adding the skill return parameter λ_t and let $\tilde{\boldsymbol{\beta}}_t \equiv \{\tilde{\boldsymbol{\beta}}_{a,t}\}_{a\in\mathcal{A}}$, where we drop $\boldsymbol{\beta}_{\tau_1}$ from each $\boldsymbol{\beta}_{a,t}$ as we already plugged in the estimate from Step 1. If any measure observed at time t^* is repeated at time t, we can replace it with $\widehat{\tau}_j(\cdot)$ and drop the relevant parameters. The parameter set is further simplified when $\varepsilon_{a,t}$ and $\eta_{j,a,t}$ are age/time-invariant since we can plug-in the corresponding estimates from Step 1. Then, we can estimate the skill return and other underlying parameters at time t by

$$
\left(\widehat{\lambda}_t, \widehat{\boldsymbol{\beta}}_t\right) = \underset{(\lambda_t, \tilde{\boldsymbol{\beta}}_t) \in \Lambda \times \tilde{\mathcal{B}}^{|\mathcal{A}|}}{\arg\max} \sum_{a \in \mathcal{A}} Q(\lambda_t, \tilde{\boldsymbol{\beta}}_{a,t}).
$$

Once we obtain estimates $\widehat{\boldsymbol{\beta}}_{\theta_{a,t}}$ for skill distributions $f_{\theta_{a,t}}(\cdot)$, we can estimate $\alpha_{a,t}$ for all $a \in \mathcal{A}$ at time t. In addition, we can estimate time effects in the wage equation by $\widehat{\gamma}_t = \widetilde{\gamma} - \widehat{\lambda}_t\widehat{\alpha}_{a,t}$, where we have already estimated all the components on the right hand side.

Step 3: Estimating Skill Dynamics

We discuss estimation of skill dynamics for two different cases, both using panel data. First, for a general Markov skill process, we can apply the same idea as above to estimate its dynamics using any repeated continuous measure of skills. Given the estimated elements of the model, we can write the joint density function of repeated continuous measure j at time t and $t+1$ as follows:

$$\begin{aligned} f_{T_{j,a,t},T_{j,a+1,t+1}}(T_t, T_{t+1}; \boldsymbol{\beta}_{\theta_{a,t},\theta_{a+1,t+1}}) = \iint_{\Theta\times\Theta} & \widehat{f}_{\eta_{j,a,t}}(T_t - \widehat{\tau}_j(\theta_t))\widehat{f}_{\eta_{j,a+1,t+1}}(T_{t+1} - \widehat{\tau}_j(\theta_{t+1})) \\ & \times f_{\theta_{a,t},\theta_{a+1,t+1}}(\theta_t, \theta_{t+1}; \boldsymbol{\beta}_{\theta_{a,t},\theta_{a+1,t+1}}) d\theta_t d\theta_{t+1}, \end{aligned} \tag{4.8}$$

where $f_{\theta_{a,t},\theta_{a+1,t+1}}(\theta_t, \theta_{t+1}; \boldsymbol{\beta}_{\theta_{a,t},\theta_{a+1,t+1}})$ is the joint density function of $(\theta_{a,t}, \theta_{a+1,t+1})$. The measure j specific objective function can be defined as

$$Q_j(\boldsymbol{\beta}_{\theta_{a,t},\theta_{a+1,t+1}}) = \frac{1}{|\mathcal{I}_c|} \sum_{i\in\mathcal{I}_c} \log f_{T_{j,a,t},T_{j,a+1,t+1}}(T_{i,j,a,t}, T_{i,j,a+1,t+1}; \boldsymbol{\beta}_{\theta_{a,t},\theta_{a+1,t+1}}).$$

Then, the parameters for the joint density function are estimated by

$$\widehat{\boldsymbol{\beta}}_{\theta_{a,t},\theta_{a+1,t+1}} = \underset{\boldsymbol{\beta}_{\theta_{a,t},\theta_{a+1,t+1}}\in\mathcal{B}_\theta}{\arg\max}\; Q_j(\boldsymbol{\beta}_{\theta_{a,t},\theta_{a+1,t+1}}).$$

Once we have estimated the joint density function, the dynamics of the skill process follow immediately from the conditional density function.

Second, if the skill process follows the AR(1) with fixed effect structure as in (4.3), the parameters for this process can be estimated following a similar strategy as above using a modified version of equation (4.8) that incorporates an additional time period to estimate the joint density $f_{\theta_t,\theta_{t+1},\theta_{t+2}}$ where the parameters for this density, $\boldsymbol{\beta}_{\theta_{a,t},\theta_{a+1,t+1},\theta_{a+2,t+2}}$ include the relevant cohort $c = t - a$ distribution for ψ_i and parameters of the AR(1) process (ρ and parameters determining distributions for $\phi_{i,a,t}$, $\nu_{a+1,t+1}$, and $\nu_{a+2,t+2}$).

Estimation of Skill Distributions, Skill Dynamics, and Returns to Skill when a Linear Measurement is Available

In our empirical context, one of the measurements, say $T_{1,a,t}$, is determined to be linear in skills from the estimation procedure described in Section 4.2.3. We use this information to facilitate estimation of the returns to skill over time and the evolution and dynamics of skills assuming the special case where $\theta_{i,a,t}$ follows the AR(1) plus fixed effect process described in equation (4.3).

Using the known linear measurement $T_{i,1,a,t} = \beta_{1,0} + \beta_{1,1}\theta_{i,a,t} + \eta_{i,1,a,t}$, our model for log wage residuals and the skill measurement can be written in terms of de-meaned skills:

$$\begin{aligned} w_{i,a,t} &= \lambda_t \bar{\theta}_{i,a,t} + \varepsilon_{i,a,t}, \\ \bar{\theta}_{i,a,t} &= \psi_i + \phi_{i,a,t}, \\ \phi_{i,a,t} &= \rho \phi_{i,a-1,t-1} + \nu_{i,a,t}, \\ T_{i,1,a,t} &= (\beta_{1,0} + \beta_{1,1}\alpha_{a,t}) + \beta_{1,1}\bar{\theta}_{i,a,t} + \eta_{i,1,a,t}. \end{aligned}$$

These imply the following covariances for (a, t):

$$\begin{aligned} Cov(w_{a,t}, T_{1,a+k,t+k}) &= \lambda_t \beta_{1,1} \left[Var(\psi | t-a) + \rho^k Var(\phi_{a,t}) \right], && \text{for } k \geq 0 \\ Cov(T_{1,a,t}, T_{1,a+k,t+k}) &= \beta_{1,1}^2 \left[Var(\psi | t-a) + \rho^k Var(\phi_{a,t}) \right], && \text{for } k \geq 1 \\ Cov(T_{1,a,t}, w_{a+k,t+k}) &= \lambda_{t+k} \beta_{1,1} \left[Var(\psi | t-a) + \rho^k Var(\phi_{a,t}) \right], && \text{for } k \geq 1. \end{aligned}$$

Assuming the distribution of skill shocks depends only on time, we define $\sigma^2_{\nu_t} \equiv Var(\nu_{a,t})$ for all (a, t) and can write

$$Var(\phi_{a,t}) = \rho^{2(t-t_1)} Var(\phi_{a-(t-t_1),t_1}) + \sum_{s=t_1+1}^{t} \rho^{2(t-s)} \sigma^2_{\nu_s}, \qquad \forall t \geq t_1 + 1,$$

where t_1 is the initial period of observation. As discussed earlier, we normalize $\lambda_{t^*} = 1$ and $\alpha_{a^*,t^*} = 0$. With these assumptions, the generalized methods of moments (GMM) can be used to jointly estimate the time-varying returns to skill (λ_t), autocorrelation for skill shocks (ρ), variances of initial skills by cohort ($Var(\psi|t-a)$ for all observed cohorts), initial variances of the persistent skill shock ($Var(\phi_{a,t_1})$ for cohorts observed in initial period t_1 and $Var(\phi_{a_1,t})$ for cohorts entering the sample at age a_1 at later dates), time-varying skill shock variances ($\sigma^2_{\nu_t}$), and the measurement function parameters ($\beta_{1,0}, \beta_{1,1}$). Further details on estimation and calculation of the standard errors are provided in Appendix C.3.

4.3 HRS Data

We use data from the Health and Retirement Study (HRS), a national U.S. panel survey of individuals over age 50 and their spouses.[15] It consists of seven cohorts with the initial cohort first interviewed in 1992. New cohorts of individuals were added in 1993, 1998, 2004, 2010, and 2016.[16] The survey has been fielded every two years since 1992 and it provides information about demographics, income, and cognition, making it ideal data for the purpose of our study. Because one of the cognitive tests (word recall) in 1992 and 1994 differs from the later years, we use data collected from 1996 to 2016.[17]

The HRS records the respondent's and spouse's wage rates if they are working at the time of the interview. We use the hourly wage rate, deflating nominal values to 1996 dollars using the Consumer Price Index.[18] The HRS also provides various cognitive functioning measures. We use four measures in our estimation: word recall, serial 7's, quantitative reasoning, and retrieval fluency. Table 4.1 provides a brief summary of these measures. The word recall test evaluates the memory of the respondents by reading a list of 10 words and asking them to recall immediately (immediate recall) and after a delay of about 5 minutes (delayed recall). We sum up the number of words the respondent recalled in the two tasks and obtain a score of 21 different values. The serial 7's test asks the respondent to subtract 7 from the previous number, starting with 100 for five trials. This test score is the number of trials that the respondent answered correctly, and it has 6 different values. Quantitative reasoning consists of three simple arithmetic questions assessing the numeracy of the respondent. We construct a test score based on the answers and the resulting score ranges from 0 to 4. The retrieval fluency test asks the

[15]More precisely, the sample does include some individuals age 50. For example, someone from the original cohort (born in 1931-1941) who was born late in 1941 may have been age 50 at the date of their first interview in 1992 if they were interviewed earlier in the calendar year.

[16]The HRS sample was built up over time. The initial cohort consisted of persons born between 1931 and 1941 (aged 51 to 61 at first interview in 1992). The Asset and Health Dynamics Among the Oldest Old (AHEAD) cohort, born before 1924 was added in 1993 and interviewed in 1993, 1995, and biennially from 1998 forward. In 1998, two new cohorts were enrolled: the Children of the Depression (CODA) cohort, born 1924 to 1930, and the War Baby (WB) cohort, born 1942 to 1947. Early Baby Boomer (EBB, born 1948 to 1953) cohort was added in 2004, Mid Baby Boomer (MBB, born 1954 to 1959) cohort was added in 2010, and Late Baby Boomer (LBB, born 1960 to 1965) cohort was added in 2016. In addition to respondents from eligible birth years, the survey interviewed the spouses of married respondents or the partner of a respondent, regardless of age.

[17]The word recall test contains a list of 20 words in 1992 and 1994, while it contains only 10 words in later years.

[18]https://www.bls.gov/cpi/research-series/home.htm#CPI-U-RS20Data

respondents to name as many animals as they can in 60 seconds. The test score is the total number of correct answers, ranging from 0 to 90. Additional details about the measures and the construction of other key variables are provided in Appendix C.4.

Table 4.1: Description of Cognitive Measures

	Meant to measure	Number of values	Available years
Word recall	Memory	21 (0-20)	1996-2016
Serial 7's	Numeracy	6 (0-5)	1996-2016
Quantitative reasoning	Numeracy	5 (0-4)	2002-2016
Retrieval fluency	Fluency	91 (0-90)	2010-2016

Our sample is restricted to age-eligible (i.e. born in eligible years when first interviewed) men. We use observations when men are ages 50-70 if their potential labor market experience is between 30 and 50 years.[19] We trim the top and bottom 1% of all wages within year by college- vs. non-college-educated status and 10-year experience cells. In estimation, we use non-imputed wages and cognitive measures only. The sample contains 9,848 individuals and 37,518 person-year observations.

Our sample consists of 64% white, 18% black, 13% Hispanic, and 5% other races with an average age of 60 years. We create five education categories based on years of education: 0-11 years (less than high school graduate), 12 years (high school graduate), 13-15 years (some college), 16 years (college graduate), and 17 or more years (above college). In our sample, 20% had less than 12 years of schooling, 30% had 12 years of schooling, 24% had some college, 14% completed college, and 13% had more than 16 years of schooling. Table 4.2 shows the mean and the standard deviation of the cognitive scores and log hourly wage, along with the correlation between these variables. The correlations between test scores range from nearly 0.3 to 0.5, with the highest correlation between Serial 7's and quantitative reasoning. Retrieval fluency has the lowest correlation with log wages (0.214), while quantitative reasoning has the highest (0.303).

[19]Potential experience is defined as age minus 6 minus years of schooling.

Table 4.2: Mean, standard deviation (S.D.), and correlations of cognitive scores and log wages

	Number of obs	Mean	S.D.	Correlations				
Word recall	35,747	10.31	3.14	1.000				
Serial 7's	35,859	3.94	1.45	0.311	1.000			
Quantitative reasoning	11,789	2.07	1.25	0.365	0.494	1.000		
Retrieval fluency	6,800	18.46	7.20	0.295	0.285	0.335	1.000	
Log hourly wage	20,796	2.73	0.69	0.225	0.220	0.303	0.214	1.000

4.4 Estimation Results

4.4.1 Cross-Sectional Results for Measurement Functions

As discussed in Section 4.2.3, we use data from a single year, $t^* = 2010$, to estimate measurement functions $\tau_j(\cdot)$, as well as $F_{\theta_{t^*}}(\cdot)$, $F_{\varepsilon_{t^*}}(\cdot)$, and $F_{\eta_{j,t^*}}(\cdot)$. We use data from 2010, because this is the only year that all four cognitive measures we consider were recorded for every respondent.[20] We further restrict the sample to have non-missing wages and at least one non-missing cognitive measure in 2010. The sample size for this 2010 analysis is 1,980.

We estimate the parameters by maximum likelihood as described in Section 4.2.3, normalizing $\lambda_{t^*} = 1$. We treat word recall (T_1) and retrieval fluency (T_2) as continuous measures, assuming both $\tau_1(\cdot)$ and $\tau_2(\cdot)$ are polynomial functions. In practice, we use likelihood ratio tests to determine the polynomial degree for each measure. We treat the serial 7's (T_3) and quantitative reasoning (T_4) scores as discrete measurements generated from ordered probits with latent index functions $\tau_j(\theta_{i,t}) + \eta_{i,j,t}$.[21]

We estimate the model for three cases. In Case 1, we assume that log wage shocks ε_t, continuous measurement errors (η_1, η_2), and unobserved skills θ_t are all normally distributed. Case 2 assumes that skill θ_t is distributed as a mixture of two normal distributions, while ε_t and (η_1, η_2) are all normally distributed. Finally, Case 3 is the most general, assuming both

[20] In other waves, either one or more of the cognitive tests were not administered or some of the testes were only administered to new interviewees and/or re-interviewees ages 65 or older.

[21] For each of the discrete measures, we have $K_j = 5$ or 6 choices, and we estimate linear terms (intercept terms are normalized to zero) for the $\tau_j(\cdot)$ functions along with $K_j - 1$ cutoff parameters. See Appendix C.2 for details.

skill θ_t and wage shocks ε_t are distributed as mixtures of two normal distributions, while the measurement errors (η_1, η_2) are normally distributed. While we assume the distributions of measurement errors are time invariant, we allow the distributions for skills and log wage shocks to vary freely over time.

For each case, we estimate different specifications by increasing the degree of polynomials for τ_1 and τ_2, starting from a linear specification until the model cannot be improved further, as determined by likelihood ratio tests. Then, the "best" specifications from each of the three cases are compared to determine the "best" overall specification, again using the likelihood ratio test. Table 4.3 reports the log-likelihood associated with several specifications and the three cases, along with likelihood ratio test statistics and p-values. Based on the likelihood ratio tests, the "best" overall specification allows both θ_t and ε_t to be distributed as mixtures of normal distributions. Furthermore, $\tau_1(\cdot)$ (word recall) is linear in skill, while $\tau_2(\cdot)$ (retrieval fluency) is a polynomial of degree 7 in skill. Table 4.4 reports parameter estimates and standard errors for this preferred specification.

4.4.2 Results for Skill Distributions, Skill Dynamics, and Returns to Skill

The fact that word recall scores are linear in skill is convenient and enables us to use the relatively simple approach described in Section 4.2.3 to estimate the returns to skill over time, evolution of skill distributions over time, and the dynamics of skills. To do so, we use the panel nature of the HRS from 1996-2016 (11 biennial surveys), further restricting the sample to include men ages 52-65. There are very few observations outside of that age range for the years we examine.

Since we only observe wages for those who are working at the time of the survey, there are natural concerns about the implications of selection for any covariance moments that include log wage residuals. (Fortunately, test scores are available regardless of work status.) To explore the potential implications of selection, we consider four different sampling schemes for this analysis:

1. "Full" Sample: This includes covariances for log wage residuals and/or test scores whenever they are available. Therefore, covariances using log wage residuals are only calcu-

Table 4.3: First Step Estimation (Selected Specifications)

Model	$\tau_1(\cdot)$	$\tau_2(\cdot)$	Log-likelihood	Compared	p-value
A. Both ε_t and θ_t are normal					
1	linear	linear	-18546.93		
2	linear	7th	-18500.87	2 vs. 1	0.0000
3	linear	8th	-18500.46	3 vs. 2	0.3615
4	quadratic	7th	-18498.71	4 vs. 2	0.0374
5	cubic	7th	-18497.37	5 vs. 4	0.1021
6	quadratic	8th	-18498.15	6 vs. 4	0.2914
B. ε_t is normal. θ_t is mixture of two normals.					
7	linear	linear	-18535.97		
8	linear	7th	-18487.42	8 vs. 7	0.0000
9	linear	8th	-18486.80	9 vs. 8	0.2680
10	quadratic	7th	-18487.25	10 vs. 8	0.5652
C. Both ε_t and θ_t are mixture of two normals					
11	linear	linear	-18504.07		
12	linear	7th	-18455.07	12 vs. 11	0.0000
13	linear	8th	-18454.50	13 vs. 12	0.2871
14	quadratic	7th	-18454.92	14 vs. 12	0.5891
Comparing A vs. B vs. C					
				8 vs. 4	0.0000
				12 vs. 8	0.0000

Note: η_1 and η_2 are normally distributed.

Table 4.4: First Step Estimation (Selected Specifications)

Description	Symbol	Value	Standard Error
Skill Function			
Word recall			
	$\beta_{1,0}$	10.38	0.06
	$\beta_{1,1}$	5.50	0.49
	$\sigma^2_{\eta_1}$	6.57	0.26
Retrieval fluency			
	$\beta_{2,0}$	18.02	0.56
	$\beta_{2,1}$	16.94	4.51
	$\beta_{2,2}$	38.30	35.75
	$\beta_{2,3}$	-58.50	73.66
	$\beta_{2,4}$	-312.69	296.94
	$\beta_{2,5}$	62.05	314.11
	$\beta_{2,6}$	728.30	693.87
	$\beta_{2,7}$	426.80	513.29
	$\sigma^2_{\eta_2}$	36.43	1.66
Serial 7's			
	$\beta_{3,1}$	3.60	0.34
	$\chi_{3,1}$	-2.77	0.11
	$\chi_{3,2}$	-1.95	0.14
	$\chi_{3,3}$	-1.40	0.15
	$\chi_{3,4}$	-0.78	0.15
	$\chi_{3,5}$	-0.05	0.15
Quantitative reasoning			
	$\beta_{4,1}$	4.63	0.45
	$\chi_{4,1}$	-1.90	0.09
	$\chi_{4,2}$	-0.70	0.12
	$\chi_{4,3}$	0.53	0.13
	$\chi_{4,4}$	1.52	0.14
Skill			
	$p_{\theta_t,1}$	0.87	0.17
	$p_{\theta_t,2}$	0.13	0.17
	$\mu_{\theta_t,1}$	0.06	0.05
	$\mu_{\theta_t,2}$	-0.42	0.63
	$\sigma^2_{\theta_t,1}$	0.04	0.01
	$\sigma^2_{\theta_t,2}$	0.04	0.06
Wage shocks			
	$p_{\varepsilon_t,1}$	0.12	0.07
	$p_{\varepsilon_t,2}$	0.88	0.07
	$\mu_{\varepsilon_t,1}$	0.41	0.13
	$\mu_{\varepsilon_t,2}$	-0.05	0.04
	$\sigma^2_{\varepsilon_t,1}$	0.83	0.28
	$\sigma^2_{\varepsilon_t,2}$	0.27	0.02

lated for workers, while covariances based only on test scores are calculated for both workers and non-workers.

2. "Worker" Sample: This eliminates covariances for test score measures unless an individual is working in both periods.

3. "Wage 50-60" Sample: This only includes covariances with log wage residuals for years when the worker is ages 50-60.

4. "Exp 30-40" Sample: This only includes covariances with log wage residuals for years when the worker has 30-40 years of potential experience.

The "Full" Sample raises the most concern about selection due to early retirement. For example, if some (e.g. lower skill) workers retire early when they experience a low wage shock, ε_t, while other workers (e.g. higher skill) do not, this can distort the covariance between log wage residuals and test score measures. Yet, the covariance between test scores is unaffected by this sort of selection. The "Worker" Sample does not address the selection problems, but it would provide more direct estimates that apply to the selected sample of workers. The "Wage 50-60" and "Exp 30-40" Samples address concerns about selective retirement to the extent that the vast majority of men are still working throughout their 50s or prior to reaching 40 years of potential experience (e.g. age 58 for a high school graduate). Even those retiring "early" typically work during these years.

Using data available for even-numbered years, we use the GMM approach of Section 4.2.3 (with identity weighting matrix) to estimate $\beta_{1,1}$ (for word recall), returns to skill λ_t, and the evolution of skill distributions. Because we allow for age/cohort variation in the distribution of "initial" $\phi_{a,t}$ skill shocks, we assume a cohort-invariant distribution of skill fixed effects (i.e. $Var(\psi|c) = Var(\psi)$ for all cohorts c). Regarding the skill distributions, we estimate the two-year autocorrelation for persistent skills ρ^2, variance of fixed effects $Var(\psi)$, variances of skill shocks $\sigma^2_{\nu_t}$, and variances of "initial" $\phi_{a,t}$ skill shocks when individuals enter the sample.[22]

[22] We estimate λ_t for $t = 1996$ to 2016, normalizing $\lambda_{t^*} = 1$ for $t^* = 2010$. We estimate $\sigma^2_{\nu_t}$ for years $t = 1998$ to 2016, normalizing $\sigma^2_{\nu_{1996}} = 0$. We estimate variances of initial AR(1) skill shocks for men first observed in 1996, $Var(\phi_{a,1996})$ for ages $a = 52$ to 65, as well as for men first observed at ages 52 and 53 in other years, i.e. $Var(\phi_{52,t})$ and $Var(\phi_{53,t})$ for $t = 1998$ to 2016.

Table 4.5 reports the estimates and standard errors for $\beta_{1,1}$, ρ^2, and $Var(\psi)$ for all four samples. The estimates are fairly similar across samples; although, the estimated $\beta_{1,1}$ mapping skills into word recall scores ranges from 9.0 for the "Worker" Sample to 11.8 for the "Exp 30-40" Sample. Of greater interest are the skill fixed effects variance estimates, which range from 0.014 to 0.025. Based on the "Full" Sample estimates, these suggest that variation in these permanent skill differences accounts for 38% of the variation in skills and 5% of the variation in log wages at age 60 in 2002. Our estimates for ρ^2, which reflect the dynamics of skill shocks, are also similar across samples at 0.86 to 0.87. These imply values for ρ of about 0.93, within the typical range of autocorrelations for log earnings innovations in the earnings dynamics literature (Meghir and Pistaferri, 2011).

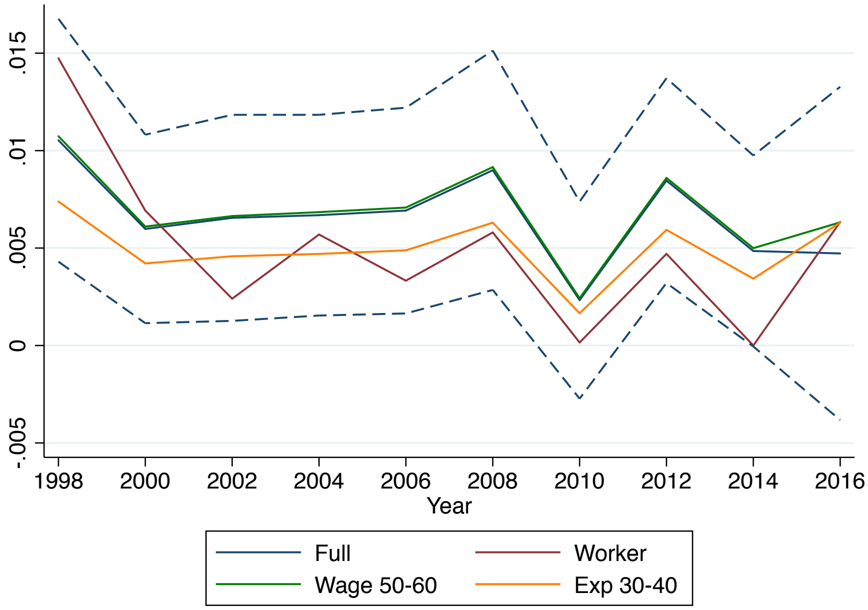
Table 4.5: GMM Estimation Results (Selected Parameters)

Parameter	Full	Worker	Wage 50-60	Exp 30-40
$\beta_{1,1}$	9.936	9.032	9.834	11.844
	(0.743)	(0.724)	(0.802)	(1.169)
ρ^2	0.861	0.867	0.861	0.861
	(0.037)	(0.050)	(0.037)	(0.037)
$Var(\psi)$	0.021	0.025	0.021	0.014
	(0.007)	(0.011)	(0.007)	(0.005)

The estimated time patterns for $\sigma^2_{\nu_t}$ are shown in Figure 4.1, while estimates for $Var(\phi_{a,t})$ are shown in Figure 4.2 (estimates for all ages in 1996) and Figure 4.3 (estimates for ages 52 and 53 for years 1998-2016). These estimates suggest considerable stability in the process for persistent skill shocks over time and across cohorts.

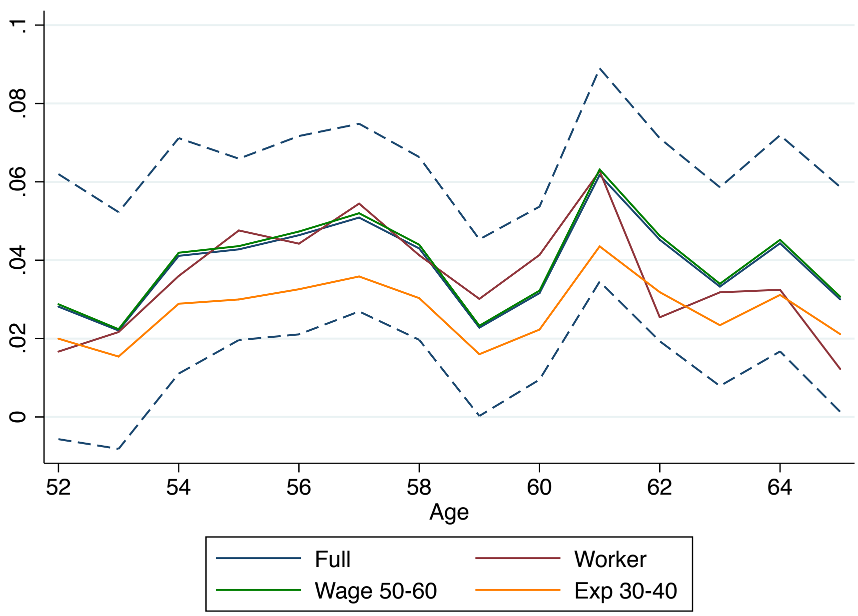
Finally, Figure 4.4 plots estimated returns to skill, λ_t, over time for all samples, along with their 95% confidence intervals. Estimated profiles for all four samples suggest relative stability in skill returns until the Great Recession, after which they appear to rise steadily through the end of our sample period. Point estimates for the "Full" Sample suggest that the returns to skill rose from a low of 0.82 in 2008 to a high of 1.21 in 2016.[23] The qualitative pattern of

[23] Using a Wald test, we reject that the return does not change from 2008 to 2016 at 1% significance level.



Note: Dashed lines represent the 95% confidence intervals for the Full sample.

Figure 4.1: Estimated Skill Shocks $\sigma^2_{\nu_t} \equiv Var(\nu_t)$ by Year



Note: Dashed lines represent the 95% confidence intervals for the Full sample.

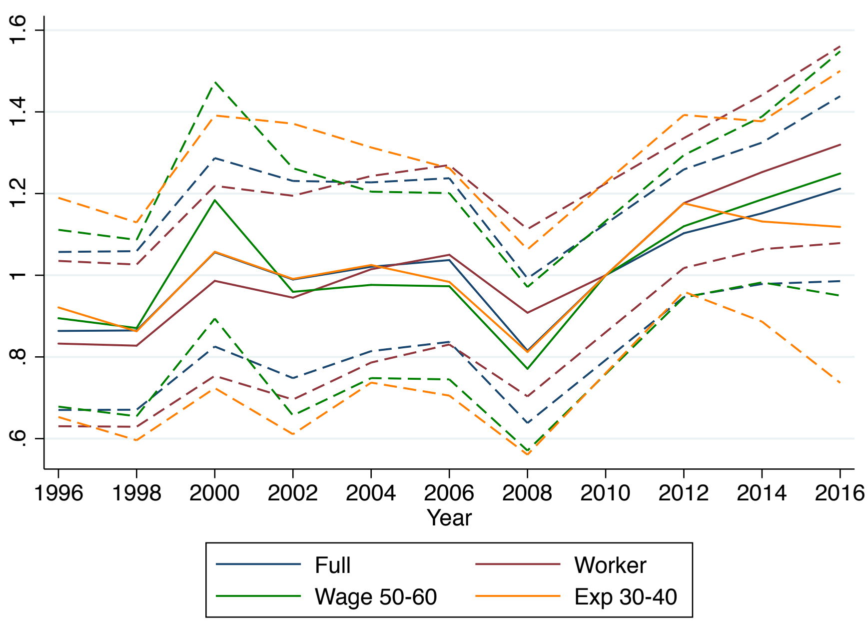
Figure 4.2: Estimated $Var(\phi_{a,t})$ by Age for $t = 1996$

(a) $a = 52$ (b) $a = 53$

Note: Dashed lines represent the 95% confidence intervals for the Full sample.

Figure 4.3: Estimated $Var(\phi_{a,t})$ by Year

relatively stable returns to skill in the late 1990s and early 2000s, followed by a rise after 2008, is consistent with the estimates of Lochner et al. (2020); however, standard errors for $\hat{\lambda}_t$ are large, making it is difficult to say how much returns actually rose after 2008, or to evaluate year-to-year changes in returns, with any confidence.



Note: Dashed lines represent the 95% confidence intervals for the Full sample.

Figure 4.4: Estimated Return to Skill (λ_t) by Year

4.4.3 Average Skill Profiles

We now explore average skill profiles by age and time for various subpopulations of interest. Because we observe cognitive test scores for individuals whether they work or not, we can examine the evolution of skills through and after retirement. Indeed, we report average lifecycle skill profiles over ages 50-70. We begin by showing profiles for different birth cohorts in the HRS, then consider different skill profiles by education and race. Finally, we explore differences in skill profiles for workers who retire at different ages.

Since $T_{i,1,a,t} = \beta_{1,0} + \beta_{1,1}\theta_{i,a,t} + \eta_{i,1,a,t}$ for word recall, we can write

$$\theta_{i,a,t} = \frac{T_{i,1,a,t} - \beta_{1,0}}{\beta_{1,1}} - \frac{\eta_{i,1,a,t}}{\beta_{1,1}},$$

where $\eta_{i,1,a,t}$ is mean zero. Linearity of the test score function implies that actual skills are simply a re-scaled measure of test scores plus idiosyncratic noise. While test score measures only allow us to obtain very noisy estimates for any specific individual's skill level, we can obtain much more precise estimates of average skill levels.

Since we normalize $\alpha_{54,1996} = 0$, we have $\beta_{1,0} = E(T_{1,54,1996})$, which can easily be estimated using the sample mean for word recall scores among individuals age 54 in 1996: $\hat{\beta}_{1,0} = \bar{T}_{1,54,1996}$. Using this along with our estimate of $\hat{\beta}_{1,1}$ from Table 4.5, we estimate average skills as

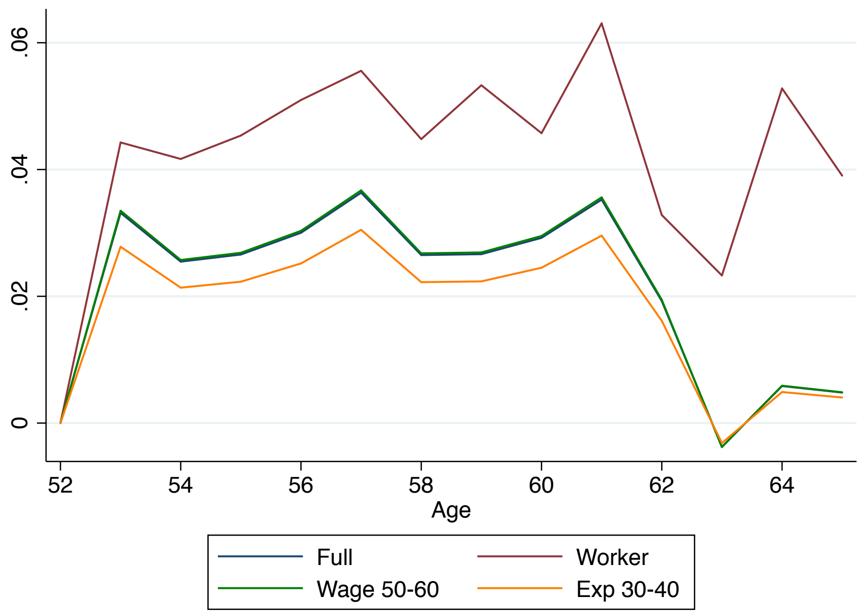
$$\widehat{\alpha}_{a,t} = \frac{\bar{T}_{1,a,t} - \hat{\beta}_{1,0}}{\hat{\beta}_{1,1}}, \tag{4.9}$$

where $\bar{T}_{1,a,t}$ reflects average $T_{1,a,t}$ for all individuals age a in year t. We can similarly obtain estimates for average skills (by age and time) conditional on any personal characteristics as long as those characteristics are independent of cognitive achievement measurement errors. In this case, we would simply use average word recall scores for the subpopulation of interest in equation (4.9). It is also worth noting that we can calculate average skills for ages outside the range we used in estimation, assuming that the measurement function mapping skills to test scores is age-invariant.

We begin by estimating average skills by age and time, $\widehat{\alpha}_{a,t}$. Recall that skills are measured in log wage units (as of 2010), so differences in skill translate roughly into percentage differences. Rather than show all of these estimates, we first regress $\widehat{\alpha}_{a,t}$ on age and year indicators (weighting by the number of observations in each age-year cell) to explore the extent to which these vary with age and time. In Figure 4.5, panel (a) plots the regression coefficients on age indicators for all four estimation samples, while panel (b) plots the regression coefficients on year dummies. The age patterns suggest relative stability, except for a roughly 3 percentage point jump up in average skills from age 52 (our base group in the regression) to age 53 and a similar sized drop between ages 61 and 63. Panel (b) shows relative stability in average skills over time with a sharp drop between 2008 and 2010 with the introduction of the Mid-Baby Boomer cohort.

Because cohorts may differ in their skills, the introduction of new cohorts to the HRS sample can produce jumps up and down in average skills like those seen Figure 4.5. We next look at the age profiles (ages 52-70) for different cohorts, which should be representative of average skills for those cohorts. Because each of these cohorts faced different educational, social, and economic conditions throughout their lives, one might expect to observe differences in their accumulated skills as of age 52 and beyond. Indeed, we see sizeable differences as documented in Figure 4.6. From age 52 to 62, average skill levels are highest for men born during World War II (WWII) and earlier, followed by the Early Baby Boomer cohorts (born 1948-1953), and then subsequent cohorts. At age 55, cohorts born before the War had average skill levels that were about 11 percentage points higher than men born between 1954 and 1959. By age 60, this gap had shrunk to about 3 percentage points, disappearing by age 63. Beyond age 63, cohort differences are small, even reversing with the earliest cohort exhibiting a much more rapid decline in skills with age compared to the War Babies and Early Baby Boomers. While average skill levels began to decline with age for men in their mid-50s for the cohorts born before, during, and immediately after WWII (by 8-13 percentage points), skill profiles remained relatively flat for the Mid-Baby Boomer cohorts (born 1954-59) throughout their late-50s and early-60s. Unfortunately, few men from the most recent cohort (born 1960-65) are observed beyond age 55, so it is difficult to say whether the apparent flattening of lifecycle skill profiles among men ages 55+ will continue.

Next, we explore whether lifecycle skill profiles among older workers differ systematically by education or race. To remove the influence of any cohort differences, we regress $\hat{\theta}_{i,a,t}$ =

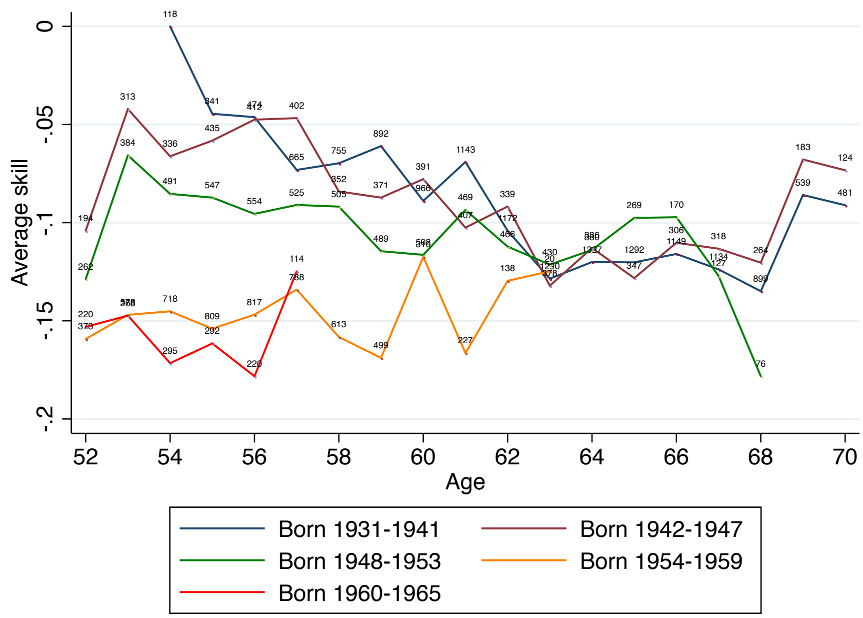


(a) Age effects

(b) Time Effects

Figure 4.5: $\widehat{\alpha}_{a,t}$ Regression Coefficients

$(T_{i,1,a,t} - \hat{\beta}_{1,0})/\hat{\beta}_{1,1}$ on HRS cohort indicators and interactions between annual age indicators and educational attainment indicators (less than high school, high school graduate, some college, college graduate, post-graduate) or race indicators (white vs. non-white).



Note: Numbers above data points reflect the number of observations.

Figure 4.6: Average Skill Profiles by Cohort

Not surprisingly, Figure 4.7 shows sizeable and statistically significant differences in skills across education groups, with college graduates possessing about 15-20% higher skill levels than high school graduates over ages 56-70.[24] High school dropouts have 10-17% lower skill levels than high school graduates. What is, perhaps, most noteworthy about this figure is the apparent parallelism in skills, even through typical retirement and post-retirement ages. Skills are systematically declining beyond age 55 with similar rates of decline for all education groups.[25]

Figure 4.8 shows the estimated average skill profiles by race. Consistent with lower wages among non-whites, we see that average skill levels are about 10-20 percentage points lower for non-white men over ages 52-57. (These gaps are statistically significant at all ages.) As with education, we see similar lifecycle profiles for both whites and non-whites.[26]

[24] F-tests for equality of skills across any education comparison at any age from 52 to 70 yield p-values less than 0.05 for all but four comparisons.

[25] Using F-tests for equality of average changes in skill (from age a to $a + 1$ for all available a shown in the figure) across education groups, we cannot reject parallelism in age profiles (at 5% significance level) for any education comparison. We also cannot reject parallelism for any education groups over subperiods, including ages 55-60, 60-65, and 65-70.

[26] Based on F-tests, we only reject parallelism over ages 52-55. We cannot reject parallelism over ages 55-60, 60-65, and 65-70.

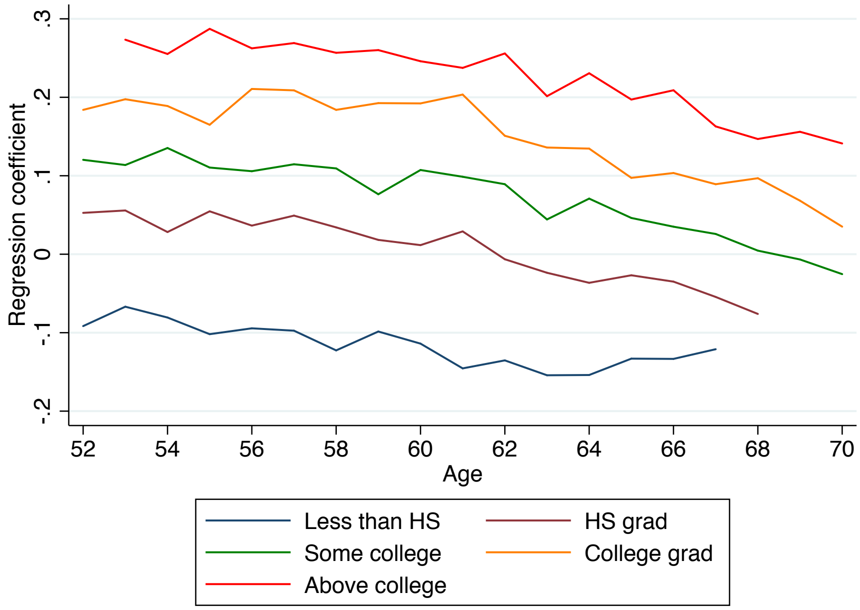


Figure 4.7: Estimated Average Skill Profiles by Education

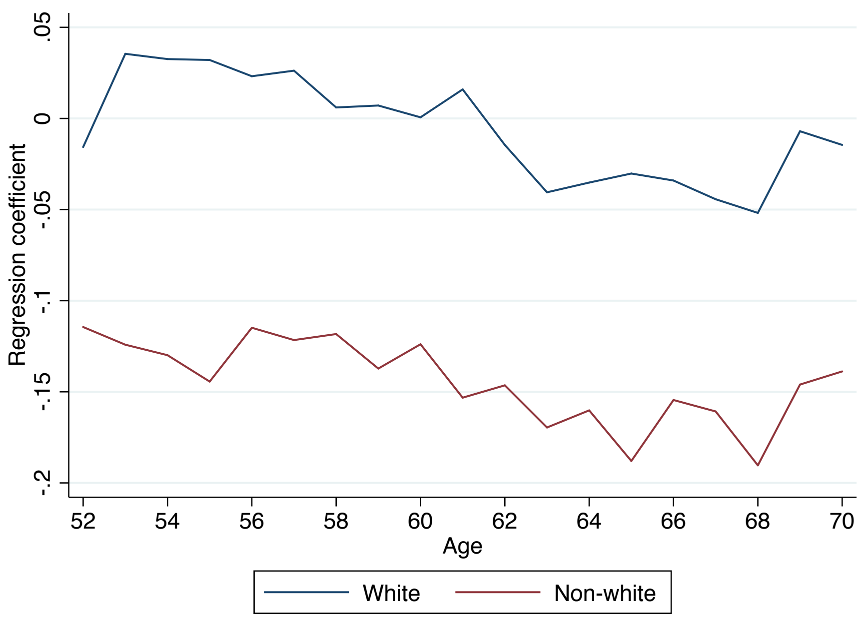


Figure 4.8: Estimated Average Skill Profiles by Race

Table 4.6 reports average estimated skill levels and log wage residuals, which net out age and time effects, by race and education for individuals ages 55-60. Column 1 reports average skills for the full sample, while column 2 reports average skills for the sample of workers (i.e. respondents reporting wages during the same periods). Average skill levels by education and race are larger for the sample of workers, but the differences are modest and vary little across education and race groups. This suggests that selection into retirement has quite modest effects on average skill levels in the workforce. As already evident in Figures 4.7 and 4.8, the average skill gap between college and high school graduates is quite similar to the skill gap between whites and non-whites, about 15%. Column (3) shows that the corresponding gaps in average log wage residuals (also at ages 55-60) are much larger — college graduates have 44% higher wages than high school graduates, while the race gap in wages is about 35%. Thus, the cognitive skills captured by our measures explain an important share of education and race wage gaps, but other factors also play an important role.

Table 4.6: Average Estimated Skill and Log Wage Residuals by Education/Race (Ages 55-60)

	Estimated Skill (Full)	Estimated Skill (Workers)	Log Wage Residual (Workers)
A. By Education			
Less than HS	-0.287	-0.241	-0.414
	(0.006)	(0.008)	(0.015)
HS grad	-0.147	-0.121	-0.186
	(0.005)	(0.006)	(0.010)
Some college	-0.085	-0.053	-0.035
	(0.005)	(0.006)	(0.012)
College grad	0.002	0.022	0.249
	(0.007)	(0.007)	(0.017)
Above college	0.083	0.089	0.483
	(0.006)	(0.007)	(0.017)
B. By Race			
White	-0.051	-0.022	0.105
	(0.003)	(0.004)	(0.008)
Non-white	-0.203	-0.160	-0.244
	(0.005)	(0.006)	(0.011)
Num. of obs.	14,047	9,064	9,064

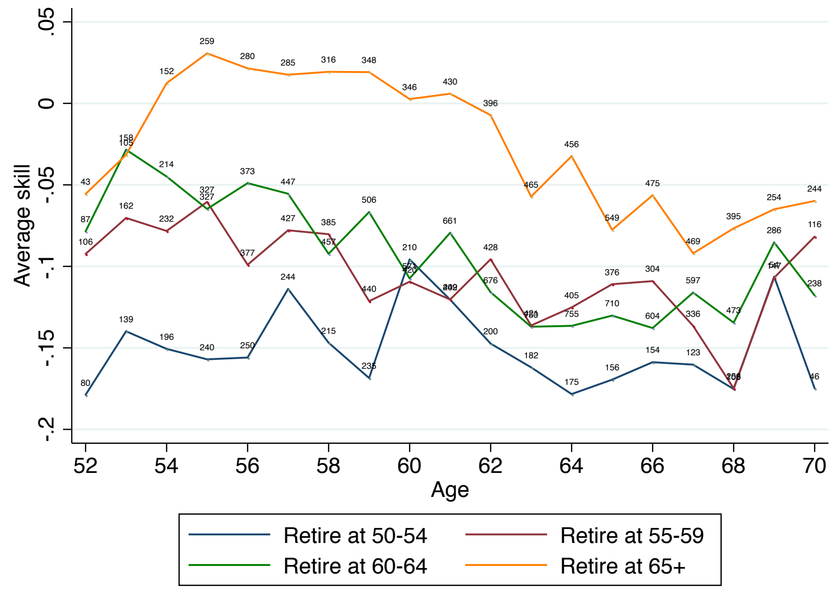
Notes: Standard errors in parentheses.

The results presented so far suggest a systematic decline in skills for men that begins when they are in their mid-50s (or earlier). Is this explained by a gradual increase in rates of retirement with sharp declines in skills for those who retire, or does it reflect more gradual declines for all workers regardless of when they decide to retire? The patterns presented in Figure 4.9 favor the latter explanation. Panel (a) shows lifecycle average skill profiles separately for workers who retire at ages 50-54, 55-59, 60-64, and 65+, while panel (b) removes cohort effects by regressing $\hat{\theta}_{i,a,t}$ on cohort indicators and interactions of age indicators with retirement age indicators (as with education and race above). In neither case do we see evidence of steep drops over the ages when individuals retire or in the time immediately following retirement. Still, the skill levels and lifecycle patterns notably differ for those who retire early and those who retire late compared to those retiring between ages 55 and 64. Those who retire before age 55 have much lower skill levels in their 50s compared to those who retire later; however, their skills continue to grow until age 61, while the skills of men retiring at ages 55-64 decline over most of these ages. Differences between very early retirees and those retiring at ages 55-64 are largely eliminated by age 60.

Those who retire at ages 65 or older possess the highest skill levels; however, their lifecycle profile over ages 52-70 looks more like that of very early retirees than those retiring in their late 50s and early 60s. These late retirees experience modest skill growth until their mid-50s, stable skill levels to about age 60, and strong declines thereafter. By age 65, about half of the difference in skills between them and those retiring over ages 55-64 is eliminated.

These patterns imply a complex relationship between retirement and skills. There is clear evidence that those with high skills in their mid-50s choose to retire late while those with low skills (a difference of more than 10 percentage points) choose to retire quite early. But, there is little evidence to suggest that retirement itself is strongly associated with a decline in skills.[27] Among those retiring very young, skills continue to increase for years after they retire, several years after they have already started declining for those retiring in their late 50s or early 60s. Over the same ages, skills are also increasing or stable for men who retire after age 65.

[27] Both Rohwedder and Willis (2010) and Bonsang et al. (2012) estimate significant negative causal effects of retirement on cognition using the HRS; however, Coe et al. (2012) does not.



Note: Numbers above data points reflect the number of observations.

(a) Average Skill

(b) Estimated Average Skill

Figure 4.9: Average Skill Profiles by Retirement Age

4.5 Conclusions

With multiple skill measures and wages each period, we have shown that if at least one measure is continuous and repeated, it is possible to nonparametrically identify the evolution of skill prices and cross-sectional skill distributions over time without any assumptions on the distributions or dynamics of skills. With panel data, the same measurements and wages can identify skill dynamics as well. Our constructive identification analysis motivates a very general multi-step estimation approach. We also show that if any of the continuous measurements is found to be linear in skills (in the first estimation step), a simple GMM approach can be used to estimate skill returns, the means and variances of skill distributions over time, and a flexible dynamic process for skills with a fixed effect and AR(1) stochastic process.

Using data from the 1996-2016 HRS, we estimate the evolution of skill returns, skill distributions, and skill dynamics for American men ages 52+ over that period. We first show that one of the repeated continuous test measures we observe is linear in skills and then use our simpler GMM estimation approach. Our estimates suggest that the returns to (cognitive) skills were fairly stable from the mid-1990s through the early 2000s, but then began to rise significantly after the Great Recession. This pattern is broadly consistent with that of Lochner et al. (2020).

We document considerable differences in average skill levels and lifecycle profiles across cohorts. More recent cohorts of men had significantly lower average skill levels in their mid-50s than did earlier cohorts when they were the same ages. However, earlier cohorts experienced much faster declines in skill with age, such that the earlier skill differences had largely disappeared by the time cohorts had reached their 60s. For the latest cohort we observe (men born in 1954–1959), we see no discernable decline in average skills prior to age 63 when they are last observed. Distinguishing individuals by education and race, we find that average skills are monotonically increasing in education and are higher for whites than non-whites. These education and race gaps are quite similar across ages and explain about one-third of the education differences and nearly half of the race differences in log wages.

We also consider the interaction of skills and retirement, showing that those who retire at older ages have substantially higher skills in their mid-50s. While skills generally decline with age for men, at least after reaching age 60, we see no sharp declines around the time men retire. To the extent that retirement does lead to cognitive decline, our results suggest that the effects are relatively modest or largely offset by other lifecycle forces.

Finally, we show that individual fixed effects account for more than a third of all skill variation at age 60. Year-to-year fluctuations are also persistent (though not a random walk) with an autocorrelation of 0.93.

In future work with the HRS data, we plan to test the validity of previous assumptions in the literature regarding the evolution of skill differences or skill growth over the lifeycle. If some of these assumptions are shown to be valid, it would provide additional credibility to previous studies and further justification for those assumptions when using other data sources without direct skill measures. We can also make more use of additional test measures (even those that are non-linear in skills), using our GMM approach to obtain more precise estimates of skill returns and skill distributions over time. In addition to measuring differences in average skill levels over time, it is straightforward to estimate changes in the distributions of skills. For example, we can estimate the distributions of skills for workers choosing to retire at different ages to better understand selection into retirement. Finally, it is possible to allow log wage equations to differ by education and/or race, accounting for the fact that other factors or skills (besides those measured by the cognitive tests in HRS) might play an important role in wage determination.

Bibliography

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L Violante (2019) "Education Policy and Intergenerational Transfers in Equilibrium," *Journal of Political Economy*, Vol. 127, pp. 2569–2624.

Adam, Stéphane, Eric Bonsang, Sophie Germain, and Sergio Perelman (2007) "Retirement and Cognitive Reserve: A Stochastic Frontier Approach Applied to Survey Data," CREPP Working Paper 2007/04.

Altonji, Joseph G, Erica Blom, and Costas Meghir (2012) "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics*, Vol. 4, pp. 185–223.

Altonji, Joseph G, Fumio Hayashi, and Laurence J Kotlikoff (1997) "Parental Altruism and Inter Vivos Transfers: Theory and Evidence," *Journal of Political Economy*, Vol. 105, pp. 1121–1166.

Autor, David H, Lawrence F Katz, and Melissa S Kearney (2008) "Trends in US Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, Vol. 90, pp. 300–323.

Bachas, Natalie (2019) "The Impact of Risk-Based Pricing in the Student Loan Market: Evidence from Borrower Repayment Decisions," Working Paper.

Blau, Francine D and Lawrence M Kahn (1997) "Swimming Upstream: TRends in the Gender Wage Differential in the 1980s," *Journal of labor Economics*, Vol. 15, pp. 1–42.

——— (2017) "The Gender Wage Gap: EXtent, Trends, and Explanations," *Journal of Economic Literature*, Vol. 55, pp. 789–865.

Bonsang, Eric, Stéphane Adam, and Sergio Perelman (2012) "Does Retirement Affect Cognitive Functioning?" *Journal of Health Economics*, Vol. 31, pp. 490–501.

Bowlus, Audra J and Chris Robinson (2012) "Human Capital Prices, Productivity, and Growth," *American Economic Review*, Vol. 102, pp. 3483–3515.

Bricker, Jesse, Meta Brown, Simona Hannon, and Karen Pence (2015) "How Much Student Debt is Out There?", Board of Governors of the Federal Reserve System.

Brown, Meta, John Karl Scholz, and Ananth Seshadri (2012) "A New Test of Borrowing Constraints for Education," *The Review of Economic Studies*, pp. 511–538.

Cagney, Kathleen A and Diane S Lauderdale (2002) "Education, Wealth, and Cognitive Function in Later Life," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, Vol. 57, pp. P163–P172.

Card, David (1999) "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, Vol. 3: Elsevier, pp. 1801–1863.

Card, David and Thomas Lemieux (1996) "Wage Dispersion, Returns to Skill, and Black-White Wage Differentials," *Journal of Econometrics*, Vol. 74, pp. 319–361.

Castex, Gonzalo and Evgenia Kogan Dechter (2014) "The Changing Roles of Education and Ability in Wage Determination," *Journal of Labor Economics*, Vol. 32, pp. 685–710.

Castora-Binkley, Melissa, Carol L Peronto, Jerri D Edwards, and Brent J Small (2015) "A Longitudinal Analysis of the Influence of Race on Cognitive Performance," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, Vol. 70, pp. 512–518.

Celidoni, Martina, Chiara Dal Bianco, and Guglielmo Weber (2017) "Retirement and Cognitive Decline. a Longitudinal Analysis Using SHARE Data," *Journal of Health Economics*, Vol. 56, pp. 113–125.

Chatterjee, Satyajit and Felicia Ionescu (2012) "Insuring Student Loans Against the Financial Risk of Failing to Complete College," *Quantitative Economics*, Vol. 3, pp. 393–420.

Chay, Kenneth Y and David S Lee (2000) "Changes in Relative Wages in the 1980s Returns to Observed and Unobserved Skills and Black–White Wage Differentials," *Journal of Econometrics*, Vol. 99, pp. 1–38.

Coe, Norma B, Hans-Martin von Gaudecker, Maarten Lindeboom, and Jürgen Maurer (2012) "The Effect of Retirement on Cognitive Functioning," *Health Economics*, Vol. 21, pp. 913–927.

Coe, Norma B and Gema Zamarro (2011) "Retirement Effects on Health in Europe," *Journal of Health Economics*, Vol. 30, pp. 77–86.

Cunningham, Alisa and Gregory Kienzl (2011) "Delinquency: The Untold Story of Student Loan Borrowing," Institute for Higher Education Policy (IHEP), Washingotn, DC.

Deming, David J (2017) "The Growing Importance of Social Skills in the Labor Market," *The Quarterly Journal of Economics*, Vol. 132, pp. 1593–1640.

Deming, David J. and Susan Dynarski (2010) *Into College, Out of Poverty? Policies to Increase the Postsecondary Attainment of the Poor*: University of Chicago Press.

Dynarski, Susan and Daniel Kreisman (2013) "Loans for Educational Opportunity: Making Borrowing Work for Todays Students," The Hamilton Project, Discussion Paper.

Edin, Per-Anders, Peter Fredriksson, Martin Nybom, and Bjorn Ockert (2017) "The Rising Return to Non-Cognitive Skill," Working Paper.

Employment and Social Development Canada (2016) "Canada Student Loans Program: Annual Report 2014 - 2015," Retrieved from: https://www.canada.ca/en/employment-social-development/programs/canada-student-loans-grants/reports/cslp-2015.html dated August 17, 2020.

——— (2017) "Canada Student Loans Program: Annual Report 2015 - 2016," Retrieved from: https://www.canada.ca/en/employment-social-development/programs/

canada-student-loans-grants/reports/cslp-annual-2015-2016.html dated August 17, 2020.

——— (2018) "Canada Student Loans Program: Annual Report 2016 - 2017," Retrieved from: https://www.canada.ca/en/employment-social-development/programs/canada-student-loans-grants/reports/cslp-annual-2016-2017.html dated August 17, 2020.

Findeisen, Sebastian and Dominik Sachs (2016) "Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans," *Journal of Public Economics*, Vol. 138, pp. 1–21.

Fortin, Nicole M (2019) "Increasing Earnings Inequality and the Gender Pay Gap in Canada: Prospects for Convergence," *Canadian Journal of Economics*, Vol. 52, pp. 407–440.

Frenette, Marc and Kristyn Frank (2016) "Earnings of Postsecondary Graduates by Detailed Field of Study," *Economic Insights*, Statistics Canada, Catalogue no. 11-626-X.

Gary-Bobo, Robert J and Alain Trannoy (2014) "Optimal Student Loans and Graduate Tax under Moral Hazard and Adverse Selection," Working Paper.

Gourieroux, Christian, Alain Monfort, and Eric Renault (1993) "Indirect Inference," *Journal of Applied Econometrics*, Vol. 8, pp. S85–S118.

Government Accountability Office (2015) "Federal Student Loans: Education Could Do More to Help Ensure Borrowers are Aware of Repayment and Forgiveness Options," Washington DC, GAO-15-663.

Grogger, Jeff and Eric Eide (1995) "Changes in College Skills and the Rise in the College Wage Premium," *Journal of Human Resources*, pp. 280–310.

Gross, Jacob PK, Osman Cekic, Don Hossler, and Nick Hillman (2009) "What Matters in Student Loan Default: A Review of the Research Literature.," *Journal of Student Financial Aid*, Vol. 39, pp. 19–29.

Hanushek, Eric A, Guido Schwerdt, Simon Wiederhold, and Ludger Woessmann (2015) "Returns to Skills Around the World: EVidence From PIAAC," *European Economic Review*, Vol. 73, pp. 103–130.

Hanushek, Eric A and Lei Zhang (2009) "Quality-Consistent Estimates of International Schooling and Skill Gradients," *Journal of Human Capital*, Vol. 3, pp. 107–143.

Heckman, James J, Lance J Lochner, and Petra E Todd (2006) "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," *Handbook of the Economics of Education*, Vol. 1, pp. 307–458.

Heckman, James J, Lance Lochner, and Christopher Taber (1998) "Explaining Rising Wage Inequality: Explorations With a Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents," *Review of Economic Dynamics*, Vol. 1, pp. 1–58.

Heckman, James and Edward Vytlacil (2001) "Identifying the Role of Cognitive Ability in Explaining the Level of and Change in the Return to Schooling," *Review of Economics and Statistics*, Vol. 83, pp. 1–12.

Herrnstein, Richard J and Charles A Murray (1994) *The Bell Curve: Reshaping of American Life by Differences in Intelligence*: Free Press.

Hu, Yingyao and Susanne M Schennach (2008) "Instrumental Variable Treatment of Nonclassical Measurement Error Models," *Econometrica*, Vol. 76, pp. 195–216.

Human Resources and Skills Development Canada (2012) "Canada Student Loans Program: Annual Report 2010 - 2011," Retrieved from: https://www.canada.ca/content/dam/canada/employment-social-development/migration/documents/assets/portfolio/docs/en/reports/cslp_cesp/cslp_2011/cslp_stats_2011.pdf dated August 17, 2020.

Ionescu, Felicia (2009) "The Federal Student Loan Program: Quantitative Implications for College Enrollment and Default Rates," *Review of Economic Dynamics*, Vol. 12, pp. 205–231.

——— (2011) "Risky Human Capital and Alternative Bankruptcy Regimes for Student Loans," *Journal of Human Capital*, Vol. 5, pp. 153–206.

Ji, Yan (2018) "Job Search under Debt: Aggregate Implications of Student Loans."

Johnson, Matthew T (2013) "Borrowing Constraints, College Enrollment, and Delayed Entry," *Journal of Labor Economics*, Vol. 31, pp. 669–725.

Juhn, Chinhui, Kevin Murphy, and Brooks Pierce (1989) "Wage Inequality and the Returns to Skill," *Journal of Political Economy*.

Kane, Thomas J (1995) "Rising Public College Tuition and College Entry: How Well Do Public Subsidies Promote Access to College?", Working Paper.

Kaplan, Greg (2012) "Moving Back Home: Insurance Against Labor Market Risk," *Journal of Political Economy*, Vol. 120, pp. 446–512.

Karlamangla, Arun S, Dana Miller-Martinez, Carol S Aneshensel, Teresa E Seeman, Richard G Wight, and Joshua Chodosh (2009) "Trajectories of Cognitive Function in Late Life in the United States: Demographic and Socioeconomic Predictors," *American Journal of Epidemiology*, Vol. 170, pp. 331–342.

Katz, Lawrence F and Kevin M Murphy (1992) "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, Vol. 107, pp. 35–78.

Kosir, Marilyn, Jerry Danette, Peter Li, and Thomas Sanford (2015) "Assessing the Feasibility of Creating a SELF Loan Refinancing Program.," Minnesota Office of Higher Education.

Kotlarski, Ignacy (1967) "On characterizing the gamma and the normal distribution," *Pacific Journal of Mathematics*, Vol. 20, pp. 69–76.

Lemieux, Thomas (2006) "Postsecondary Education and Increasing Wage Inequality," *American Economic Review*, Vol. 96, pp. 195–199.

Leuven, Edwin, Hessel Oosterbeek, and Hans Van Ophem (2004) "Explaining International Differences in Male Skill Wage Differentials by Differences in Demand and Supply of Skill," *Economic Journal*, Vol. 114, pp. 466–486.

Lochner, Lance J and Alexander Monge-Naranjo (2015) "Default and Repayment among Baccalaureate Degree Earners," in B. J. Hershbein and K. Hollenbeck eds. *Student Loans and the Dynamics of Debt*, Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, Chap. 8, pp. 235–284.

Lochner, Lance and Alexander Monge-Naranjo (2016) "Student Loans and Repayment: Theory, Evidence, and Policy," Vol. 5: Elsevier, pp. 397–478.

Lochner, Lance, Youngmin Park, and Youngki Shin (2020) "Wage Dynamics and Returns to Unobserved Skill," Working Paper.

Lochner, Lance and Youngki Shin (2014) "Understanding Earnings Dynamics: Identifying and Estimating the Changing Roles of Unobserved Ability, Permanent and Transitory Shocks," Working Paper.

Lochner, Lance, Todd Stinebrickner, and Utku Suleymanoglu (2018) "Parental Support, Savings and Student Loan Repayment," Working Paper.

Loury, Linda Datcher and David Garman (1995) "College Selectivity and Earnings," *Journal of Labor Economics*, Vol. 13, pp. 289–308.

Lucas, Deborah and Damien Moore (2010a) "Guaranteed versus Direct Lending: The Case of Student Loans," in Deborah Lucas ed. *Measuring and Managing Federal Financial Risk*: University of Chicago Press, Chap. 7, pp. 163–205.

——— (2010b) "Guaranteed versus Direct Lending: The Case of Student Loans," in Deborah Lucas ed. *Measuring and Managing Federal Financial Risk*: University of Chicago Press, Chap. 7, pp. 163–205.

Lucas, Deborah and Marvin Phaup (2010) "The Cost of Risk to the Government and Its Implications for Federal Budgeting," in Deborah Lucas ed. *Measuring and Managing Federal Financial Risk*: University of Chicago Press, Chap. 3, pp. 29–54.

Mazzonna, Fabrizio and Franco Peracchi (2012) "Ageing, Cognitive Abilities and Retirement," *European Economic Review*, Vol. 56, pp. 691–710.

——— (2017) "Unhealthy Retirement?" *Journal of Human Resources*, Vol. 52, pp. 128–151.

McGarry, Kathleen (2016) "Dynamic Aspects of Family Transfers," *Journal of Public Economics*, Vol. 137, pp. 1–13.

Meghir, Costas and Luigi Pistaferri (2011) "Earnings, Consumption and Life Cycle Choices," in David Card and Orley Ashenfelter eds. *Handbook of labor economics*, Vol. 4B: Elsevier, Chap. 9, pp. 773–854.

Moffitt, Robert A and Peter Gottschalk (2012) "Trends in the Transitory Variance of Male Earnings Methods and Evidence," *Journal of Human Resources*, Vol. 47, pp. 204–236.

Mueller, Holger M and Constantine Yannelis (2019) "Reducing Barriers to Enrollment in Federal Student Loan Repayment Plans: Evidence from the Navient Field Experiment," Working Paper.

Murnane, Richard J, John B Willett, Frank Levy et al. (1995) "The Growing Importance of Cognitive Skills in Wage Determination," *The Review of Economics and Statistics*, Vol. 77, pp. 251–266.

Neal, Derek A and William R Johnson (1996) "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy*, Vol. 104, pp. 869–895.

Office of the Chief Actuary (2012) "Actuarial Report on the Canada Student Loans Program as at 31 July 2011," Retrieved from: https://www.osfi-bsif.gc.ca/Eng/Docs/CSLP_2011.pdf dated August 17, 2020.

——— (2019) "Actuarial Report on the Canada Student Loans Program as at 31 July 2018," Retrieved from: https://www.osfi-bsif.gc.ca/Eng/Docs/CSLP_2018.pdf dated August 17, 2020.

Park, Youngmin (2016) "Heterogeneity in Parental Altruism and Optimal Higher Education Subsidies," Working Paper.

Rohwedder, Susann and Robert J Willis (2010) "Mental Retirement," *Journal of Economic Perspectives*, Vol. 24, pp. 119–38.

Rumberger, Russell W and Scott L Thomas (1993) "The Economic Returns to College Major, Quality and Performance: A Multilevel Analysis of Recent Graduates," *Economics of Education Review*, Vol. 12, pp. 1–19.

Schwartz, Saul and Ross Finnie (2002) "Student Loans in Canada: An Analysis of Borrowing and Repayment," *Economics of Education Review*, Vol. 21, pp. 497–512.

Situ, Jerry (2006) "Canada Student Loans Repayment Assistance: Who Does and Does Not Use Interest Relief?", Culture, Tourism and the Centre for Education Statistics: Research Papers. No. 47.

Taber, Christopher R (2001) "The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?" *The Review of Economic Studies*, Vol. 68, pp. 665–691.

Yannelis, Constantine (2015) "Asymmetric Information in Student Loans," New York University.

Zsembik, Barbara A and M Kristen Peek (2001) "Race Differences in Cognitive Functioning Among Older Adults," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, Vol. 56, pp. S266–S274.

Appendix A

Chapter 2 Appendix

A.1 Details on Field of Study

The following tables list the detailed areas of study for the CSLP and PSIS field of study categories.

Table A1: CSLP Field of Study Classification

Field of Study	Area of Study
Administration/Business	Business Administration; Accounting; Commerce; Communications; Court Reporter; Hotel Administration; Hospitality; Marketing; Media Resources; Municipal Government; Public Relations; Realty Appraisal; Stenography; Programming; Systems Analysis
Agriculture	Agriculture; Veterinary Medicine; Environmental Sciences; Forestry; Horticulture; Dairy Farming; Fish and Wildlife
Arts/Science	Archaeology; Astronomy; Biology; Chemistry; Ecology; Economics; Fine Arts; Geography; Geology; History; Linguistics; Material Science; Mathematics; Philosophy; Physics; Psychology; Sociology; Statistics; Physiology; Library Science; Theatre; Visual Arts; Crafts; Interpreting Journalism
Community Service/Education	Community Services; Counselling; Social Work; Rehabilitation; Visiting Bookmaker; Outdoor Recreation
Dentistry	Dentistry; Dental Technology; Dental Nursing
Engineering/Technology	Engineering; Electronics; Instrumentation; Kinetics; Pilot Training; Architecture; Urban/Regional; Planning; Navigation; Marine Engineering; Waste Technology
Health Sciences	Dietetics; Home Economics; Chiropractice; Chiropody; Podiatry; Registered Nursing; Registered Assistant Nurse; Optometry; Pharmacy; Laboratory Techonology; X-Ray Technology; Audiology; Speech Therapy; Medical Secretary; Food Products; BIO-Medical Sciences; Environmental Food; Medical Radiography
Law	Law; Bar Admission; Para-Legal Studies; Criminology

Table A2: PSIS Field of Study Classification

Field of Study	Area of Study
Education	Education
Visual and Performing Arts, and Communications Technologies	Communications Technologies/Technicians and Support Services; Visual and Performing Arts
Humanities	Aboriginal and Foreign Languages, Literatures and Linguistics; English Language and Literature/Letters; Liberal Arts and Sciences, General Studies and Humanities; Medieval and Renaissance Studies; Holocaust and Related Studies; Classical and Ancient Studies; Maritime Studies; Philosophy and Religious Studies; Theology and Religious Vocations; History; French Language and Literature/Letters
Social and Behavioural Sciences	Area, Ethnic, Cultural, Gender, and Group Studies; Communication, Journalism and Related Programs; Family and Consumer Sciences/Human Sciences; Peace Studies and Conflict Resolution; Biopsychology; Gerontology; Museology/Museum Studies; Science, Technology and Society; Behavioural Sciences; International/Global Studies; Intercultural/Multicultural and Diversity Studies; Cognitive Science; Cultural Studies/Critical Theory and Analysis; Dispute Resolution; Human Computer Interaction; Sustainability Studies; Psychology; Social Sciences
Business, Management and Public Administration	Accounting and Computer Science; Public Administration and Social Service Professions; Business, Management, Marketing and Related Support Services
Physical and Life Sciences, and Technologies	Biological and Biomedical Sciences; Biological and Physical Sciences; Natural Sciences; Nutrition Sciences; Human Biology; Marine Sciences; Physical Sciences; Science Technologies/Technicians
Mathematics, Computer and Information Sciences	Computer and Information Sciences and Support Services; Library Science; Mathematics and Statistics; Systems Science and Theory; Mathematics and Computer Science; Computational Science
Architecture, Engineering and Related Technologies	Architecture and Related Services; Engineering; Engineering Technologies and Engineering-Related Fields; Historic Preservation and Conservation; Construction Trades; Mechanic and Repair Technologies/Technicians; Precision Production
Agriculture, Natural Resources and Conservation	Agriculture, Agriculture Operations and Related Sciences; Natural Resources and Conservation
Health and Related Fields	Parks, Recreation, Leisure and Fitness Studies; Health Professions and Related Programs; Dental, Medical and Veterinary Residency Programs
Law	Legal Professions and Studies
Other	Pre-Technology Education/Pre-Industrial Arts Programs; Multidisciplinary/Interdisciplinary Studies, Other

Appendix B

Chapter 3 Appendix

B.1 State Group

Table B1: State Grouping by Tuition

Group	States	2-year Annual tuition	4-year Annual tuition
1	Alabama, Alaska, Arkansas, Hawaii, Idaho, Kansas, Kentucky, Mississippi, Montana, Nevada, New Mexico, North Dakota, Oklahoma, Texas, Utah, West Virginia, and Wyoming.	1,396	4,900
2	Arizona, California, Colorado, Delaware, Florida, Georgia, Louisiana, Nebraska, North Carolina, South Dakota, and Washington.	963	7,508
3	Illinois, Indiana, Iowa, Maryland, Michigan, Missouri, Oregon, South Carolina, Tennessee, Virginia, and Wisconsin.	2,006	8,718
4	Connecticut, District of Columbia, Maine, Massachusetts, Minnesota, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Rhode Island, and Vermont.	2,741	13,359

B.2 Moments

The moments used are listed below:

- Fraction of high school graduates, 2-year college graduates, and 4-year college graduates.
- Fraction of high school graduates, 2-year college graduates, and 4-year college graduates by AFQT quartile, parental income quartile, and state of residence group.
- Tuition elasticity of education: change in the fraction of 2-year college or 4-year college graduates when 2-year college or 4-year college increases tuition by 1,000 dollars.
- Fraction of youth working full-time, part-time, and not working while enrolled in 2-year and 4-year colleges.
- Fraction of youth working full-time, part-time, and not working after finishing school by education level (high school, 2-year college, and 4-year college).
- The probability of working conditional on not working in the previous year and the probability of not working conditional on working in the previous year by education level.
- Mean and variance of log wages of college students working full-time.
- Mean of log wages of college students working part-time.
- Mean of log wages of youth working full time by education level and year after school.
- Mean of variance of log wages of youth working full time across all years and education levels.
- Mean of covariance of log wages of youth working full time at two consecutive periods across all years and education levels.
- Coefficients of regressing log wages on AFQT, education, cumulative work years, and interaction of part-time work status and education.

- Average parental transfers across all education levels and all years.
- Average youth assets at age 20, 25, and 30.
- Average youth assets if negative at age 20, 25, and 30.
- Average annual amount of student loans by college type.
- 2-year cohort default rate among college graduates.
- Coefficients of regressing 2-year college grants on parental income.
- Coefficients of regression 4-year college grants on AFQT and parental income.

B.3 Parameter Estimates

Table B1: Estimated Parameters

Description	Symbol	Value	Standard Error
Altruism factor	η	0.0517	0.0008
Utility function:			
Work preferences:			
Preference for working part-time	α_{v1}	-0.1015	0.0023
Preference for working full-time	α_{v2}	-0.0084	0.0017
Work part-time while attending 2-year college	α_{v3}	0.0426	0.0085
Work full-time while attending 2-year college	α_{v4}	-0.0797	0.0070
Work part-time while attending 4-year college	α_{v5}	0.0348	0.0047
Work full-time while attending 4-year college	α_{v6}	-0.1413	0.0043
Variance of shocks for part-time working	σ^2_{v1}	0.0176	0.0002
Variance of shocks for full-time working	σ^2_{v2}	2.82E-05	7.32E-07
Preference for college attendance:			
2-year college:			
Intercept	α_{z1}	-0.1566	0.0244
AFQT	α_{z2}	0.0010	0.0004
Log parent income	α_{z3}	0.0149	0.0017
Variance of shocks	σ^2_{z1}	0.0151	0.0074
4-year college:			
Intercept	α_{z4}	-1.2212	0.0113
AFQT	α_{z5}	0.0107	0.0004
Log parent income	α_{z6}	0.0662	0.0020
Variance of shocks	σ^2_{z2}	0.1690	0.0204
2-year college grants:			
Constant	α^c_{g1}	1693.49	17.14
Family income	α^c_{g2}	-0.0017	0.0043
4-year college grants:			
Constant	α^u_{g1}	4779.01	23.20
AFQT	α^u_{g2}	32.2306	0.1747
Family income	α^u_{g3}	-0.0203	0.0042
College wages:			
Mean	$\bar{w}^c$	9.2207	0.0228
Variance	σ^2_c	0.7922	0.0029
Human capital function:			
Constant	$\alpha_{\psi 1}$	9.7308	0.0163

AFQT	$\alpha_{\psi 2}$	0.0029	0.0003
2-year college	$\alpha_{\psi 3}$	0.1798	0.0149
4-year college	$\alpha_{\psi 4}$	0.3283	0.0060
Experience	$\alpha_{\psi 5}$	0.0701	0.0013
Job offer probability:			
Constant	α_{j1}	-0.7704	0.0021
Human capital	α_{j2}	0.2525	0.0017
Not working in previous period	α_{j3}	-1.0833	0.0023
Wage:			
Part-time wage ratio (college student)	α_p^{col}	0.5995	0.0047
Part-time wage ratio (hs grad)	α_p^{nc}	0.4993	0.0015
Part-time wage ratio (2 year grad)	α_p^{c}	0.4964	0.0029
Part-time wage ratio (4 year grad)	α_p^{u}	0.4356	0.0040
Permanent shocks	σ_ξ^2	0.0345	0.0002
Transitory shocks	σ_w^2	0.1525	0.0005
Borrowing limit:			
Constant	α_{x1}	2.5612	0.0033
Human capital	α_{x2}	0.2440	0.0011
Age	α_{x3}	0.1315	0.0007
Student loans:			
Utility cost of borrowing in 2-year college	α_{b1}	-0.0617	0.0124
Utility cost of borrowing in 4-year college	α_{b2}	-0.0647	0.0049
Utility cost of default	α_d	-0.0020	0.0011

B.4 Computational Details

Model solution

Given a set of parameters, the model is solved backward numerically from the terminal value function. Since the state space includes multidimensional continuous state variables, I solve the value function at a subset of the points in the state space for each period. The integrals taken over the distribution of the shocks are approximated by Gauss-Hermite quadrature. I then use multilinear interpolation to approximate the continuation values at points off the discrete grid.

Estimation

To estimate the parameters, I use simulated methods of moments. For the minimization of the criterion function, I use both the Simulated Annealing and the Nelder-Mead algorithms with different starting values to ensure that a global minimum is achieved.

To get the standard errors of the estimates, I calculate the variance covariance matrix for parameter estimates as follows:

$$V = (1 + \frac{1}{S})\left[\frac{\partial m(\theta)'}{\partial \theta} W^* \frac{\partial m(\theta)}{\partial \theta}\right]^{-1},$$

where $\frac{\partial m(\theta)}{\partial \theta}$ is the derivative of the vector of moments with respect to the parameter vector, S is the number of simulations which is set to 30, and W^* is the weighting matrix.[1] The weighting matrix is the inverse of the diagonal matrix with the variance of the corresponding moments as the elements. Derivatives are calculated numerically.

[1] Gourieroux et al. (1993)

Appendix C

Chapter 4 Appendix

C.1 Identification Details

In this appendix, we provide the regularity conditions for the identification result in Section 4.2.2. The result is based on Theorem 1 in Hu and Schennach (2008). For completeness of the arguments, we rewrite all regularity conditions using notation in the current setup. Suppose that we have two test measures, $J = 2$, which is the minimum requirement. The model at (a^*, t^*) can be rewritten as follows:

$$\begin{aligned} w_{i,a^*,t^*} &= \theta_{i,a^*,t^*} + \varepsilon_{i,a^*,t^*} \\ T_{i,1,a^*,t^*} &= \tau_1(\theta_{i,a^*,t^*}) + \eta_{i,1,a^*,t^*} \\ T_{i,2,a^*,t^*} &= G_2\left(\tau_2(\theta_{i,a^*,t^*}) + \eta_{i,2,a^*,t^*}\right). \end{aligned}$$

We collect the necessary regularity conditions below:

Assumption 1 *The observations $(w_{i,a^*,t^*}, T_{i,1,a^*,t^*}, T_{i,2,a^*,t^*})$ generated from the model above satisfy the following conditions:*

i. The joint density of $(\theta_{i,a^*,t^*}, w_{i,a^*,t^*}, T_{i,1,a^*,t^*}, T_{i,2,a^*,t^*})$ is bounded, and so are all their marginal and conditional densities. Furthermore, the joint density of $(\theta_{i,a^*,t^*}, w_{i,a^*,t^*}, T_{i,1,a^*,t^*})$ is continuous, and so are all their marginal and conditional densities.

ii. *The random variables* w_{i,a^*,t^*}, $T_{i,1,a^*,t^*}$, *and* $T_{i,2,a^*,t^*}$ *are mutually independent conditional on* θ_{i,a^*,t^*}.

iii. *The conditional density functions* $f_{w_{i,a^*,t^*}|T_{i,1,a^*,t^*}}(w|t)$ *and* $f_{\theta_{i,a^*,t^*}|w_{i,a^*,t^*}}(\theta|w)$ *form a bounded complete family of distributions indexed by t and w, respectively.*

iv. *For all* $\theta_1, \theta_2 \in \Theta$, *the set* $\{t_2 : f_{T_{i,2,a^*,t^*}|\theta_{i,a^*,t^*}}(t_2|\theta_1) \neq f_{T_{i,2,a^*,t^*}|\theta_{i,a^*,t^*}}(t_2|\theta_2)\}$ *has positive probability whenever* $\theta_1 \neq \theta_2$.

v. *We normalize that* $E[w_{i,a^*,t^*}|\theta_{i,a^*,t^*}] = \theta_{i,a^*,t^*}$ *and that* $E[\varepsilon_{i,a^*,t^*}|\theta_{i,a^*,t^*}] = E[\eta_{i,1,a^*,t^*}|\theta_{i,a^*,t^*}] = E[\eta_{i,2,a^*,t^*}|\theta_{i,a^*,t^*}] = 0$.

Condition (i) is a mild restriction on the distribution and allows $T_{i,2,a^*,t^*}$ to be discrete. Conditions (ii), (iv), and (v) are immediately satisfied from the model construction. For example, the strict monotonicity of $\tau_2(\cdot)$ implies condition (iv). The completeness assumption in condition (iii) is widely used in the nonparametric identification literature and is satisfied in many classes of distributions, e.g. the exponential family. See, Hu and Schennach (2008) for further discussions on the completeness assumption.

Under Assumption 1, Theorem 1 in Hu and Schennach (2008) holds, and we can identify the joint and conditional densities $f_{T_{2,a^*,t^*},\theta_{a^*,t^*}}(\cdot,\cdot)$, $f_{w_{a^*,t^*}|\theta_{a^*,t^*}}(\cdot|\cdot)$, and $f_{T_{1,a^*,t^*},|\theta_{a^*,t^*}}(\cdot|\cdot)$ from Equation (6) therein. The measurement function $\tau_1(\cdot)$ is the conditional mean function of T_{1,a^*,t^*} given θ_{a^*,t^*} and can be identified by

$$\tau_1(\theta) = E[T_{1,a^*,t^*}|\theta] = \int t_1 f_{T_{1,a^*,t^*}|\theta_{a^*,t^*}}(t_1|\theta)dt_1.$$

The marginal density $f_{\theta_{a^*,t^*}}(\cdot)$ is identified by integrating the joint density:

$$f_{\theta_{a^*,t^*}}(\theta) = \int f_{T_{2,a^*,t^*},\theta_{a^*,t^*}}(t_2,\theta)dt_2.$$

Finally, the marginal densities $f_{\varepsilon_{a^*,t^*}}(\cdot)$ and $f_{\eta_{1,a^*,t^*}}(\cdot)$ are identified by the standard deconvolution method:

$$
\begin{aligned}
\varphi_{\varepsilon_{a^*,t^*}}(t) &= \varphi_{w_{a^*,t^*}}(t)/\varphi_{\theta_{a^*,t^*}}(t) \\
\varphi_{\eta_{1,a^*,t^*}}(t) &= \varphi_{T_{1,a^*,t^*}}(t)/\varphi_{\tau_1(\theta_{a^*,t^*})}(t),
\end{aligned}
$$

where $\varphi_x(\cdot)$ denotes the characteristic function of x.

C.2 Step 1 Estimation

Derivation of Equation (4.4):

$$
\begin{aligned}
& f_{w_{a^*,t^*},T_{1,a^*,t^*},\ldots,T_{J,a^*,t^*}}(w, T_1, \ldots, T_J) \\
&= \int_\Theta f_{w_{a^*,t^*},T_{1,a^*,t^*},\ldots,T_{J,a^*,t^*},\theta_{a^*,t^*}}(w, T_1, \ldots, T_J, \theta) d\theta \\
&= \int_\Theta f_{w_{a^*,t^*}|T_{1,a^*,t^*},\ldots,T_{J,a^*,t^*},\theta_{a^*,t^*}}(w|T_1, \ldots, T_J, \theta) \\
&\quad \times f_{T_{1,a^*,t^*}|T_{2,a^*,t^*},\ldots,T_{J,a^*,t^*},\theta_{a^*,t^*}}(T_1|T_2, \ldots, T_J, \theta) \times \cdots \times f_{T_{J,a^*,t^*}|\theta_{a^*,t^*}}(T_J|\theta) f_{\theta_{a^*,t^*}}(\theta) d\theta \\
&= \int_\Theta f_{w_{a^*,t^*}|\theta_{a^*,t^*}}(w|\theta) f_{T_{1,a^*,t^*}|\theta_{a^*,t^*}}(T_1|\theta) \times \cdots \times f_{T_{J,a^*,t^*}|\theta_{a^*,t^*}}(T_J|\theta) f_{\theta_{a^*,t^*}}(\theta) d\theta \\
&= \int_\Theta f_{\varepsilon_{a^*,t^*}|\theta_{a^*,t^*}}(\varepsilon|\theta) f_{\eta_{1,a^*,t^*}|\theta_{a^*,t^*}}(\eta_1|\theta) \times \cdots \times f_{\eta_{J,a^*,t^*}|\theta_{a^*,t^*}}(\eta_J|\theta) f_{\theta_{a^*,t^*}}(\theta) d\theta \\
&= \int_\Theta f_{\varepsilon_{a^*,t^*}}(w - \theta; \boldsymbol{\beta}_{\varepsilon_{a^*,t^*}}) \\
&\quad \times f_{\eta_{1,a^*,t^*}}(T_1 - \tau_1(\theta; \boldsymbol{\beta}_{\tau_1}); \boldsymbol{\beta}_{\eta_{1,a^*,t^*}}) \times \cdots \times f_{\eta_{J,a^*,t^*}}(T_J - \tau_J(\theta; \boldsymbol{\beta}_{\tau_J}); \boldsymbol{\beta}_{\eta_{J,a^*,t^*}}) f_{\theta_{a^*,t^*}}(\theta; \boldsymbol{\beta}_{\theta_{a^*,t^*}}) d\theta.
\end{aligned}
$$

Consider mixtures of normal distributions for the distributions of ε_{a^*,t^*}, η_{j,a^*,t^*} of continuous measure, and θ_{a^*,t^*}:

$$
f_{\varepsilon_{a^*,t^*}}(w - \theta; \boldsymbol{\beta}_{\varepsilon_{a^*,t^*}}) = \sum_{n_\varepsilon} p_{\varepsilon_{a^*,t^*},n_\varepsilon} \frac{1}{\sqrt{2\pi\sigma^2_{\varepsilon_{a^*,t^*},n_\varepsilon}}} \exp\left(-\frac{(w - \theta - \mu_{\varepsilon_{a^*,t^*},n_\varepsilon})^2}{2\sigma^2_{\varepsilon_{a^*,t^*},n_\varepsilon}}\right),
$$

$$
\begin{aligned}
&f_{\eta_{j,a^*,t^*}}(T_j - \tau_j(\theta;\boldsymbol{\beta}_{\tau_j});\boldsymbol{\beta}_{\eta_{j,a^*,t^*}}) \\
&= \sum_{n_{\eta_j}} p_{\eta_{j,a^*,t^*},n_{\eta_j}} \frac{1}{\sqrt{2\pi\sigma^2_{\eta_{j,a^*,t^*},n_{\eta_j}}}} \exp\left(-\frac{(T_j - \tau^j(\theta;\boldsymbol{\beta}_{\tau_j}) - \mu_{\eta_{j,a^*,t^*},n_{\eta_j}})^2}{2\sigma^2_{\eta_{j,a^*,t^*},n_{\eta_j}}}\right),
\end{aligned}
$$

$$
f_{\theta_{a^*,t^*}}(\theta;\boldsymbol{\beta}_{\theta_{a^*,t^*}}) = \sum_{n_\theta} p_{\theta_{a^*,t^*},n_\theta} \frac{1}{\sqrt{2\pi\sigma^2_{\theta_{a^*,t^*},n_\theta}}} \exp\left(-\frac{(\theta - \mu_{\theta_{a^*,t^*},n_\theta})^2}{2\sigma^2_{\theta_{a^*,t^*},n_\theta}}\right),
$$

where $\boldsymbol{\beta}_x := (p_{x,n_x}, \mu_{x,n_x}, \sigma_{x,n_x})$ and $\sum_{n_x} p_{x,n_x} = 1$ for $x = \varepsilon_{a^*,t^*}, \eta_{j,a^*,t^*}$, and θ_{a^*,t^*}.[1] Note that the location restriction (i.e. $E[x] = 0$) implies that $\sum_{n_x} p_{x,n_x}\mu_{x,n_x} = 0$.

If the test measure T_j is discrete, we assume that it is generated from an ordered probit model. Suppose it has K_j discrete values: $T_j \in \{1, ..., K_j\}$. We need to estimate $K_j - 1$ cutoff values for the ordered probit, i.e., $\chi_j := (\chi_{j,1}, \chi_{j,2}, ..., \chi_{j,K_j-1})$. The density function is:

$$
f_{\eta_{j,a^*,t^*}}(T_j - \tau_j(\theta;\boldsymbol{\beta}_{\tau_j});\chi_j) = \sum_{k=1}^{K_j} \mathbb{1}(T_j = k)\left[\Phi\left(\chi_{j,k} - \tau_j(\theta;\boldsymbol{\beta}_{\tau_j})\right) - \Phi\left(\chi_{j,k-1} - \tau_j(\theta;\boldsymbol{\beta}_{\tau_j})\right)\right],
$$

with $\chi_{j,0} = -\infty$ and $\chi_{j,K_j} = \infty$. $\Phi(\cdot)$ is the cdf of the standard normal distribution.

Since we have 4 measures: two are continuous (T_1 and T_2) and two are discrete (T_3 and T_4), the log-likelihood for individual i at age a^* in year t^* is:

$$
\begin{aligned}
\ell_{i,a^*,t^*} = \log \int_{-\infty}^{\infty} &\left[\sum_{n_\varepsilon} p_{\varepsilon_{a^*,t^*},n_\varepsilon} \frac{1}{\sqrt{2\pi\sigma^2_{\varepsilon_{a^*,t^*},n_\varepsilon}}} \exp\left(-\frac{(w_{i,a^*,t^*} - \theta - \mu_{\varepsilon_{a^*,t^*},n_\varepsilon})^2}{2\sigma^2_{\varepsilon_{a^*,t^*},n_\varepsilon}}\right)\right] \\
&\times \left[\sum_{n_{\eta_1}} p_{\eta_{1,a^*,t^*},n_{\eta_1}} \frac{1}{\sqrt{2\pi\sigma^2_{\eta_{1,a^*,t^*},n_{\eta_1}}}} \exp\left(-\frac{(T_{i,1,a^*,t^*} - \tau_1(\theta;\boldsymbol{\beta}_{\tau_1}) - \mu_{\eta_{1,a^*,t^*},n_{\eta_1}})^2}{2\sigma^2_{\eta_{1,a^*,t^*},n_{\eta_1}}}\right)\right] \\
&\times \left[\sum_{n_{\eta_2}} p_{\eta_{2,a^*,t^*},n_{\eta_2}} \frac{1}{\sqrt{2\pi\sigma^2_{\eta_{2,a^*,t^*},n_{\eta_2}}}} \exp\left(-\frac{(T_{i,2,a^*,t^*} - \tau_2(\theta;\boldsymbol{\beta}_{\tau_2}) - \mu_{\eta_{2,a^*,t^*},n_{\eta_2}})^2}{2\sigma^2_{\eta_{2,a^*,t^*},n_{\eta_2}}}\right)\right] \\
&\times \left[\sum_{k_3=1}^{K_3} \mathbb{1}(T_{i,3,a^*,t^*} = k_3)\left[\Phi\left(\chi_{3,k_3} - \tau_3(\theta;\boldsymbol{\beta}_{\tau_3})\right) - \Phi\left(\chi_{3,k_3-1} - \tau_3(\theta;\boldsymbol{\beta}_{\tau_3})\right)\right]\right]
\end{aligned}
$$

[1] For expositional purposes, we assume that the number of distributions for each random variable mixture (i.e. n_ε, n_{η_j}, and n_θ) do not vary with age and time.

$$
\begin{aligned}
&\times \left[\sum_{k_4=1}^{K_4} \mathbb{1}(T_{i,4,a^*,t^*} = k_4) \left[\Phi \left(\chi_{4,k_4} - \tau_4(\theta; \boldsymbol{\beta}_{\tau_4}) \right) - \Phi \left(\chi_{4,k_4-1} - \tau_4(\theta; \boldsymbol{\beta}_{\tau_4}) \right) \right] \right] \\
&\times \left[\sum_{n_\theta} p_{\theta_{a^*,t^*,n_\theta}} \frac{1}{\sqrt{2\pi\sigma^2_{\theta_{a^*,t^*,n_\theta}}}} \exp\left(-\frac{(\theta - \mu_{\theta_{a^*,t^*,n_\theta}})^2}{2\sigma^2_{\theta_{a^*,t^*,n_\theta}}} \right) \right] d\theta \\
&= \log \sum_{n_\varepsilon, n_{\eta_1}, n_{\eta_2}, k_3, k_4, n_\theta} p_{\varepsilon_{a^*,t^*,n_\varepsilon}} p_{\eta_{1,a^*,t^*,n_{\eta_1}}} p_{\eta_{2,a^*,t^*,n_{\eta_2}}} p_{\theta_{a^*,t^*,n_\theta}} \\
&\times \frac{1}{\sqrt{2^4 \pi^4 \left(\sigma_{\varepsilon_{a^*,t^*,n_\varepsilon}} \sigma_{\eta_{1,a^*,t^*,n_{\eta_1}}} \sigma_{\eta_{2,a^*,t^*,n_{\eta_2}}} \sigma_{\theta_{a^*,t^*,n_\theta}} \right)^2}} \\
&\times \mathbb{1}(T_{i,3,a^*,t^*} = k_3) \times \mathbb{1}(T_{i,4,a^*,t^*} = k_4) \\
&\times \int_{-\infty}^{\infty} \exp\Bigg(- \frac{(w_{i,a^*,t^*} - \theta - \mu_{\varepsilon_{a^*,t^*,n_\varepsilon}})^2}{2\sigma^2_{\varepsilon_{a^*,t^*,n_\varepsilon}}} - \frac{(T_{i,1,a^*,t^*} - \tau_1(\theta; \boldsymbol{\beta}_{\tau_1}) - \mu_{\eta_{1,a^*,t^*,n_{\eta_1}}})^2}{2\sigma^2_{\eta_{1,a^*,t^*,n_{\eta_1}}}} \\
&- \frac{(T_{i,2,a^*,t^*} - \tau_2(\theta; \boldsymbol{\beta}_{\tau_2}) - \mu_{\eta_{2,a^*,t^*,n_{\eta_2}}})^2}{2\sigma^2_{\eta_{2,a^*,t^*,n_{\eta_2}}}} - \frac{(\theta - \mu_{\theta_{a^*,t^*,n_\theta}})^2}{2\sigma^2_{\theta_{a^*,t^*,n_\theta}}} \Bigg) \\
&\times \left[\Phi \left(\chi_{3,k_3} - \tau_3(\theta; \boldsymbol{\beta}_{\tau_3}) \right) - \Phi \left(\chi_{3,k_3-1} - \tau^3(\theta; \boldsymbol{\beta}_{\tau_3}) \right) \right] \\
&\times \left[\Phi \left(\chi_{4,k_4} - \tau_4(\theta; \boldsymbol{\beta}_{\tau_4}) \right) - \Phi \left(\chi_{4,k_4-1} - \tau^4(\theta; \boldsymbol{\beta}_{\tau_4}) \right) \right] d\theta.
\end{aligned}
$$

$\boldsymbol{\beta}_{a^*,t^*}$ is estimated by maximizing the log-likelihood function:

$$
\widehat{\boldsymbol{\beta}}_{a^*,t^*} = \underset{\beta_{a^*,t^*} \in \mathcal{B}}{\arg\max} \frac{1}{|\mathcal{I}_{c^*}|} \sum_{i \in \mathcal{I}_{c^*}} \ell_{i,a^*,t^*}.
$$

Standard Errors

Define the score of the log-likelihood for observation i as follows:

$$
\widehat{\boldsymbol{S}}_{i,a^*,t^*} = \boldsymbol{S}_{i,a^*,t^*}(\widehat{\boldsymbol{\beta}}_{a^*,t^*}) = \frac{\partial \ell_{i,a^*,t^*}(\widehat{\boldsymbol{\beta}}_{a^*,t^*})}{\partial \boldsymbol{\beta}^\top_{a^*,t^*}},
$$

and the Hessian:

$$
\widehat{\boldsymbol{H}}_{i,a^*,t^*} = \boldsymbol{H}_{i,a^*,t^*}(\widehat{\boldsymbol{\beta}}_{a^*,t^*}) = \frac{\partial^2 \ell_{i,a^*,t^*}(\widehat{\boldsymbol{\beta}}_{a^*,t^*})}{\partial \boldsymbol{\beta}^\top_{a^*,t^*} \partial \boldsymbol{\beta}_{a^*,t^*}}.
$$

The asymptotic variance matrix is:

$$\widehat{\boldsymbol{V}}_{ML,a^*,t^*} = \left(\sum_{i\in\mathcal{I}_{c^*}} \widehat{\boldsymbol{H}}_{i,a^*,t^*}\right)^{-1} \left(\sum_{i\in\mathcal{I}_{c^*}} \widehat{\boldsymbol{S}}_{i,a^*,t^*}\widehat{\boldsymbol{S}}_{i,a^*,t^*}^{\top}\right)\left(\sum_{i\in\mathcal{I}_{c^*}} \widehat{\boldsymbol{H}}_{i,a^*,t^*}\right)^{-1}.$$

C.3 GMM Estimation with a Linear Measure

Let $\boldsymbol{\Lambda}$ be a vector of parameters to be estimated in the second stage. We use the generalized methods of moments (GMM) to estimate $\boldsymbol{\Lambda}$. Suppose the total number of covariances is M and let $m = 1, ..., M$ be the index of the covariances. Define the theoretical covariance vector as $\boldsymbol{h}(\boldsymbol{\Lambda}) = (h_1(\boldsymbol{\Lambda}), ..., h_M(\boldsymbol{\Lambda}))^{\top}$. Let $d_{i,m}$ be the indicator of whether individual i contributes to the m^{th} covariance. Then we can write individual i's contribution to the m^{th} moment as $g_m(\boldsymbol{z_i}, \boldsymbol{\Lambda})$ where $\boldsymbol{z_i}$ includes $d_{i,m}$, individual i's log wage residuals, and cognitive measures. $g_m(\boldsymbol{z_i}, \boldsymbol{\Lambda})$ is equal to $d_{i,m}$ times the difference between the product of corresponding de-meaned variables and the theoretical covariance. For example, individual i's contribution to the moment involving covariance $Cov(w_{a,t}, T_{j,a+k,t+k})$ is $g_m(\boldsymbol{z_i}, \boldsymbol{\Lambda}) = d_{i,m}[(w_{i,a,t} - \bar{w}_{a,t})(T_{i,j,a+k,t+k} - \bar{T}_{j,a+k,t+k}) - h_m(\boldsymbol{\Lambda})]$.

Let $\boldsymbol{g}(z, \boldsymbol{\Lambda}) = (g_1(z, \boldsymbol{\Lambda}), ..., g_M(z, \boldsymbol{\Lambda}))^{\top}$. Then the following moment condition holds at the true parameter $\boldsymbol{\Lambda}_0$:

$$\mathbb{E}[\boldsymbol{g}(z, \boldsymbol{\Lambda}_0)] = 0.$$

The GMM estimator $\widehat{\boldsymbol{\Lambda}}$ solves

$$\min_{\boldsymbol{\Lambda}} \left[\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{g}(\boldsymbol{z_i}, \boldsymbol{\Lambda})\right]^{\top} \boldsymbol{W} \left[\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{g}(\boldsymbol{z_i}, \boldsymbol{\Lambda})\right],$$

where $\boldsymbol{W}$ is the weighting matrix.

Standard Errors

The GMM estimator $\widehat{\boldsymbol{\Lambda}}$ is asymptotically normal with a variance matrix

$$\boldsymbol{V}_{GMM} = (\boldsymbol{G}^{\top}\boldsymbol{W}\boldsymbol{G})^{-1}(\boldsymbol{G}^{\top}\boldsymbol{W}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{G})(\boldsymbol{G}^{\top}\boldsymbol{W}\boldsymbol{G})^{-1}/N,$$

where $\boldsymbol{G}$ is the Jacobian of the vector of moments, $\mathbb{E}[\partial \boldsymbol{g}(z, \boldsymbol{\Lambda}_0)/\partial \boldsymbol{\Lambda}_0^\top]$, and $\boldsymbol{\Omega} = \mathbb{E}[\boldsymbol{g}(z, \boldsymbol{\Lambda}_0)\boldsymbol{g}(z, \boldsymbol{\Lambda}_0)^\top]$. To calculate the asymptotic variance matrix, both expectations are replace by sample averages and evaluated at the estimated parameters:

$$\widehat{\boldsymbol{G}} = \frac{1}{N}\sum_{i=1}^{N} \frac{\partial \boldsymbol{g}(z_i, \widehat{\boldsymbol{\Lambda}})}{\partial \boldsymbol{\Lambda}^\top} = -\boldsymbol{W}^{-\frac{1}{2}} \frac{\partial \boldsymbol{h}(\widehat{\boldsymbol{\Lambda}})}{\partial \boldsymbol{\Lambda}^\top},$$

$$\widehat{\boldsymbol{\Omega}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{g}(z_i, \widehat{\boldsymbol{\Lambda}})\boldsymbol{g}(z_i, \widehat{\boldsymbol{\Lambda}})^\top.$$

We can test r linear parameter restrictions $H_0 : \boldsymbol{R}\boldsymbol{\Lambda} = 0$ using Wald test statistic:

$$(\boldsymbol{R}\widehat{\boldsymbol{\Lambda}})^\top(\boldsymbol{R}\widehat{\boldsymbol{V}}_{\boldsymbol{GMM}}\boldsymbol{R})^{-1}(\boldsymbol{R}\widehat{\boldsymbol{\Lambda}}) \xrightarrow{d} \chi^2_r.$$

C.4 Data

C.4.1 Cognitive Measures

Details on the construction of the four cognitive measures are as follows:

- Word recall. In the data, there are two separate tasks to assess respondent's memory: one is immediate word recall and the other is delayed word recall. During the interview, the interviewer read a list of 10 nouns to the respondent and asked the respondent to recall as many words as possible from the list in any order. After approximately 5 minutes of answering other survey questions, the respondent was asked to recall the nouns previously presented. We construct a single measure which is the sum of the number of nouns that the respondent recalled in the two tasks. This measure ranges from 0 to 20.

- Serial 7's. This test asks the respondent to subtract 7 from the prior number, beginning with 100 for five trials. Correct subtractions are based on the prior number given, so that even if one subtraction is incorrect subsequent trials are evaluated on the given (perhaps wrong) answer. This test score ranges from 0 to 5.

- Quantitative reasoning. In HRS 2002, three questions were added to the core survey to assess respondents numerical ability:

1. "Next I would like to ask you some questions which assess how people use numbers in everyday life. If the chance of getting a disease is 10 percent, how many people out of 1,000 would be expected to get the disease?"

2. "If 5 people all have the winning numbers in the lottery and the prize is two million dollars, how much will each of them get?"

3. "Let's say you have $200 in a savings account. The account earns ten percent interest per year. How much would you have in the account at the end of two years?"

We construct a single measure called quantitative reasoning using the answers from these three questions. For the first two questions, the respondent gets 1 if the answer is correct and 0 otherwise. For the last question, the respondent gets 2 if the answer is correct. If the respondent used 10% as a simple interest rate rather than a compound interest rate, i.e., answered 240 instead of 242, he gets 1. The quantitative reasoning measure is the sum of scores of all three questions and it ranges from 0 to 4.

- Retrieval fluency. This task was first incorporated in the HRS in the 2010 wave. During this task, respondents were asked to name as many animals as they could withing a 60-second time limit. The retrieval fluency measure is constructed as the number of total animal answered minus the number of incorrect names. The value of this measure ranges from 0 to 90.

C.4.2 Age

The age variable we use is the age at the end of the interview. According to the HRS, when there are different beginning and ending interview dates, most of the interview is usually conducted on the ending date. So it is recommended to use age at the end of interview date for respondent age at each interview.

The interval between interviews is usually 2 years. But about 5-10% of the sample was interviewed a year later than the wave year. For example, the normal case would be someone at age 52 interviewed in 1998 and age 54 interviewed in 2000. But it could be the case that he was interviewed in 2001 for the second interview at age 55. Another case could be he was aged 53 when interviewed in 1999 and aged 54 when interviewed in 2000. In these cases, we assume that age at the first interview is the age at that wave year and the subsequent interviews are two years apart. So for the first case, the wages we observe are $w_{a=52,t=1998}$ and $w_{a=55,t=2001}$ and we assume they are $w_{a=52,t=1998}$ and $w_{a=54,t=2000}$. For the second case, we observe $w_{a=53,t=1999}$ and $w_{a=54,t=2000}$ and we assume they are $w_{a=53,t=1998}$ and $w_{a=55,t=2000}$.

Another approach is to use the birth year to calculate age at each wave year. Then we would assume that we observe $w_{a=52,t=1998}$ and $w_{a=54,t=2000}$ for both the first and second cases. The results are quite similar and do not drive any particular patterns using this alternative approach.

Curriculum Vitae

Name:	Qian Liu
Post-Secondary Education and Degrees:	Renmin University of China Beijing, China 2006-2010 B.A. in Economics
	Renmin University of China Beijing, China 2010-2012 M.A. in Economics
	University of Western Ontario London, ON 2012 - 2020 Ph.D. in Economics
Honours and Awards:	Ontario Trillium Scholarship 2012-2016
Related Work Experience:	Research and Teaching Assistant The University of Western Ontario 2012 - 2019
	Instructor The University of Western Ontario 2017 - 2018
	Lecturer (tenure-track) Brock University 2019 -