Essays in Household Finance

by

Natalie M. Cox

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Benjamin Handel, Chair
Professor David Sraer
Professor Ulrike Malmendier
Professor Emmanuel Saez

Spring 2017

Essays in Household Finance

Abstract

Essays in Household Finance

by

Natalie M. Cox

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Benjamin Handel, Chair

The use of technology by firms is changing the way insurance and lending markets function. I study the financial technology, or "fin-tech", industry, which is characterized by a growing number of online lenders who use data on educational, employment, and financial outcomes to quickly assess the risk of prospective borrowers and offer individualized loan terms. In many ways, their financial "innovations" can be thought of as movements towards more personalized products: interest rates that better reflect individuals' risk, payment plans that are tailored to individuals' monthly income and expenditures, and user-friendly interfaces that make financial decisions more intuitive and uncomplicated. On an individual level, as firms expand and customize product offerings, there is the potential for large efficiency gains. These innovations could also have wider implications for market structure; for example, if more accurate risk-based pricing creates clear winners and losers, it will change the distribution of consumer surplus.

Advances in data-driven underwriting have both efficiency and equity implications for consumer lending markets where private and public credit options coexist. In the $1 trillion student loan market, private lenders now offer a growing distribution of risk-based interest rates, while the federally-run loan program sets a break-even, uniform interest rate. In my first chapter, I measure the overall gains in consumer surplus from such risk-based pricing and quantify the redistributional consequences of low-risk types refinancing out of the government pool into the private market. The empirical analysis is based on a unique applicant-level dataset from an online refinancing firm that contains information on loan terms, household balance sheets, and risk-based interest rates. I first leverage a series of firm-conducted interest rate experiments to estimate the sensitivity of borrowers' maturity and refinancing choices to interest rates. Using the maturity response, I then estimate a structural model of borrowers' repayment preferences. Using the estimated model, I show that comprehensive risk-based pricing generates large absolute gains in welfare of $480 per borrower relative to a break-even uniform price, and $400 relative to a coarser method of FICO-based pricing. If the federal pool conducts breakeven pricing, these efficiency gains come at a direct equity cost – low risk surplus will increase on average by $2,300, while high risk surplus will fall by

$2,100. In order to maintain access to the current uniform rate, the government would have to transition from break-even pricing to an average net subsidy of $2,080 per borrower.

In the second chapter, I empirically analyze the fixed and variable rate decisions of borrowers who are financing large personal loans, and are given the option to switch rate types at any point. Many online lending firms now offer financial products that are more flexible and personalized than traditional loans; however, little is known about how consumers will interact with these more complete, but also more complex, contracts. Over my sample time period, the market index interest rate for the fixed and variable rate loans changed considerably. I first present reduced form evidence on the determinants of borrowers' initial rate decisions and the presence of switching costs, and then estimate a structural model that maps these findings to the coefficient of absolute risk aversion and a switching cost parameter. I compare the active and inactive rate choices of borrowers in different interest rate environments to separately identify switching costs from risk preferences.I show that while initial rate choices are very responsive to the prevailing interest rate environment, very few borrowers ever take advantage of the option to switch rate types even when interest rates increase. Specifically, I estimate a risk aversion parameter of .0564, which implies that borrowers are very risk averse, and lower and upper bounds on switching costs from $166 to $1,185. I also show that both the initial probability of choosing a variable rate loan and the probability of never switching are positively correlated with borrower liquidity constraints, which suggests that these borrowers are more focused on current monthly payments than future interest rate risk.

This thesis is dedicated with thanks and admiration to my selfless parents, who always made my education a priority and taught me to think critically and creatively.

Contents

List of Figures

List of Tables

Acknowledgments

I thank my advisors Ben Handel and David Sraer for their incredible guidance and support, generosity, insightful suggestions, and for providing such a fruitful and open environment for research. I also have benefited greatly from comments of faculty at UC Berkeley, including my committee members, Ulrike Malmendier and Emmanuel Saez, as well as Alan Auerbach, Kei Kawai, Patrick Kline, Jon Kolstad, Danny Yagan and Gabriel Zucman.

This work has benefited particularly from the comments of my classmates Zarek Brot-Goldberg, Jennifer Kwok, Julien Lafortune, Waldo Ojeda, Jon Schellenberg, Avner Schlain, and Katalin Springel, and seminar participants at UC Berkeley.

I am incredibly grateful to Pierre Bachas, who is my constant source of morale, friendship, and economic wisdom, and who made graduate school a delight.

Finally, I am very grateful to my family for their encouragement, teaching, advice and support throughout my life. I thank my parents, Camille and Allen, my three inspiring sisters Genevieve, Veronica, and Katherine, who are my role models and best friends, my sweet nieces Camilla, Evelyn, and Juliette, and Finn.

Chapter 1

Pricing, Selection, and Welfare in the Student Loan Market: Evidence from Borrower Repayment Decisions

1.1 Introduction:

Technological improvements in data-driven underwriting are transforming consumer lending, giving rise to a new financial technology, or "fin-tech", industry. Online lenders quickly assess the risk of prospective borrowers using data on educational, employment, and financial outcomes, and offer more individualized loan terms. More accurate, sophisticated pricing algorithms have been shown in theory and practice to reduce asymmetric information and correct inefficiencies in many areas of consumer lending, including credit cards, mortgages, and auto lending. This paper focuses on the growing number of firms who use this information to underwrite and refinance the student debt of borrowers who have finished schooling.[1] For student borrowers, who are young and have under-developed credit histories, the gains from non-traditional scoring methods and the ability to refinance their federal debt at lower interest rates could be large.

However, advances in risk-based pricing could have complex implications for how private and public lending options coexist. This is especially true in the student loan space, where the federally-run Direct Loan program currently dominates the student loan origination market, but does not use risk-based pricing. It instead pools all borrowers together at uniform interest rates with the goal of being revenue neutral.[2] This uniform interest rate policy

[1]Borrowers take out debt when they begin school, but begin repaying that debt only upon completion. When a loan is refinanced, the federal government, which does not have a pre-payment penalty, is paid off by the private firm, which then takes over the servicing and liabilities associated with the loan.

[2]This document describes the break-even goal in detail: http://www.gao.gov/assets/670/660548.pdf

has clear equity implications: it allows all borrowers to access credit for higher education despite observable variation in the expected costs of lending. However, the "break-even" nature of the price produces a potential tension between the public program's equity goals and emerging private-sector developments in risk-based pricing: as low risk types refinance into the private sector, the average risk, and break-even interest rate, of the remaining federal borrowers will rise. The relatively high unsubsidized interest rates in the Direct Loan program, which have reached levels over 7% in the last decade, perhaps already reflect the impact of sorting between sectors.

In this paper, I study this efficiency-equity tradeoff with a proprietary dataset of applicants from an online student loan refinancing firm that employs comprehensive risk-based pricing. Using a series of firm-conducted pricing experiments, I first show that observationally similar borrowers choose shorter maturities when the interest rates they are offered increase.[3] I then estimate a structural choice model that ties this reduced form maturity elasticity to the underlying parameters governing expected utility, namely the intertemporal elasticity of substitution. In a series of counterfactuals, I use this model to both measure the gains from private sector innovations in risk-based pricing and analyze how these gains are distributed over borrowers of different risk types. I show that current advances in risk-based pricing generate absolute gains in welfare relative to a break-even uniform pricing scheme by $480 on average per borrower, and that these gains are sizable in comparison to coarser modes of pricing based only on FICO score. However, because low risk types will sort into the private sector, these gains will be concentrated amongst low risk borrowers.

While I focus on the market to refinance student loans, this analysis is applicable to several lending and insurance markets where risk-scoring technologies are advancing and where public and private credit options coexist. The closely linked private student loan *origination* market also shows signs of selection: in 2011, 40% of new borrowers had FICO scores greater than 770, while less than 5% had scores below 670[4]. In car insurance, firms now use data from tracking devices to accurately identify and cherry-pick the lowest risk drivers. My methods could also be used to analyze government policy responses observed in similar settings: in the mortgage market, FHA-backed loan eligibility is predicated on risk-related factors like FICO score, but subsidies are provided for lower income households. The health insurance market takes a more regulatory approach, by limiting the set of risk-related factors that can determine premiums.

The dataset I use provides precise, borrower-level information on risk-score inputs and formulas, interest rates, repayment decisions, and balance sheets. It is powerful for several reasons: first, the dataset contains both borrowers who fully refinance their loans (N=12,000), as well as a larger, more representative sample of borrowers who are shown an interest rate quote and may or may not fully refinance (N = 200,000). Second, the dataset allows me to

[3] Maturity is the relevant response for student borrowers in repayment, given that their loan principal is already fixed.

[4] The source for these numbers is the 2012 CFPB Private Student Loan Report. The report notes that "During the boom years, the lowest credit deciles were the most heavily populated. After the financial crisis, the distribution reversed...Today, only a very good credit is likely to be approved"

link maturity and refinancing decisions to a wide set of socioeconomic and risk-related variables describing the borrower, including their income, degree type, occupation, employment history, location, age, credit score and report, assets, liabilities, and savings behavior. Firms refinancing student debt can price on many more observable characteristics (applicants' financial accounts, employment outcomes, and educational histories) than firms originating student debt to new borrowers, who have yet to realize risks like education completion and finding initial employment. These additional variables not only provide valuable econometric controls in many specifications, but also allow for a more comprehensive analysis of how the student debt interacts with other elements of the household balance sheet.

Finally, the dataset contains exogenous variation in interest rates *within* risk type that allows me to measure how borrower surplus changes under different interest rate regimes. This variation consists of 10 interest rate changes that were conducted at a firm-wide level to gather quasi-experimental evidence on maturity and application elasticities. The repeated cross-sectional nature of the data means that I can compare the repayment choices of observably similar borrowers (matched on characteristics like debt, income, risk, and age) under different price regimes. Taken together, these features of the data allow for a detailed analysis of how credit-scoring and refinancing innovations will impact borrowers of different risk and income levels, influence sorting into the private sector, and impact federal costs. They also permit a novel micro-analysis of student borrowers' financial behavior and intertemporal optimization.

I leverage the experimental variation in price schedules *within* risk type to measure how borrowers respond to interest rates on two budget relevant repayment responses: 1) loan maturity, and 2) the propensity to refinance. Loan maturity is perhaps the most fundamental decision made by the borrower during repayment, given that the choice of how much to borrow is already fixed. This means maturity is the only means a borrower has during repayment to lower monthly payments and increase immediate liquidity. Borrowers reveal their intertemporal preferences when making a maturity decision: by extending maturity they reduce their monthly payment, but increase their total interest paid. Reduced form evidence reveals that the borrowers in my sample decrease maturity when interest rates increase, and that this sensitivity increases with borrower quality. This suggests that low risk borrowers are focused on minimizing interest payments, while borrowers who have lower incomes are more focused on minimizing monthly payments. It also suggests that low risk borrowers would have larger welfare gains from refinancing at lower interest rates.

Next, I use this maturity response to estimate an expected utility model of borrowers' repayment preferences, in which borrowers first choose a maturity, and then whether or not to refinance in the private sector. The structural model dictates how differences in income, debt, and interest rates should impact maturity choices under the assumption of CRRA utility. It allows me to map the reduced form maturity elasticity to the underlying parameter of interest, the intertemporal elasticity of substitution (IES). At their optimal term, borrowers balance the gain in marginal utility from having a slightly lower monthly payment over the life of the loan against the cost of paying more interest overall. I use the first order condition that captures this tradeoff to estimate the model using non-linear least squares. Heterogeneity in income levels, growth and volatility, which I model and estimate as a function of observable characteristics, generates differences in the level of maturity choices across borrowers. I use the exogenous *within* risk type interest rate variation, and resulting shifts in the term distribution, to identify the IES. The value of the IES is central to my question, since it governs borrowers' interest rate sensitivity. I find a moderately high IES of .8, which suggests borrowers are sensitive to interest rate levels, and seems plausible in a setting where borrowers are well aware of the intertemporal tradeoffs they are making.

I use the estimated model of borrower utility to perform a series of counterfactual analyses that investigate how innovations in pricing and government policy responses impact selection and welfare. In addition to the expected utility model, the counterfactuals use the observed risk-specific interest rates that were offered to each borrower by the refinancing firm. These prices allow me to precisely measure the price "wedge" borrowers of different risk types face in the public sector vs the private sector. To use these observed prices, I must assume that the private refinancing sector is perfectly competitive, and therefore the prices I observe reflect firms' expected costs of lending.[5]

My first counterfactual measures how borrower sorting between public and private repayment options, and consequently consumer surplus, changes as private sector firms price on additional borrower characteristics (including income, savings, school rank, and degree type) beyond FICO score. My findings highlight how developments in the private sector's ability to price borrower risk will simultaneously i) improve welfare for low risk borrowers and ii) increase sorting of low risk borrowers out the public repayment pool. The use of comprehensive risk-based pricing improves borrower welfare on average by \$480 relative to a setting where only a uniform, breakeven interest rate is available, more than 50% of the average monthly payment in my sample. I show these gains are sizable even in comparison to coarser, more traditional modes of risk based pricing – using FICO score for example, improves welfare by only \$64 on average per borrower. This suggests that using less traditional risk based pricing algorithms that consider additional borrower characteristics like savings, income, and education can benefit individuals who are low risk, but have underdeveloped credit histories (e.g. the student borrower population, which is high debt, but young).

In a second counterfactual, I show that how these welfare gains are distributed across borrowers of different risk levels, and the degree to which the Direct Loan program unravels, is predicated on the government's policy response. If the government continues to set a breakeven, uniform rate, risk-based pricing innovations will come at a direct equity cost – average gains of \$2,300 will be concentrated amongst low risk borrowers, while high risk borrowers will suffer an average welfare loss of \$2,100 after the federal interest rate increases by .6 pp. I show that transitioning instead to a net interest rate subsidy would both prevent unraveling in the public sector and allow low risk types to benefit from risk-based pricing. To maintain the current uniform rate would require a \$2,090 subsidy on average per borrower.

[5] I argue this assumption seems reasonable given the rapid growth in the sector over the last 4 years – the first refinancing firms emerged in 2011, and since then the market has expanded to include more than a dozen major players.

My model highlights how the effective size of an interest rate subsidy can deviate from the mechanical size once borrower selection and maturity responses are accounted for. For example, lowering the lower federal interest rate mechanically by .6 percentage points will reduce refinancing into the private sector, reduce the average risk of the remaining federal borrowers, and therefore translate into an effective subsidy of only .2 pp.

I also consider how refinancing frictions impact selection and welfare. I use empirical refinancing elasticities to model a setting with frictions, where the propensity to refinance increases with the size of interest rate savings. With refinancing frictions, the extent of low-risk selection into the private sector is reduced significantly, which reduces selection into the private sector and generates an average surplus gain of only $208 per borrower. Frictions also mean that coarse vs. fine-grained pricing algorithms have different equity implications. In a setting with no refinancing frictions these innovations increase efficiency at little equity cost: they extend larger savings to low risk types, but do not expand the pool of refinancers at the extensive margin and thus have little impact on selection. However, in a setting with frictions, low risk types have a higher propensity to refinance when offered comprehensive pricing, leading to greater low-risk selection into the private market.

This paper relates to several literatures: first, it contributes research on how borrowers finance their higher education with student loans. While work has primarily focused on optimal borrowing limits and repayment structures (Lochner and Monge-Naranjo 2015; Rothstein & Rouse 2011; Avery & Turner 2012; Yannelis 2015; Lucas & Moore 2010, Beyer, Hastings, Neilson, & Zimmerman (2015)), and are therefore applicable to policies that increase or decrease loan limits, my paper is the first empirical analysis of how interest rates can be used as a policy instrument. The question of how borrowers more generally respond to interest rates, and how these responses are determined by credit constraints, is central to the household finance literature (Adams, Einav & Levin; Gross & Souleles 2002; Martins & Villanueva 2006; Karlan & Zinman 2005). Several papers that study the role of maturity in consumer finance outside the realm of student loans (Attanasio, Goldberg, & Kyriazidou 2008; Hertzberg, Liberman, & Paravisini 2016) show borrowers' maturity choices are influenced by borrowers' liquidity constraints, and are preferred by riskier borrowers. Demonstrating how student borrowers manage repayment once principal is fixed by manipulating loan maturity is important since maturity extensions are now used as a policy instrument by the Federal government as a means to relieve repayment burdens.

I also contribute to a growing literature on how technological advances in credit-scoring can generate both efficiencies in consumer lending, and impact market structure (Einav, Jenkins, Levin, 2012, 2013; Einav, Finkelstein, Levin, 2010; Edelberg 2006; Paravisini, Schoar 2013, Phillipon 2016). The literature that studies uniform and average cost pricing schemes in the presence of heterogeneous risk (Bundorf, Levin, & Mahoney 2012; Einav, Finkelstein, & Cullen 2008; Hurst, Keys, Seru, & Vavra 2015) relates directly to my analysis of the Direct Loan program's break-even rate. These studies have shown in several markets (health insurance, mortgages) that while uniform pricing policies achieve cross-sectional redistribution, they can also distort consumer choices and generate welfare loss. Finally, my paper contributes to a literature that structurally estimates parameters relating to risk aversion and consumption smoothing using micro-data on consumer choices and quasi-experimental variation in prices (Einav & Cohen 2007; Best et. al. 2015; Gruber 2008). As noted earlier, my estimates are novel because I can explicitly control for borrower liquidity, and they come from a previously un-modeled setting: maturity choice.

The rest of the paper proceeds as follows. Section II describes the setting and data with an emphasis on the variables impacting borrowers' maturity choices, and the use of exogenous interest rate variation to identify maturity elasticities. Section III describes the theoretical framework, estimation of the maturity demand model, and discusses the results. Section IV outlines a welfare framework, a simple model of loan costs, and analyzes several counterfactuals. Section V concludes.

1.2 Setting and Data:

Institutional Background:

While this paper focuses on the *repayment* of student loans, it is first necessary to understand their origin. From the borrower's perspective, funding for education can come from several sources - private savings, family contributions, state government, the college they attend, a non-profit or private organization, or the federal government. Loans that come from the federal government to finance post-secondary education is by far the most popular option - over 90 % of the student loan market consist of Federal Direct Loans. This paper concentrates on loans that were originated by the Federal government, but may subsequently be refinanced in the private sector.

There are two key loan-related facts that motivate my paper: growth in the originated volume of Federal loans and heterogeneity in borrower risk. Origination rates in the Direct Loan program, where there is no risk-based underwriting and generous lending limits, have skyrocketed over the past decade. Driven by growth on both the extensive and intensive margins of borrowing, the outstanding volume of student debt has quadrupled in the last 12 years, and the median borrower's holding has grown from \$14,000 to \$19,500 (Looney, Yannelis 2015). This growth has made student debt one of the largest forms of household debt, second only to mortgages, and left many borrowers with sizable monthly payments and large amounts of accumulated interest.

Growing in line with origination rates have been average delinquency rates – the average three-year cohort default rate (CDR)[6], peaked at 14.7% for the 2010 cohort, compared to a rate of 5.2% in 2002. However, these average trends mask important heterogeneity: they are primarily driven by a small group of "non-traditional" borrowers attending for-profit schools who do not compete their degrees (Looney, Yannelis 2015). Default rates amongst graduate students and individuals at 4 year institutions have remained consistently low. These relatively low-risk borrowers make up the majority of the dollars lent by the Direct loan program – graduate students are some of the biggest borrowers, holding 33% of dollars

[6] The percentage of loans in delinquency 3 years after entering repayment

outstanding. In general, debt amount and borrower risk are *negatively* correlated, which means a large portion of the Direct Loan Portfolio will be "overpriced" by a break-even price regime.

Delinquency is associated with significant costs for both the government and the borrower. In the event of delinquency, the Federal government can garnish wages and seize any federal payments to the borrower. Student loans are also not dischargeable in bankruptcy, even when refinanced privately. In 2015, the Department of Education estimated their net recovery rate would be approximately 80% (Dept. of Ed. Loans Overview).

Student Loan Repayment Options:

Repayment of federal debt, and importantly choice of repayment plan, does not occur until a student has finished schooling (either undergraduate or graduate school). In most cases a student can postpone repaying their debt for more than a 6-month grace period after graduation without incurring interest. A student chooses a repayment plan only once repayment begins, *not* upon taking out the loan. They also have the option of changing repayment plans as time progresses.

Federal repayment plans fall into two general categories: fixed payment plans, which adjust the monthly payment level to ensure that the full amount of the original loan will be paid off in a specified number of years, and income-based plans, which scale the monthly payment in proportion to the borrower's income[7]. The absence of a prepayment penalty means that even if a borrower is in a fixed payment plan the effective term of their loan might be much shorter or longer. In an analysis of the Federal loan portfolio, Deborah Lucas and Damien Moore note that "time to repayment varies widely, from less than a year to over thirty years," with "approximately 8 % of originated loans closing in less than five years, and approximately 60% within fifteen years". During my analysis I assume that term choice in the Direct Loan program is "flexible" – this somewhat understates the gains borrowers receive in the private market where there is flexible term choice, and is a generous assumption since over 50% of borrowers in the Direct Loan Portfolio remain in the 10 year fixed maturity plan[8].

A new repayment option that is growing in popularity consists of refinancing Federal debt in the private sector. Refinancing can take place at any point over the life of a loan – immediately when a borrower begins repayment to the Federal government, or in the midst of a repayment schedule. Federal loans, which do not carry a pre-payment penalty, are paid off by the private firm which takes over the servicing and liabilities associated with the loan. It is important to note that student loans that are refinanced in the private market are still not dischargeable in the case of bankruptcy.

A key contributor to the growth of the private refinancing sector has been the development of comprehensive, low-cost risk-based pricing. The majority of refinancing firms are

[7]A description of the federal repayment plan options currently available is provided in the Appendix.

[8]I discuss the welfare implications of limiting term contracts (for example to only a 10 year fixed term) in the Appendix.

online lenders, who digitally link to applicants' financial accounts and credit reports, and use extensive amounts of data to quickly and thoroughly assess their risk. These online applications pull information from a wide array of data sources (from employers to credit card accounts) that would have been unavailable or impossible to process in a traditional lending setting. They reduce the frictions of refinancing both for the lender, who face lower underwriting costs, and for the borrower, who can apply and review competing interest rate quotes in a matter of minutes. By making technology integral to the lending process, the student loan refinancing sector is reflective of a larger entrepreneurial trend often referred to as "fin-tech". It describes a growing number of new, online lenders in a variety of sectors (mortgages, personal loans, etc) that use proprietary data-driven lending algorithms and offer an alternative to traditional lenders.

Interest Rates and Loan Costs:

Another important feature that distinguishes the private and public repayment sectors is cost. When choosing a loan maturity, borrowers trade off between the monthly "cost" of a loan, aka the payment made each month towards principal and interest, and the total interest cost paid over the life of that loan. Total cost and monthly payment are inversely related – a longer maturity loan will have a lower monthly payment, but a higher overall interest cost (as more interest accumulates, at potentially a higher rate). This means that different loan maturities will appeal to different types of borrowers. For example, individuals with lower incomes who are more liquidity constrained may prefer a long maturity with lower monthly payments, despite additional interest costs.

The federal government charges a single interest rate for all loan terms, whereas the private sector charges an increasing rate for longer maturities - this means that the "total" cost differential of a long vs. short term loan will be larger in the private refinancing sector than in the federal sector.

Holding maturity constant, the private sector can also offer either a lower or higher interest rate *depending* on the borrower's expected risk. Low risk borrowers can therefore decrease both the monthly and total cost of their loan by moving to the private sector and refinancing at a lower interest rate. For higher risk individuals who would actually face a higher interest rate under risk-based pricing, the uniform federal interest rate is preferable and refinancing will likely not occur.

In the private refinancing sector, the variables that determine an individual's risk-based interest rate are proprietary and company-specific. Traditionally, the basis of risk-based pricing formulas has been credit score - for example, the interest rate an individual gets on a mortgage is typically a function of FICO and loan size. However, firms are also able to consider variables like employment, income, liabilities, and educational background when estimating a borrower's risk [9]. The risk scores in my dataset are typical of "data-driven"

[9] Legally prohibited risk-based pricing factors under the Equal Credit Opportunity Act are: race, color, religion, national origin, sex, marital status, age, and receipt of income from any public assistance program

underwriting, and while a function of thousands of data points, are primarily driven by free cash flow, degree, income, savings, FICO, and debt.[10]

Federal interest rates follow a "one-size-fits-all" formula that is specified under the Higher Education Act. Specifically, each year they are determined by an index rate (currently the 10 year Treasury note), plus an add-on margin that varies by loan type (see Table 1.1). A report by the GAO notes that while these margins attempt to "break-even" in expectation, they have recently generated a profit for the Direct Loan program. Undergraduates are able to borrow at lower interest rates up to a certain limit ($5,500 to $7,500, depending on their year in school), and then must borrow at higher interest rates. Graduate students face higher federal interest rates regardless of loan amount; the fact that they are some of the largest, lowest risk borrowers makes them a population especially prone to refinancing in the private sector.

Dataset:

I use a proprietary dataset from a student loan refinancing firm that contains extensive information on interest rates, risk score inputs and outputs, and maturity and refinancing decisions. The dataset describes individuals who both decide to refinance in the private sector, *and* on the wider population of those who apply and view an initial interest rate quote. While the first group is very low risk, the second sample allows me to measure the distribution of market-priced risk, and the refinancing propensity, of a more representative sample of Federal borrowers.

The main dataset that I use when measuring maturity elasticities is a repeated cross-section of all new borrowers refinancing with the firm over the period of a year; it links background financial information (debt amount, income, assets, credit score) about borrowers with the menu of interest rates they faced and the ultimate maturity choices they made when refinancing with the firm. When estimating refinancing propensity, I use a similarly structured dataset which includes all applicants to the firm, including those who do not necessarily choose not to refinance. The sections below describe the data and price variation used during estimation.

Descriptive Statistics

Refinancer Population: The population of borrowers who ultimately refinance are high income, high debt, and highly educated. The majority (70%) hold higher than a bachelor degree, are in their early to mid-30s (IQR = [29, 35]), and earn a post-tax median income of $67,500. Given that the majority are graduate students and have attended many years of schooling, it is not surprising that they also hold large amounts of student debt, with the

[10] These variables explain over 85% of the risk score.

median borrower owing over $50,000[11]. The median monthly payment on refinanced debt is $600 per month.

The richness of the data allows for a more thorough description of these borrowers beyond monthly income and debt amount. Table 2.3 shows that 40% are home owners, they spend a median of $1,300 on housing each month, and they have an average FICO score of 780. In terms of assets and liabilities, the median borrower holds $38,000 in assets, $0 in investments (the 75th percentile has $15,000 in investments), owes $89,000 in liabilities, and has a median monthly free cash flow (post tax income minus fixed monthly payments like housing) of $3,100. Borrowers hold a host of degrees and occupations; JDs (lawyers) make up 13% of the sample, MBAs are 17%, MDs (doctors) 5%, pharmacists 6%, and dentists 4%. The majority of borrowers finished school in the last 4 years, with 25% graduating in 2016 and 50% since 2012; however, some are refinancing older loans, with 25% of borrowers having graduated before 2010.

The impressive background of these candidates translates into them obtaining considerably lower interest rates when refinancing. The average previous interest rate on the loans (before refinancing) was 6.7%, which is representative of the range of interest rates charged on Federal Loans over the past decade, and the average refinancer saved 2.21 percentage points when refinancing.

Applicant Population: In addition to refinancers, I also observe a larger, less selected sample of website visitors who see an interest rate quote but do not necessarily proceed with the refinancing process. This sample is more representative of the population of graduate student borrowers who have federal loans. Figure 1.3 compares the debt and income quantiles of my applicant sample to a nationally representative sample of graduate student borrowers. They look very similar. Panel (b) in Figure 1.2 plots the range of quoted 10 yr fixed APRs for this applicant group. It is interesting to note that the average risk-based APR for this sample is 6.5%, which is very close to the uniform graduate rate charged by the Direct loan program – this again suggests that the applicant sample is representative of the distribution of risk underlying the federal portfolio.

Maturity Choices: Borrowers in my setting are asked to choose from a continuum of maturities from 5 to 20 years. This allows them to customize their payments and provides a precise revelation of their repayment preferences.

Given the novelty of the choice set and complexity of interest rate/monthly payment tradeoff, one might wonder if borrowers are making completely informed decisions – for instance, if they only see how maturity impacts monthly payment and are unaware of the impact on interest rate, they may choose much longer maturity then they would have in a full information scenario. I argue there are several aspects of the user interface that borrowers interact with that make this unlikely: for one, borrowers are provided with the

[11] The majority of these individuals hold graduate degrees, making them representative of 33% of the $1 trillion student loan portfolio. Graduate students are an important part of the student borrower population, since despite holding the largest amounts of debt (a median of $46,000 amongst 2014 graduates), they have the lowest rates of default and delinquency.

monthly payment, APR, and total paid for the maturity the choose at many points during the refinancing process. Borrowers use a "slider" to adjust their monthly payment, and are shown how maturity, APR, and total payments change simultaneously – Figure 1.4 shows an example of this user interface. This means they are aware not only of the tradeoffs inherent when choosing any given maturity, but also the rate at which these tradeoffs change when they adjust maturity. In this way the information provided to the borrower is very similar to the information necessary for the first order condition calculations in our model: they see both the change in monthly payment associated with a maturity increase and the change in total interest paid. Borrowers are also asked to contemplate and modify their maturity choice at several points during the refinancing process.

This being said, the distribution of maturity choices still shows some evidence that not all borrowers are optimizing in the sense of our model, and may instead be using other heuristics, like rule-of-thumb accounting, when making maturity choices. For example, there are small spikes in the term distribution at the 10 yr and 15 yr marks, suggesting that some borrowers prefer a standard, rather than customized, maturity[12]. Other borrowers seem to have specific monthly payments in mind when choosing a term: the distribution of chosen monthly payments has distinct spikes at "round" monthly payments, like $500, $1000, or $1500. While these behavioral borrowers are in the minority, it is important to acknowledge that our rational model of financial decision making does not apply to all households – those with specific maturity or monthly payment targets will be less sensitive to interest rate changes, and this would bias elasticity estimates downward.

Price Variation:

An important element of this study is the use of firm-wide price changes to identify consumers' elasticity of demand for maturity. The observed distribution of term choices described in the summary statistics above is informative for understanding how different types of customers select into various repayment contracts. However, they reveal little about how consumers would respond if faced with a *new* set of interest rates. Measuring this response requires variation in interest rates.

There are two main types of interest rate variation in my dataset: risk based and experimental. Using risk based variation to identify maturity elasticities is potentially misleading, since individuals of differing risk types may also differ on unobservable dimensions (like expectations about future income growth or volatility) that will impact their maturity choices. I instead focus on using 10 small shifts in interest rates *within* risk score that were conducted at a firm-wide level, and were thus unrelated to the characteristics of any given borrower. Figure 1.5 provides a graphical explanation of why within risk-score price variation is necessary for identification, and how it can be used.

[12] This may also be somewhat of an "inertia" or "default" effect, since most borrowers are refinancing from settings that offer only traditional term choices like 10 years, and may prefer to stay with their current term.

The experimental price changes were conducted primarily to gather quasi-experimental evidence for the firm on maturity choice and application volume elasticities with respect to interest rates. The price changes occurred over time, not simultaneously for different groups of borrowers, and at a frequency of once to twice a month. This frequency helps alleviate concerns about significant changes in the composition of customers over time (a period of rapid growth), but there are still changes in the observable characteristics of the population over the full set of price changes (see below). While not all price regimes lasted the same amount of time or effected the same number of borrowers, on average they each impacted 1,100 borrowers. Borrowers were not aware of the timing of these price changes, and therefore could not respond by adjusting when they refinanced.

To demonstrate the importance of using only the experimental price changes (and not risk based price variation), I run a series of regressions of maturity choice on observable characteristics and offered interest rates. To quantify the interest rate an individual is offered, I calculate the average fixed rate APR (P_i) over all maturities that individual i with risk type p_i faces. Again, there are only *two* sources of variation in P_i: risk type, and price changes. I first regress[13] maturity choice on P_i (and observables), pooling both sources of variation:

$$T_i = \alpha + \beta P_{it} + X_i' \mu + \epsilon_{it}$$

The results (see column 1 of Table 1.3) show individuals who face higher interest rates as measured by P_i are actually *more* likely to choose a longer term, even conditional on income and loan amount. This seems counterintuitive, since it implies that demand for longer loans is essentially upward sloping, even when controlling for different income and debt levels. However, by omitting risk score from the regression, it also implicitly assumes that all risk types have the same level of demand for maturity. If instead, if higher risk types have a higher demand for maturity due perhaps to expected income variability, then our price coefficient would suffer from omitted variable bias.

My next specification includes risk score, thus controlling for risk-based price differences and using *only* the remaining exogenous price variation to identify the price coefficient. The coefficient on the price variable now has a significant, negative sign – this means that when faced with higher interest rates, similar individuals choose shorter loans. My measure of borrower risk score has a strong, negative coefficient, meaning less risky individuals are much less likely to choose a long term maturity.[14]

The elasticity that corresponds with this coefficient (see Table 1.4) says that a one percent increase in the average offered APR causes a .8% decrease in the average maturity chosen. In unit terms, this means that increasing the average APR from 5.5% to 5.6% would decrease the average term by 1.7 months (from a mean of 108 months). For the average $70,000 loan, this would increase monthly payments by 1.6%, but keep total interest payments relatively

[13]I use a tobit specification to account for the truncation of the choice set at 60 and 240 months.

[14]A higher score means an individual is less risky and faces a lower level of risk based prices.

constant, increasing by only .005%. This shows that the majority of borrowers place more weight on minimizing total interest payments than on minimizing monthly payments.

In a final specification, I allow risk type and price to interact – this allows for different price elasticities across risk types and lets us unpack the aggregate elasticity of -.8% estimated in the previous specification. The significant negative interaction term shows that not only do riskier types have a higher willingness to pay for maturity, but their choice of maturity is less price sensitive then their low-risk counterparts'. Table 1.4 expresses these results in terms of elasticities – these show that the highest risk individuals are essentially inelastic to price changes, whereas the lowest risk individuals exhibit a much higher elasticity and reduce term when interest rates increase. This is interesting, since it suggests that while the lowest risk individuals are interest rate sensitive, the highest risk individuals are primarily driven by the level of monthly payment. This difference in price elasticities may at least partially explain the difference in the levels of term choices across risk types.

Figure 1.6 offers a more graphical representation of these elasticities: it plots the average residualized variation in interest rates over rate maps that does not come from risk based pricing[15] and the average residualized variation in term choices that is not explained by risk type or other observable characteristics like debt amount. These residuals are normalized relative to the average APR and average term over my full sample. Figure 1.6 shows that while the experimental price changes are small (from 1-4% of their starting value), they are strongly correlated maturity choices – the average term choice decreased with an increase in the level of prices, while controlling for other relevant variables like loan amount, income, and risk score[16].

One identification concern is that if these changes occurred over time there may be simultaneous changes in the composition of the applicant population. If these compositional changes impact maturity choices (e.g. if some regimes have more high debt individuals choosing longer loans), they could confound our estimates of the impact of APR. Figure 1.23 in the Appendix shows how three variables relevant to term choice, log income, log debt, and FICO, changed within the borrower pool over price regimes – while the differences are not huge, they are statistically significant across some regimes. Therefore in all empirical specifications, I control for changes in observable characteristics so that I don't attribute effects stemming from compositional changes to the price changes. On this point, it is helpful to note that the price changes were not monotonic: interest rates both increased and decreased over time, and therefore will not be confounded by other monotonic trends occurring over time like growth of the company.

Using temporal variation presents a second selection concern: while some individuals may respond to interest rate changes on the intensive margin by adjusting maturity, others may respond on the extensive margin by no longer refinancing or refinancing with a different

[15] I first regress the average APR over all terms on risk score, which predicts most of the variation with an $R^2 = .97$. I use the residual from this regression, which contains variation coming only from temporal price changes, in this plot.

[16] In the appendix I also show how the share choosing a variable rate moved in close tandem with the fixed variable spread.

company. If those who join or leave the population after a price change have systematically higher or lower maturity preferences, then this extensive margin response will bias our intensive margin estimates. My model, which considers both the extensive and intensive margin decisions, provides a framework for thinking about this effect - it suggests that the optimal term choice in any given market is a function of observable and unobservable borrower characteristics, and the market choice is then a discrete comparison of these optimal utilities. Conditional on an individual's "optimal" term, the market with the lowest prices at that term will be chosen. This means a level shift up in the interest rate schedule could make an individual with an 80 month loan preference as likely to leave the market as an individual with a preference for a 180 month loan. In other words, even if we see changes in the size of the pool of borrowers in response to a price change, this should not bias our intensive margin results so long as extensive margin attrition comes from a range of individuals with various term preferences.

One empirical way to gauge the extent of extensive margin responses is to test whether changes in observable borrower characteristics over price regimes are correlated with the experimental variation in APR. If the composition of observable characteristics is predicted by the price changes, then we would be worried that there may also be selection on unobservables. Table 1.6 tests whether four important observable characteristics, income, debt, FICO, and savings, are predicted by the price regime shifts. These insignificant results show that price changes did not cause any differential attrition across observable characteristics: while characteristics like income and FICO did vary over price regimes, this variation was not correlated with the price level.

I also predict individuals' maturity choices, $\hat{T}_i$, using all observable characteristics other than APR, and test whether this variable is predicted by the price regime shifts. Again, these results are insignificant. Figure 1.23 graphically shows the lack of correlation between observable characteristics like debt and income and the price shifts – the fact that the distribution of observable characteristics stayed relatively constant over price regimes means that, for example, low income individuals were as likely to leave the market in response to a price increase as high income individuals. This makes sense since the highly competitive nature of the refinancing market, and growth of risk based pricing, means that even "high" risk borrowers have outside options that are close to their quoted price.

1.3 Structural Repayment Model with Income Heterogeneity

The reduced form estimates showed that individuals who were low income, high debt, and high risk preferred longer loans, without imposing structural assumptions on how or why these variables should impact maturity choice. They also used exogenous price changes to identify the relationship between maturity choice and interest rates, showing that on average maturity choices were sensitive to changes in the level of interest rates, and that this sensitivity decreased with risk. In this section I use the same relationships and exogenous price variation to estimate a structural model which dictates how observed differences in income, debt, and risk based prices should impact maturity choices under the assumption of CRRA utility. This builds upon the reduced form findings by relating borrowers' observed maturity choices, income and expenditures directly to the underlying parameters governing expected utility.

I first outline a simplified model of maturity and refinancing demand to understand what borrower characteristics influence maturity and refinancing choices. This is a two-stage model – borrowers choose a loan term, and then between repaying in the private or public sector. I model this decision process in the reverse order: I first solve for the borrower's optimal maturity choice in both the public and private sectors, and then compare utilities across the two sectors to determine their refinancing choice. The model incorporates the two key differences on the supply side between the private and public sector - the private sector charges term and risk-specific interest rates, whereas the public sector charges a single break-even interest rate to all.

I model all repayment decisions conditional on debt, schooling, and educational choices, which are made at an earlier period before repayment begins. This equates to the assumption that these decisions are fixed and not impacted by the level of interest rates or ability to refinance debt. This assumption is valid for the population of student borrowers who have already made their loan principal decisions and are yet to make repayment choices (i.e. those currently in school or beginning repayment) – however, it does not apply to individuals who have yet to make borrowing decisions (prospective borrowers who have yet to start school). In a full equilibrium analysis, the level of interest rates and refinancing options could also impact choices like loan principal.

Basic Set-up

Borrowers entering repayment have incurred a fixed amount of student debt while attending school, D_i, are now finished with school, and are beginning repayment. Income follows a unit root process: each period log income grows at a constant rate off of the previous period's level, and also experiences a per-period shock. Specifically:

$$ln(w_{it}) = ln(w_{it-1}) + u_{it}$$
$$u_{it} \sim N(g_i, \sigma_i^2)$$

where g_i is a yearly growth rate and σ_i^2 is individual-specific income variance.

All borrowers have the same per-period CRRA utility function $u(c) = \frac{c^{(1-\gamma)}}{(1-\gamma)}$, and discount factor, β. Upon entering repayment, they choose a repayment maturity, T_i, to maximize their present discounted stream of expected future utility. In addition to choosing a maturity, borrowers can choose between a public and private repayment sector. Two main things distinguish the public and private repayment options:

- Risk based pricing: the private sector offers interest rates that are increasing in a borrower's observed expected cost. This expected cost is represented by a borrowers' risk score, p_i, which is a function of observable characteristics: $p_i = f(X_i)$. The government offers a single price for all risk types, g.
- Maturity based pricing: the private company offers maturity specific interest rates, $r(T, p_i)$, that are increasing in T. The government does not vary interest rates over maturity.

Monthly and total payments could be lower or higher for a given individual in the private vs. public sector - this depends on their risk type, maturity preference, and the resulting private interest rate.

Maturity Demand

When choosing a maturity T, individuals maximize the discounted stream of yearly utility over the next 20 years (the length of the longest contract)[17]. Specifically, they solve:

$$\max_T E[\sum_1^T \beta^t u(w_{it} - d_i) + \sum_{T+1}^{20} \beta^t u(w_{it})]$$

$$s.t.\ d_i = D_i * \frac{r(T, p_i)}{(1 - (1 + r(T, p_i))^{-T})}$$

where d_i is the yearly payment associated with maturity T. As borrowers extend maturity, each periods' payments become lower, but they pay more over the life of the loan.

Solving the maximization problem results in the first order condition:

$$0 = -E[\sum_0^T \beta^t \frac{\partial d}{\partial T} u'(w_{it} - d_i) + \beta^{T+1} u(w_{i(T+1)} - \frac{d_i}{2}) - \beta^{T+1} u(w_{i(T+1)})]$$

which can be rewritten as:

$$E[\sum_0^T \beta^t \frac{\partial d}{\partial T} u'(w_{it} - d_i)] = E[\beta^{T+1}(-d_i) u'(w_{i(T+1)} - \frac{d_i}{2})]$$

This condition says that at the optimal loan maturity, the sum of marginal utility gained from a slightly lower monthly payment (from a slightly longer term) is equal to the marginal utility lost from paying additional interest for an extra year.[18]

[17] In the public sector the interest rate $r(T, p_i)$ is replaced with g

[18] To make this condition empirically tractable, I approximate the second term, $\beta^{T+1} u(w_{i(T+1)} - d_{jit}) - \beta^{T+1} u(w_{i(T+1)})$, with the expression:

$$\beta^{T+1} u(w_{i(T+1)} - d_i) - \beta^{T+1} u(w_{i(T+1)}) \approx \beta^{T+1}(-d_i) u'(w_{i(T+1)} - \frac{d_i}{2})$$

The Influence of Interest Rate on Maturity Choice:

My analysis is primarily concerned with understanding how the level of interest rates impacts borrower welfare and repayment choices. The first order condition captures how maturity choices, and therefore utility levels, change under various price regimes. All else constant, as the level of interest rates increases, individuals must decrease maturity to maintain the optimality condition. The exact formula for the response $\frac{dT}{dr}$ does not have an analytical solution, but Figure 1.7 shows how simulated optimal term choices vary over interest rate levels – as the level of interest rates increases, the optimal maturity choice decreases.

The Influence of Non-Interest Rate Factors on Maturity Choice:

The other non-interest rate factors in our model that influence maturity demand and interact with the interest rate elasticity are: income level, debt level, income growth and volatility, and the intertemporal elasticity of substitution $\frac{1}{\gamma}$. Due to concave utility, individuals who are low income or high debt gain more marginal utility from decreasing yearly payments, and thus have a higher willingness to pay for long maturities. Individuals who expect income to grow in the future will also prefer a long maturity, since it acts as a means to transfer consumption from the future to the present. Individuals with higher income variability have both higher and less elastic maturity demand due to the fact that longer loans help to smooth consumption across a more variable income profile.

Note that the income-related factors that drive demand (income levels, growth, and volatility) are very correlated with risk score p_i. Therefore, the same variables that increase demand for maturity on the borrower's side will also increase interest rates on the supply side. This means that even when faced with higher risk-based prices, high risk borrowers may choose longer loans. This is in line with our reduced form evidence, which showed that, all else constant, riskier borrowers had higher demand for long maturities.

The optimal term condition also helps us understand how the intertemporal elasticity of substitution ($\frac{1}{\gamma}$) influences demand. As γ increases, an individual will prefer a longer maturity holding price constant. Intuitively, this is because an individual with concave utility will prefer to smooth consumption by lowering yearly payments, even if it means paying more interest overall. A high level of γ (ie a low intertemporal elasticity of substitution) also means that the term choices of individuals will be less responsive to price changes. Thus γ is essential for understanding how a borrower's decisions, and utility, would respond to *changes* in price, the central goal of this study.

Refinancing Choice

Borrowers also decide whether or not to refinance by comparing the *absolute* levels of interest rates at the optimal maturity across sectors. For individuals on the high or low end of the risk distribution, the risk based price differential determines whether they should refinance. For very low risk individuals there will be a clear incentive to refinance in the private sector, and for very high risk individuals, there will be no incentive to refinance because government interest rates will be significantly lower than private rates.

For marginal individuals who face a similar level of prices in the private and public sectors, the term based price differential (and whether they prefer shorter or longer loans) could also determine whether they will sort into the private sector. If low risk individuals also prefer shorter maturities, they will gain substantial consumer surplus in the private sector where there are lower risk-based *and* lower maturity-based prices. The combination of risk *and* maturity-based price differences could exacerbate risk-based sorting between markets.

While the model describes the decision to refinance as a discrete choice problem, in reality borrowers might face frictions (inertia, search costs) or have idiosyncratic preferences that prevent them from refinancing even when they would receive lower interest rates. In the counterfactual section I therefore estimate and use empirical refinancing elasticities with respect to price to capture more realistic refinancing rates.

Modeling Borrower Delinquency

The above model does not explicitly model how delinquency might impact borrowers' optimal maturity choice. Rather, borrower income levels and risk impact maturity decisions because of the possibility that a low income draw minus a large debt payment will generate a very high marginal utility. Explicitly modeling borrower delinquency would generate very similar predictions – in the Appendix, I derive the exact conditions such that allowing for delinquency will not change optimal term choice.

One reason why allowing for delinquency does not generate vastly different predictions is that it does not benefit the borrower and in many ways extends a loan's maturity. Borrowers are not able to default on their debt – student loans are not dischargeable in bankruptcy even when they are refinanced in the private sector, which removes the possibility of "strategic default". However, borrowers do sometimes become delinquent on their loans, which means they are late on their payments. These delinquent payments generate costs for the lender, who may eventually have to transfer delinquent loans to a collections agency in the private sector, or exert effort to recover payments in the public sector. The current recovery rate in the public sector, where the government is able to garnish wages and seize federal payments, is 80%. Delinquency also has negative consequences for the borrower, who receives a penalty on their credit score and must repay the missed portion in the future with additional interest.

Contemporaneous Financial Decisions

The above model defines yearly consumption as income minus the student debt payment; in reality individuals may be faced with many other fixed expenses and monthly payments that could impact their effective level of liquidity and thus their maturity choices.

One way I address this empirically is by using a measure of borrowers' "free-cash-flow" (FCF) rather than monthly income when estimating the demand model empirically. Free cash flow, defined as the remaining income an individual has after paying taxes and other fixed monthly expenses, accounts for the fact that housing payments and/or other debt payments will substantially lower some borrowers' effective monthly free cash flow and could influence their maturity choices. This is an important empirical adjustment: the median monthly free cash flow ($3,100) is less than half the median monthly income in my dataset. Over 40% of borrowers have a mortgage (which on average translates into a $1,900 payment), and the median monthly fixed expenses for borrowers is $2,400. All of the borrowers have some sort of fixed monthly payment on their credit reports: 40% of borrowers have monthly auto payments which are on average $450, 75% have credit card payments, and 90% have uncategorized installment debt.

While FCF is a more accurate depiction of monthly borrower liquidity, the model also assumes that borrowers are not readjusting on other financial margins when refinancing. In other words, contemporaneous savings and debt decisions are assumed to be exogenous, predetermined, and unaffected by maturity and refinancing decisions. I can test this assumption by looking at borrowers' other monthly payments before and after refinancing, and measuring whether they adjust immediately during refinancing. Table 1.26 in the appendix describes changes in other monthly payments (mortgages, auto loans, credit cards, etc) before vs. after refinancing for individuals who had positive monthly payments to begin with, and shows that for the vast majority of borrowers these stayed constant. This makes sense, since many of these payments are fixed installments, and it would take active work on the borrower's part to readjust them.

I can also observe the savings and investment behavior of borrowers in my sample: because individuals in my sample are young, they have relatively low levels of savings to begin with. Slightly under 40% have a formal retirement savings account – for example 25% have a 401k, with a median balance of $24,000. The number of individuals with investment holdings increases with age. Figure 1.27 shows that while the median borrower continues to not have substantial savings through age 60, the 75th percentile has accumulated over $80,000 by age 50. However, 90% of my borrowers are under 40 years old, and therefore even the most active savers have investment holdings that are much smaller than their student debt amount.

Evidence on Permanence of Term Choice:

Our model assumes that borrowers make a term choice in year 1 to maximize expected utility over the life of the loan. One might question whether borrowers are actually optimizing over such a long time horizon, or if they are in fact choosing a monthly payment to fit their *current* income level, with the intent to refinance and change term yet again in the future when their income level changes.

To address this, I look at payment patterns over time within my sample of refinancers – in other words, do any individuals keep their payment level over time constant, or do they systematically make higher or lower payments on their debt. I find that there are some extra payments in the data, but they are small and do not vary systematically over time. Figure 1.24 in the Appendix shows that each month borrowers pay on average 1.5% more than their regular payment, and this is driven by on average only 1% of borrowers making a extra payment each month. There is also no systematic trend in the extra payments. One might expect payments to increase with time as income increases, but here the level of extra payments stays constant over the two year period.

Estimation

Our reduced form evidence expressed maturity choice as a linear function of observables, risk type, and interest rate (X_i, p_i, and $r(T,p)_{it}$), identifying the interest rate coefficient off price shifts orthogonal to p_i. In the structural estimation, T_i is instead expressed as an implicit non-linear function of X_i, p_i, and $r(T,p)_{it}$ derived from the borrower's first order condition, and the price shifts now serve to identify γ. This structural exercise will allow us to estimate borrower utility using the same price variation and maturity response as in the reduced form exercise. While it imposes stronger assumptions on the borrower's problem (i.e. parametrizing the income process and assuming CRRA utility), it allows us to map the reduced form maturity elasticity to a parameter of economic interest, γ, and to ultimately measure changes in consumer surplus. It also allows economic theory to dictate how borrower liquidity and income risk should impact repayment decisions.

This model assumes that unobservable differences in future income create heterogeneity in maturity choices. This allows individuals who were observationally equivalent *today* (i.e. the same income, debt amount, and risk type) to choose different terms if they have different expectations about future earnings growth or volatility.

The structural estimation faces the same endogeneity concerns as the reduced form estimation – risk-based interest rate variation suffers from omitted variable bias, since the factors that impact risk score could also impact maturity choices. In order to isolate only exogenous price variation when estimating γ, I allow the slope and volatility of income growth to vary across observable characteristics X_i and risk type p_i. I assume that any remaining unobserved heterogeneity that influences maturity choices is uncorrelated with prices, and therefore will not bias my results. The model also implicitly assumes that individuals choose T to conform to future income and consumption paths, but do not choose these paths to conform to T.

Empirical Framework:

Recall that individuals choose T to maximize a discounted stream of yearly utility, which lead to the first order condition:

$$\sum_{1}^{T_i} \beta^t \frac{\partial d}{\partial T} E[(w_{it} - d_i)^{-\gamma}] = E[\beta^{T_i+1}(-d_i)(w_{iT+1})^{-\gamma}]$$

$$s.t.\ d_i = T_i * D_i * \frac{r(T_i,p_i)}{(1-(1+r(T_i,p_i))^{-T_i})}$$

This FOC provides our main estimating moment. We observe most elements of this equation, including: T_i, the optimal term choice, d_i which represents the yearly payment for individual i at term T_i, $r(T_i, p_i)$ which is the risk, term specific interest rate faced by individual i at term T_i, $\frac{dd}{dT_i}$, and w_{i0} which is defined as after tax income. We do not observe future income, w_{it}, but I assume log income follows a unit root process[19] and grows at a yearly rate potentially specific to a vector of observable individual characteristics including: highest degree type , current disposable income, student loan amount, age, age^2, FICO score, home ownership, and number of dependents. Specifically:

$$ln(w_{it}) = ln(w_{it-1}) + (X_i'\mu) + u_{it}$$

where $X_i'\mu$ is a yearly growth rate specific to observable characteristics and

$$u_{it} \sim N(0, \sigma_u^2)$$
$$\sigma_u^2 = (\omega - \upsilon * p_i)^2$$

By modeling income growth as a function of X_i, our specification allows individuals described by any of these characteristics to have higher or lower *levels* of demand for maturity. This will explain why the term choices of individuals with these characteristics vary systematically in the data. Including these characteristics will also control for observable changes in sample across price regimes that might impact term choice, but that we don't want to confound with experimental changes in interest rates.

I model the individual-specific yearly income shock u_{it} as a function of observable risk type. This allows individuals who are similar on all observables X_i, but have different risk scores p_i to choose different terms because of expectations about future income volatility. Again, it is important to include observed risk type in the model because it is perfectly correlated with risk based prices *and* could impact maturity demand. If I didn't include observed p_i in the estimation, the model would wrongly attribute differences in maturity choices across risk type to differences in offered APR, and our estimate of γ would suffer from omitted variable bias.

In my main specification, I use a certainty equivalence approach to write the first order condition as a closed form analytical expression. Specifically I rewrite the expected marginal utility as the marginal utility of a certainty equivalent given by:

$$E[(w_{i0} * e^{t*(X_i'\mu)} * e^{\sum_1^t u_{it}} - d_i)^{-\gamma}] = (w_{i0} * e^{t*(X_i'\mu)} * e^{\pi_{it}} - d_i)^{-\gamma}$$

where π_{it} is the certain amount an individual would have to be given in that period to make the certainty equivalent equal to the expected marginal utility. Specifically[20]:

[19] In the robustness checks I relax this assumption and try other specifications.

[20] for derivation of π_{it} see Appendix

$$\pi_{it} = \frac{1}{2} * t * \sigma^2[1 - (1+\gamma)\frac{w_{i0} * e^{t*(X_i'\mu)}}{w_{i0} * e^{t*(X_i'\mu)} - d_i}] \qquad \text{for t<T+1}$$
$$\pi_{it} = \frac{1}{2} * t * \sigma^2(-\gamma) \qquad \text{for t}\geq\text{T+1}$$

One can see that as income volatility, risk aversion, and the debt to income ratio ($w_{i0} * e^{t*(X_i'\mu)} - d_i$) ratio increases, the certainty equivalent becomes more negative.

Using this expression, our analytical estimating moment becomes:

$$g_i(\theta) = \sum_1^T \beta^t \frac{\partial d}{\partial T}(w_{i0} * e^{t*(X_i'\mu)} * e^{\pi_{it}} - d_i)^{-\gamma} - \beta^{T+1}(-d_i)(w_{i0} * e^{(T+1)*(X_i'\mu)} * e^{\pi_{i(T+1)}})^{-\gamma}$$

This makes the first order condition a nonlinear function of observable variables, $(r, T_i, w_{i0}, D_i, X_i, p_i)$, and unobservable parameters, $\theta = \{\gamma, \mu, v, \omega\}$, that we need to estimate[21]. To estimate the model, I use nonlinear least squares, choosing the parameters $\theta = \{\gamma, \mu, \omega, v\}$ that minimize the quadratic form:

$$b = \arg\min_\theta g_i(\theta)' g_i(\theta)$$

Identification:

Ideally, to identify a parameter like γ we would observe the maturity choices of identical borrowers under multiple price regimes. Instead, we observe maturity choices and observable characteristics, X_i, of similar borrowers under multiple, independently varied price regimes. Our first identification concern is to separately identify the impact of interest rates on maturity decisions from the impact of risk type on maturity decisions, since risk type and risk based interest rates are perfectly correlated. A second concern is to correctly attribute what portion of change in term choice across price regimes comes from actual changes in interest rate, and what portion comes from changes in sample composition.

Both (μ and v) are identified off of how the static *level* of term choice varies across these observable characteristics: we can identify the coefficients in μ off the fact that we observe individuals with different characteristics (X_i, X_i') but similar risk type p_i choosing different maturities when faced with the same interest rate r. The parameter v, which scales income volatility with respect to risk type, is identified off the maturity choices of individuals who face the same prices r and are similar in characteristics X_i, but are different risk types (p_i vs p_i'). Both μ and v help to control for how *observable* heterogeneity amongst borrowers, that are possible changing over price regimes, could impact maturity choice.

In contrast, γ, which represents how consumers trade off consumption now vs. the future, is identified off of *shifts* in the maturity distribution over price regimes, and not

[21] I calibrate $\beta = .98$

level differences in maturity choices. Because our model controls for risk type, which is perfectly correlated with risk based prices, the only remaining price variation comes from the temporal, within-risk type, price experiments. These price changes provide moment conditions in which observationally identical individuals (in terms of both X_i and p_i) face different price regimes (r vs r') and make potentially different maturity decisions. They allow us to compare the term choices of two populations who face different prices, but who we assume are in expectation are similar in their unobservables characteristics. Using this price variation requires the assumption that conditional on X_i and p_i, any unobserved characteristics of borrowers across price regimes are uncorrelated with their maturity choices.

Results

The results from the structural estimation are shown in Table 1.10. The first column estimates come from our preferred specification, which models log income as a unit root process with a growth rate specific to a host of observable characteristics, and the remaining columns report results from specifications with alternative assumptions, discussed in more detail below.

The estimate of γ in the primary, and in all specifications, falls on the moderate to low end of estimates in the existing literature. It translates into an IES of .89, whereas most recent micro estimates have found a IES from .2-.6 (the highest estimate in the literature is an IES equal to 2, or $\gamma = .5$). This value implies that on average there is a sizable consumption response to changes in interest rates. The small estimate of γ is not surprising in light of our sample and setting. It is reflected in both the distribution of term choices, in which over one quarter of borrowers choose the shortest 5 year term, and the reduced form results, which found a relatively large maturity elasticity. In addition, my dataset is unique in that I can fully observe household's balance sheets and explicitly control not only for borrower income, but also for other monthly fixed expenses. Using pre-tax income, rather than this more accurate measure of monthly free cash flow, would overstate borrower liquidity and bias estimates of γ upwards.

Individuals actively refinancing their loans are also likely more cognizant of the interest rate tradeoffs they are making then individuals in studies that examine credit card use or saving rates. The online interface in my setting explains how interest rates and debt maturity interact, potentially making my sample more informed than those studied in a traditional loan setting. For example, when borrowers compare maturities, the website calculates and displays the monthly payment, total interest paid, and APR associated with each. These are complex calculations that the borrower may not make independently, and make the total interest/monthly payment tradeoff studied in our model extremely salient to the borrower.

It is also possible that the *type* of debt studied here could also have a unique psychological impact on estimates of γ. The amount of student debt a borrower has is more often determined by "necessity" (due to the level of tuition or financial aid available at the individual's school), rather than choice and thus be perceived as more burdensome and unwanted. Therefore borrowers might treat their student loans differently than other forms of debt or savings, and want to pay it off more quickly. This results underscore the importance of considering several models of consumer behavior when analyzing saving and borrowing decisions - while our lifecycle model of repayment rationalizes these choices under the assumption of full information and rational expectations, there may in fact be behavioral tendencies, for example debt aversion or rule of thumb accounting, that are driving some portion of individuals' behavior.

In addition to γ, the model also estimates income growth and volatility parameters – they imply that the median borrower expects income to grow yearly by $4,500, which translates into a median 5.2% yearly growth rate. Figure 1.8 shows the range of these estimates. The income growth coefficients in the specification can be understood using the following logic: when income grows at a faster rate, future consumption becomes higher relative to current consumption, and the marginal utility "tradeoff" of paying a loan today versus tomorrow also becomes steeper. Thus rationalizing a shorter term choice (all else constant) requires estimating a lower rate of income growth. The model's CRRA assumption tries to reconcile term choices primarily using current income, debt levels, and γ, with any remaining heterogeneity in term choice explained by variation in these growth rates. A reduced form regression of term choice on X_i (holding p_i and r constant) would reveal that older individuals, those with lower FICO scores, and those with non-graduate degrees all choose significantly longer loans; this model reconciles those same features of the data by assigning those same groups higher rates of income growth[22].As a robustness check, I later compare these coefficients to actual observed changes in income growth over time using a separate cross-sectional sample of similar borrowers.

The coefficients on the age and age squared variables suggest that there is a "hump" shaped age-earnings profile, with estimated earnings increasing and then decreasing with age. This corresponds with the fact that we observe term choices increasing and then decreasing with age, all else equal, and echoes the pattern actually observed in most empirical work on the age-earnings profile. Degree type also has a significant impact on term choice[23], acting above and beyond the differences it generates in current income and debt levels. This makes sense if careers have consistently different earning trajectories; for example, doctors have a more delayed increase in earnings than lawyers, which is reflected in our estimates by them having a higher future income growth rate.

While the magnitude and direction of most of these coefficients seems reasonable, some do appear counterintuitive. For example, do individuals with a Masters degree really expect faster income growth then those with a JD degree? First, it is important to note that these are differences in the income growth rate, not in the level of income. Second, while in the general population these trends may not hold, our sample is "selected" in the sense that BA degree recipients who are approved to refinance may have exceptional income expectations relative to the remaining BA degree population. Finally, these coefficients might suggest there are limitations to our model, which attempts to rationalize term choices under the

[22]These coefficients essentially restate the mean T conditional on X_i in terms of income growth.
[23]Omitted degree type is BA/BS.

assumption that all heterogeneity comes from differences in future income expectations. It is possible that there is also heterogeneity in the discount factor, or bias in future income expectations, that is misspecified by our model and makes these estimates unreasonable. It is interesting to note that individuals with higher starting incomes and lower starting debt amounts are estimated to have higher levels of future income growth. This could mean that the estimate of γ, which is currently identified off of *changes* in term choices in response to price shifts, doesn't perfectly rationalize the *levels* of term choices in the data – high income, low debt individuals choose terms that are too long relative to their low income, high debt counterparts, and the model therefore rationalizes these choices with higher levels of income growth.

In addition to future cash flow growth coefficients, we also estimate v, which controls how income variance varies across risk type. These imply that the certainty equivalent is decreasing in risk type – after 10 years, the certainty equivalent is 97.5% of expected income for the best risk score and 92.5% of expected income for the lowest risk score.

Figures 1.19 in the Appendix analyze the fit of the model. They compare predicted to observed term choices, and show that in general the model slightly overpredicts maturity, but otherwise seems to perform well. All counterfactual exercises use these predicted term choices as a comparison point.

Robustness Analyses:

In this section I test the robustness of my modeling assumptions in several ways: first, I draw on external pieces of evidence that lend context and credibility to the structural model. I compare the cash flow paths implied by the estimated parameters to actual cross sectional age-earnings paths from a dataset of similar borrowers. I also use these external income paths to calibrate the model, and use the remaining variation to identify heterogeneity in γ. Second, I estimate alternative specifications that directly relax key structural assumptions of the econometric model. Finally, I use an additional borrowing choice made by borrowers, between a fixed or variable interest rate, to provide a second estimate of γ and further evidence of interest rate sensitivity.

Evidence on Cross-Sectional Income Paths: One of the key assumptions of our model was that expectations about future income growth influence and generate heterogeneity in maturity choices that can't be explained by observable income, debt, and interest rate levels. In this section I test that assumption with a second dataset, to see whether cash flow does in fact seem to change with age, and whether the shapes of these income paths vary across different types of borrowers in the ways predicted by our model.

While one would ideally use panel data on the income, asset, and liability paths of my borrowers for this exercise, the long time horizon of the debt contracts (up to 20 years) is a limiting factor. I instead use a cross section of observationally similar individuals at various ages to create a pseudo age-income profile. The dataset I use to estimate these profiles contains individuals who are similar to my refinancing applicants in many important respects (high income, high FICO, mainly graduate degree recipients), but who are applying instead for small personal loans rather than applying to refinance student debt. This distinction is important when estimating cross-sectional age profiles – if I used a cross-section of the student loan borrower population to estimate these profiles, one might worry that individuals refinancing student debt at age 40 have very different income trajectories than those refinancing at age 30. Here the worry is that individuals borrowing small amounts ($5,000 - $15,000) at different ages have fundamentally different earning trajectories. While this selection concern is valid, one must weigh it against the fact that this population is similar to my borrowers in many unique respects that would be difficult to find and match to in a survey dataset like the CPS. These include both tangible characteristics, like degree type, income level, or FICO score, as well as intangible characteristics. For example, my population is refinancing with a new internet-based bank, which makes them potentially different, or more tech savvy, then a population that uses only traditional banks. Furthermore, because my sample has a high socioeconomic status, they make up only a small percentage of most representative survey samples.

Figure 1.9 plots the median yearly income for individuals in this personal loan sample of different ages, controlling for state and degree type - this plot shows that cash flow *does* change with age, increasing considerably between the ages of 25-35. The primary model estimates had a median growth rate of $4,500, which is not all that different from the yearly income growth rate seen in the external data. In our primary specification I also modeled this growth rate as a linear function of age and age^2, which again seems to be a reasonable assumption given that income increases at a rather constant rate from 25 to 35, but then somewhat levels off. The graph plots these estimated trends at age 25 and age 30, and shows that the growth rates estimated by our model are in fact very close linear approximations to the observed income growth rates in the data.

Our primary specification allowed income paths to vary by many characteristics, from number of dependents to FICO score. While we would like to compare trends observed in this dataset to those estimated in our model, we can use cross-sectional data to only compare on *time-invariant* characteristics like degree type, occupation, student debt amount, and state. Figure 1.25 in the Appendix plots monthly free cash flow (FCF) after separating individuals into 4 degree levels: associates, bachelors, masters, and professional. These types not only have different levels of free cash flow, but also somewhat different trends – professional degree recipients have more rapid FCF growth between 20-40. Even amongst professional degrees the paths can vary considerably – the figure below compares the free cash flow growth of the three largest occupational groups represented in the data, doctors, lawyers and pharmacists. It shows that doctors experience lower incomes and more rapid growth through their mid-thirties, perhaps due to the long residency process. Our structural model provided a similar estimate – it showed that MDs had much higher predicted FCF growth than JDs.

These plots show that our empirical specification was correct in several assumptions: income seems to grow with age (contrary to a constant income assumption), and these growth rates vary considerably with observable characteristics like degree and occupation. The direction of many of our estimated coefficients is reinforced by this observed dataset: for instance, income growth increases and then decreases with age, and doctors experience more delayed income growth then other professional degrees. However, the growth rates estimated in our model were slightly larger then those observed in the data – this could either mean that there is some misspecification of our model, or it could mean that using a age-earnings cross-section of borrowers understates earnings growth. This could occur if older borrowers applying for loans have lower incomes then their young borrower counterparts will have in several years.

Estimation using Cross-Sectional Income Paths: While my model estimates implied income paths from term choices, another option would be to calibrate income paths and variance using an external data source. The model would then return estimates for γ, but no longer estimate v and μ. This approach imposes more assumptions (for example that these external cross-sectional income paths are actually representative of individual specific income trends in my data), but offers another means for estimating γ and verifying the robustness of our initial estimate.

I use the external dataset described above to calibrate my model, estimating income growth rates that are specific to age and occupation. I calibrate income variance by looking at the cross-sectional variance of income across observed FICO categories.[24] On average, this means that median income grows by 6.6% for a 25 year old and 5.6% for a 30 year old. I then use these growth rates to predict future income paths for individuals in my estimation sample. I still use the observed starting income levels of individuals in my sample, and calibrate only the growth rates.

This exercise estimates $\gamma = 1.2245$ with a 95% CI of (1.2099 , 1.2391). This is close, but slightly higher, than the estimates from our main specification. One reason why this may yield a slightly higher estimate of γ is that a cross-sectional income path may be less steep than an actual *within*-individual path. This could occur if the older individuals in my sample of personal loan applicants are selected on having lower savings and potentially lower incomes. Their incomes would therefore be lower then the future incomes of the 20 or 30 year old in my sample. All else constant, a less steep projected income path requires a higher level of γ to rationalize the same maturity choices, since longer maturities now play less of a role in redistributing and smoothing consumption from the future to the present.

Estimation of Heterogenous γ: Calibrating, rather than estimating, these income paths also allows us to instead use the variation in term choices to estimate heterogeneity in γ (rather than heterogeneity in implied income growth). When I estimate γ as a function of observables (see Table 1.5), including risk score, log income, and log debt, I find that individuals who are riskier or have a higher debt to income ratio have a higher level of γ (or a lower IES). This supports the idea that individuals who have less free cash flow will be less responsive to the level of interest rates and more focused on using their maturity choice to maintain a certain monthly payment. The resulting distribution of estimated values of γ range from 1 to 1.9.

Alternative Specifications: Our primary specification models log income growth as a autoregressive process. Table 1.10 presents results from several alternative specifications which test

[24] I do not observe the same risk score variable in this second dataset, and therefore use FICO instead.

the sensitivity of my primary estimates to the underlying assumptions and structure of the model.

In specifications 2 and 3, I allow income to be normally, rather than lognormally distributed. This means that income follows the process:

$$w_{it} = X_i'\mu + w_{it-1} + u_{it}$$

and therefore the estimated income growth variables are expressed in dollar terms, rather than as growth rates. This change in functional form does not change the predictions of the model – when expressed in dollar terms, the estimated income growth rates from our main specification are very close to those estimated using this specification. The two specifications also return very similar estimates of γ. The details of this estimated method (including the derivation of the analytical moment) are included in the appendix.

In specification 3, I make the assumption that income is constant over time, conditional on a vector of observable characteristics that I include outside the utility function. These controls are included to account for the fact that certain observable characteristics of the sample, like income or debt level, changed over price regimes. This makes our first order condition estimating moment:

$$(T_i * \frac{(w_{0i} - d_i)}{(w_{0i})}^{-\gamma} - \frac{-d_i}{\frac{\partial d}{\partial T}_i} + X_i'\mu) = 0$$

The results from this specification give a slightly higher estimate of γ. The coefficients on the controls are no longer interpretable as "income growth", but they are similar in sign to those from our reduced form specifications.

Finally, specification 4 uses simulated method of moments for estimation, rather than nonlinear least squares. More details describing this approach can be found in the appendix. I again find a similar estimate of γ, which suggests that the analytical approximations underlying the primary specification do not have a huge impact on estimates.

Evidence on Fixed/Variable Rate Choice: To provide a second estimate of the level of risk aversion (captured by the parameter γ), I analyze the same borrowers' choice of a fixed or variable interest rate. At the same time as choosing a term, borrowers also choose between a fixed and variable rate, but have the flexibility to change this decision once a year. Like term choice, this decision again captures the preferences of borrowers trading off between interest rate savings and levels of consumption – however, this is a choice which smooths consumption over states of uncertainty, rather over time. By locking in the current prevailing interest rate, the fixed rate provides "insurance" to the borrower against future volatility in interest rates. It is therefore priced at a premium to the variable rate, which is pegged to the prevailing market rate and thus can change over time. A more risk averse individual would be willing to pay a higher fixed rate premium to insure against this uncertainty, just as in the term choice scenario they would be willing to pay a higher maturity premium to smooth payments over time.

In the appendix, I write down and estimate a model that describes the rate choices made by borrowers. In the spirit of the term choice estimation, I again use small exogenous changes in the fixed variable spread that were *not* based on or correlated with market-wide interest rate changes to measure how these prices changes impact what proportion of borrowers choose a fixed rate loan and offer a second estimate of γ. I estimate γ as both a constant parameter and a function of observables using maximum likelihood estimation; the results are shown in Table **??**. The average value of γ estimated using the fixed variable choice is .85, similar to that found when examining term choice: again, borrowers seem very price sensitive and fast to switch between the fixed and variable rates as prices change. Borrowers who are lower risk and have more free cash flow (higher income and lower debt) have a higher estimated value of γ, perhaps due to the fact that they are less liquidity constrained, and thus less likely to choose a variable rate for interest rate savings.

1.4 Welfare Analysis:

In this section, I use the estimates from the structural demand model to analyze how borrower welfare will change as the private refinancing sector expands and develops more comprehensive risk-based pricing. My benchmark for these comparisons is an entirely uniform, break-even regime, which represents how the Direct Loan program would operate with complete pooling and no private refinancing sector.

When I compare the benchmark to a scenario in which the public and private options coexist, I show two main effects: the first is a net increase in consumer surplus, as low risk types refinance into lower, more efficient risk-based prices. The second is an equity loss: as low risk types select into the private market it will increase average costs for the Direct Loan program, and increase the break-even interest rate for the remaining borrowers. I show that as the ability of the private sector to price on more observables increases, the impact on these two effects is not necessarily equal – if pricing innovations primarily lower rates for individuals who would have already refinanced, efficiency gains will not come at an equity cost.

I next analyze how both of these factors, growth and distribution of consumer surplus, change when 1) we account for frictions that reduce the propensity to refinance, and 2) under an alternate government pricing policy. I show that if the government transitions from a break-even interest rate to a net subsidy it can stem unraveling and strike an equity/efficiency tradeoff. I also show how the size of a interest rate subsidy must account for the two behavioral, and budget-relevant, responses highlighted in our model: maturity choice and refinancing.

For these exercises, I use the sample of all refinancing applicants – individuals who received a refinancing price quote, but who did not necessarily complete the entire refinancing process. This is different from my estimation sample, which included only approved, agreed refinancers. The applicant sample is much more representative of the federal loan portfolio - it includes individuals with expected market interest rates both above and below the current graduate interest rate (see Figure 1.2). However, this exercise requires that I extrapolate estimates taken from the refinancing sample to a group with a much wider distribution of income, FICO score, and debt amount. To limit the extent of the extrapolation, I restrict the applicant sample to individuals who have a debt-to-income ratio that overlaps with the support of the refinancing sample.

Welfare Framework:

A welfare comparison of these scenarios requires developing a money metric measure of changes in consumer utility under different price regimes as well as a simple model of loan costs in the private and public sector, which I outline below.

Quantifying Changes in Consumer Surplus:

To quantify and calculate the change in consumer surplus under the various counterfactual pricing scenarios, we need to translate our ordinal measures of utility into monetary values. First, note that two interest rate regimes, (r, r'), will generate two different term choices, (T, T'), and therefore two different monthly payments (d, d'). One can think of quantifying the utility difference between these two price regimes as the sum given to an individual today that would make their level of utility under the second price regime the same as it was under the first price regime. Specifically, the CV is given by:

$$u(w_1 - d) + E[\sum_{t=2}^{T} \beta^t u(w_t - d) + \sum_{t=T+1}^{T_{max}} \beta^t u(w_t)] =$$

$$u(w_1 - d' + CV) + E[\sum_{t=2}^{T'} \beta^t u(w_t - d') + \sum_{t=T'+1}^{T_{max}} \beta^t u(w_t)]$$

I express the CV both in absolute dollar terms, and also as a percent of the each borrower's total interest payments.[25,26]

[25] An alternative measure of CV, but not one that I use in this paper, would be to calculate the amount one would need to give the borrower *each period* they make payments under the new price regime to make them as well off as they were before the price change:

$$E[\sum_{t=1}^{T'} \beta^t u(w_t - d(r', T') + CV) + \sum_{t=T'+1}^{T_{max}} \beta^t u(w_t)] = E[U(r, T)]$$

[26] I use a certainty equivalent approach when calculating expected utility, where:

$$E[u(w_t - d)] \approx u(w_{i0} * e^{t*(X_i'\mu)} e^{\pi_{it}} - d)$$

where now

Private Loan Pricing and Break-even Pricing

The analysis of the demand side focused on understanding how borrowers' preferences vary across *risk type*. The emphasis on this dimension stems from the fact that risk type also determines costs on the supply side, since risk represents an increased probability that part of the loan will not be recovered. Therefore, as our counterfactual models how different risk types choose maturities and sort across sectors, it requires us to model how the supply side will set and adjust prices in response to individuals' maturity choices and sorting patterns.

Private Sector Pricing: I use the observed interest rates schedules from my dataset to represent the risk and maturity specific prices charged in the private sector in our model. From the perspective of the model, the heterogeneity I observe in these risk based prices comes from differences in income levels, growth, and variability across borrowers that generates differential costs for lenders. In reality, the firm estimates borrower risk using a predictive algorithm to estimate the probability of delinquency. This algorithm produces a risk score p_i for each individual based on a vector of characteristics, X_i. This score maps to a schedule of risk, maturity specific interest rates $r(T, p_i)$.[27]

In order to use these observed prices as inputs into the government's break-even pricing rule, I must assume firms in the private sector set each maturity, risk-specific interest rate s.t.:

$$r(T, p_i) = \alpha + E[c(T, p_i)]$$

where α represents costs that are invariant to term or risk, such as the cost of capital, overhead, or servicing, and $E[c(T, p_i)]$ is the expected cost of lending to individual i over maturity T.

My assumption that observed interest rates reflect the expected costs of lending hinges on the assumption of a perfectly competitive refinancing market - i.e. if a firm charged above cost to lend to an individual, another firm could enter the market and offer a slightly lower price to the same individual while still breaking even. The refinancing market displays most features of perfect competition, including rapid entry into the industry by many firms, and little product differentiation. It is also very easy for consumers to price shop and compare

---

$$\pi_{it} = \frac{1}{2} * t * \sigma^2 [1 - \gamma \frac{w_{i0} * e^{t*(X_i'\mu)}}{w_{i0} * e^{t*(X_i'\mu)} - d_i}] \qquad \text{for t<T+1}$$

$$\pi_{it} = \frac{1}{2} * t * \sigma^2 (1 - \gamma) \qquad \text{for t}\geq\text{T+1}$$

[27] In addition to credit scores, the non-linear scoring formula also incorporates additional factors such as education, employment history, income, debt to income, free cash flow, and both the stock and flow of other financial accounts (e.g. credit cards, savings, investments, loans) to estimate a risk score (p_i) for every applicant that reflects their the probability of delinquency. These variables are similar to the variables (income level, volatility, and liquidity) driving costs in our model.

across refinancing firms, due to their online nature and the fact that they all offer quick, personalized price "quotes".

By using these market prices, I focus only on risk that is observable and priced in the private sector but unpriced in the public sector. I also assume that expected losses given risk type and term are the same in the private and public sector. This assumption seems reasonable given that both sector treat default and delinquency similarly: student loans are not dischargeable in bankruptcy, even when they have been privately refinanced.[28] One concern when using private sector costs as the basis for g is that the fixed costs of lending, α, in the private market are different then those in the public sector. In the private sector α represents costs that are invariant to term or risk, such as the cost of capital, overhead, or servicing, and that may not apply to the federal Direct loan program.

Uniform Break-even Pricing: I assume the government sets its interest rate g to be revenue neutral: at g the sum of the per-borrower subsidies will equal to zero. I use risk-adjusted discount rates to estimate the size of each per-borrower subsidy – specifically, I discount future cash flows under the uniform price regime with the risk, maturity-specific interest rates that would be assigned to that loan in the private sector. These risk-adjusted discount rates will be higher than g for high risk and long maturity borrowers, generating a positive subsidy, and lower than g for low risk, short term borrowers, generating a negative subsidy. The risk-adjusted stream of cash flows, where monthly payments under uniform pricing are given by $d_i(g)$ and term choice is T, is:

$$PRDV_i(g) = \frac{d_i(g)}{r(T,p_i)}[1 - \frac{1}{(1+r(T,p_i))^T}]$$

The value of the subsidy is given by the difference between the risk-adjusted present value of the loan, and the loan principal (which is equivalent to the present value of the loan without risk adjustment). The breakeven interest rate g is thus defined by:

$$\tilde{g} = \{g : \sum_{i=1}^{N} D_i - \sum_{i=1}^{N} PRDV_i(g) = 0\}$$

Baseline: Welfare under Fully Uniform Pricing

As a benchmark for my all my analyses, I assume individuals in my sample are forced into a uniform pricing scheme at the rate that is breakeven (according to the definition above) for my sample. This equates to an interest rate of 6.6%, which is in the range of existing

[28] Private sector risk scores are also highly correlated with another much cited measure of expected costs used by Federal loan program: the cohort default rate (CDR). The CDR is calculated and published by the Federal government at the school level, and reflects the student loan default rate of a cohort of students from that school after 3 years of completion. The difference in the CDR between the highest and lowest risk types in my sample is roughly 2 percentage points, which is similar to the spread in their risk-based interest rates (Figure 1.20). This means any risk-based cost differential would have to stem from differences in the recovery rates across sectors.

Federal Interest Rates for graduate students[29]. This is slightly higher than the average 10 yr APR in the sample, since it takes into account that maturity choices change under uniform pricing and impact total costs.

Our benchmark represents the market at an extreme where there is full pooling and no price differentiation. Figure 1.11 compares the uniform benchmark to the opposite efficient extreme: a fully risk-based setting[30]. The graph characterizes riskier types as having both higher and less elastic demand ($D(p_H)$ vs $D(p_L)$) for maturity, as motivated by our model and confirmed empirically. Efficient risk-based prices ($r(p_H)$ and $r(p_L)$), which I assume coincide with expected costs, are increasing in term and risk. The equilibrium term choices in the efficient setting are denoted by T_L^* and T_H^* – here riskier types choose longer loans even when the cost differential is reflected in the pricing. In the benchmark setting, where the uniform price is represented by g, the high and low risk term choices are pushed further apart.

Uniform pricing will create inefficiency as some individuals choose maturities at a price that is above or below the expected cost of providing to them - the low risk types end up choosing shorter loans than they would in a setting where they are charged the cost of providing the loan, and this distortion is large because of their elastic demand. This graph also highlights the equity motive of the uniform price: it acts as a means to redistribute surplus between the high and low risk types, effectively "taxing" low risk types in order to "subsidize" high risk types.

Using our model, we can calculate the quantities represented here graphically[31]: on average, individuals who are low risk gain \$2298, or 12% of their loan interest payments, in consumer surplus under risk-based pricing relative to uniform pricing, whereas individuals who are high risk (and thus face lower interest rates under uniform pricing) lose an average of \$832 (or 4.2% of their total interest payments) in consumer surplus. We can get a measure of deadweight loss under fully uniform pricing by aggregating the CV across the entire population[32]. The net change in surplus under the two schemes is equal an average of \$1006

[29]Actual interest rates on Direct PLUS Loans ranged from 6.41-7.9% from 2006-2016, while those on Graduate or Professional Direct Unsubsidized Loans ranged from 5.4 to 6.8%

[30]This figure somewhat simplifies the problem by making demand a function of the level of interest rates, rather than a function of both the level and spread, but is helpful for understanding how the two pricing schemes compare. Recall our demand function modeled above, $\{T : E[\sum_0^T \frac{\beta^t}{\beta^{T+1}}(\frac{w_{it}-d(T)}{w_{i(T+1)}-\frac{d(T)}{2}})^{-\gamma}] - \frac{-d(T)}{\frac{\partial d}{\partial T}} = 0\}$, optimizes with respect to both the level of prices at term T represented by $d(T)$ and the slope of prices at term T represented by $\frac{\partial d}{\partial T}$.

[31]To quantify changes in consumer surplus during repayment under uniform vs. perfectly competitive risk-based pricing I first predict borrowers' maturity choices in either setting. Figure 1.21 in the Appendix shows the average predicted maturity choice made by different risk bins under each regime – as shown in the theoretical graph above, lower risk borrowers decrease their terms, while higher risk borrowers increase their terms. These term responses translate into a term elasticity of -.9%, very close to the reduced form elasticity estimated earlier. I next calculate the per borrower compensating variation (CV) that would make each borrower indifferent between uniform pricing and risk based pricing.

[32]The change in producer surplus under the two pricing regimes is equal to 0. This is because the uniform pricing scheme is revenue neutral for the government by design ($\sum_{i=1}^{N(g)} D_i - \sum_{i=1}^{N(g)} PRDV_i(g) = 0$) and we

per borrower, or 5% of interest paid.

While uniform pricing redistributes surplus over risk type, it achieves more modest redistribution over another equity-relevant variable: income. Figure 1.28 in the appendix plots the average subsidy given to each borrower under uniform pricing over both borrower risk type and borrower income. The lowest income borrowers get a subsidy of slightly more than \$1000, while the riskiest borrowers get an average subsidy of almost \$2,000. This is because income is not perfectly correlated with risk type or maturity preferences (the two dimensions that differentiate costs and thus directly generate redistribution), and therefore the uniform rate is an imperfect instrument for achieving redistribution over income.

Counterfactual I: Innovations in Risk-Based Pricing and Expansion of the Private Refinancing Market

Our benchmark scenario compared the market at two extremes: fully uniform or fully risk-based pricing. While this helped illustrate the efficiency-equity tradeoffs under either pricing regime, it is an unrealistic portrayal of the student loan market, where in fact the two schemes coexist and are evolving.

I next analyze what happens to sorting and welfare in the market as risk-based pricing technologies advance and refinancing firms are able to price on more characteristics. Panel (a) of Figure 1.12 shows how innovations in risk-based pricing increase the distribution of interest rates charged in the private sector relative to a more coarse measure of borrower quality like FICO score. Here I calculate and plot the 10 year fixed interest rate each borrower would face if the firm could only price on FICO score, as well as the 10 year fixed interest rate each borrower would face if the firm could price on a more comprehensive set of variables, like monthly free cash flow, assets, degree type, and occupation. The graph shows that more comprehensive risk-based pricing expands the distribution of interest rates, in particular extending lower interest rates to the least risky types. The gains to considering additional characteristics is especially large for the student borrower population: because these borrowers are young and have less-developed FICO scores, this allows them to signal their risk type through other characteristics like degree type, savings behavior, or income.

Using these prices, we can calculate changes in borrower surplus as low risk individuals refinance out of the public sector to take advantage of lower rates. In this initial analysis I assume all individuals who would benefit from refinancing ($CV > 0$) do so. Table 1.7, Column 1, displays the average gain from risk-based pricing for individuals who refinance. By leaving the uniform regime, these consumers gain on average \$2298 in consumer surplus, or 12% of their total interest payments. Using comprehensive risk-based pricing, rather than FICO-based pricing, increases these gains dramatically by \$980 per borrower. It is helpful to frame these changes in surplus in terms of overall interest paid on the loan – Figure 1.13 plots the distribution of compensating variation as a percentage of total interest paid on the

make the assumption of perfect competition in the private sector. Thus the change in total welfare is equal to the change in total consumer surplus borrowers' loan, assuming a 10 year term at an APR of 6.6%, for individuals with a positive CV under full risk-based pricing. The savings are substantial, over 12% for the average borrower under full risk based pricing and 7.5% under FICO based pricing.

From the government's perspective, when these low-risk individuals leave the break-even program to refinance, the average risk of the remaining pool increases. Under the assumption that the government continues to use break-even pricing, I calculate the new break-even rate for the remaining set of borrowers in column 2. The initial exit of low risk types under full risk-based pricing increases the break-even interest rate by .63 percentage points (almost 10%). While innovations in risk-based pricing increased the gains for individuals who refinanced relative to FICO-based pricing, the two pricing schemes generate relatively similar changes in the new break-even rate (.63 vs .61). This is because innovations in risk-based pricing do not increase the extensive *number* of refinancers, but rather extend lower interest rates to individuals who would have already benefited from FICO-based pricing and exited the public pool. Therefore the non-refinancing borrowers face a similar decrease in welfare (column 3) under either FICO or risk-based pricing. This means that total welfare (expressed as the average change in consumer surplus in column 4) increases as risk-based pricing methods become more comprehensive, without necessarily increasing selection out of the break-even program.

Accounting for Refinancing Frictions: Use of Empirical Refinancing Rates

Above I use the rule that eventually all individuals will refinance if $CV > 0$, which implicitly assumes there are no refinancing frictions. Next I use empirical refinancing elasticities that reflect the fact that not everyone who would gain utility from refinancing in fact does so. There are potential switching costs associated with refinancing, and some individuals may value certain aspects of the Federal repayment program more than the interest rate savings they could achieve by refinancing. Figure 1.22 in the Appendix plots the observed probability that an individual in our sample attempts to refinance (i.e. they fully fill out an application after seeing a risk-specific interest rate quote) against their estimated CV, and shows that the observed propensity to refinance increases gradually with the associated financial incentives.

To provide a refinancing probability that reflects these frictions, I estimate empirical refinancing elasticities wrt. APR with a series of regressions that again utilize the firm-conducted price experiments. The dependent variable is full application to refinance after seeing a "quick" price quote, and the dependent variable is the average APR an individual was shown. I consider the process of fully applying equivalent to an intent on the borrowers' part to fully refinance.

Specifically, I use a logistic regression specification, where R_{ij} is an indicator that individual i applied to the firm j after being shown an interest rate quote $r(T,p)_{ij}$; p_i is the individual's risk type; and X_i are individual characteristics. Let :

$$\begin{aligned} U_{ij} &= X_i'\mu - \alpha r(T,p)_{ij} + v * p_i + \epsilon_{ij} \\ &= V_{ij} + \epsilon_{ij} \end{aligned}$$

be the utility that individual i gets from refinancing with firm j, where ϵ_{ij} is a type one extreme value error. An individual will choose firm j if $U_{ij} \geq U_{ik} \forall k \neq j$. Specifically:

$$Pr(R_{ij} = 1) = \frac{exp(V_{ij})}{\sum_{k=1,N} exp(V_{ik})}$$

I next make two simplifying assumption that reduce the problem down to a binomial discrete choice problem. First, note that there may be multiple options for refinancing ($j = 1, 2, ...N$) – for instance multiple private firms or the default option of repaying to the federal government. I make the assumption that, conditional on interest rate, all refinancing firms and the Federal loan program provide the same utility to a borrower. In other words, the only differentiating product characteristic is price.

Second, I isolate and use only experimental price variation to estimate the change in the probability of refinancing when interest rates change by controlling for the borrowers' risk type. This assumes that these price shifts were made independently of price changes at other firms and the Direct Loan program. This assumption seems reasonable, since the experiments were conducted to collect information on borrower elasticities, and not for competitive reasons.

This allows me to assume that $U_{ik} = U_{ik'} \forall k \neq j$ – in other words there is only a "single" outside option. I can then rewrite the problem as:

$$Pr(R_{ij} = 1) = \frac{exp(V_{ij})}{exp(V_{i0}) + exp(V_{ij})}$$

I estimate the model using a logistic regression framework. The resulting estimates (Table 1.8) show that when the average price increases by 1%, borrowers are 2.5% less likely to fully apply. Conditional on the above assumptions, the elasticity captures the change in the probability of an individual repaying to a specific entity when the price of that entity changes, holding all other prices in the market constant. Importantly, I focus on how the application rate *changes* with prices, rather than the level of this rate.

I use these estimated refinancing elasticities to repeat the refinancing exercise described above: I calculate the probability that each individual in my sample would refinance into the private sector when faced with an alternative break-even price regime set at 6.6% APR. I assume the probability of refinancing is 0 if $g = r(T, p_i)$, and that it increases according to the empirical refinancing elasticity as the spread $g - r(T, p_i)$ increases. The results are shown in rows 3-4 of Table 1.7.

With frictions, the extent of selection is reduced significantly – the break-even rate increases by only .12 percentage points. The *size*, rather than just the absolute *sign*, of gains now impacts the decision to refinance. This means only the least risky individuals, who have the most to gain from refinancing, will select into the private sector. But while frictions reduce inequity in surplus across risk type, it also reduces total efficiency. More individuals remain in the uniform price regime, which reduces the total change in average surplus by more than 50% relative to a setting with no frictions.

Because the size of the gains from refinancing now impact borrower decisions, frictions also differentiate the impact of more fine-grained pricing on selection. FICO-based pricing, which offers less savings to low risk borrowers, causes fewer borrowers to refinance and therefore has a very small impact on the break-even interest rate.

Counterfactual II: Transitioning to a Net Subsidy

The previous counterfactual showed that as risk-based pricing advances, low risk types will exit the public sector and the break-even rate will be forced to increase. In theory, this process could continue, and lead to the portfolio eventually unraveling.

One policy response that would prevent unraveling is a transition from break-even pricing to a net subsidy. Table 1.9 contains the per borrower and total subsidy that the government would need to provide if they did not readjust their interest rate after low risk types exited, but rather kept it at 6.6%. Without refinancing frictions, this equates to an average subsidy of slightly over $2,000 per borrower.

This policy is more efficient than fully uniform pricing, since it allows low risk types to reap the gains associated with lower risk-based prices. It is less efficient than fully risk-based pricing, since the government continues to subsidize high risk types. However, this policy achieves greater equity: the high risk borrowers no longer will lose surplus as risk-based pricing advances. In addition, the government now has more control over which funds subsidize the low risk types – rather than implicitly "taxing" low risk borrowers as they would under a break-even scheme, the subsidy could be funded by an income tax and by individuals who are not necessarily borrowers. A subsidy also allows for intergenerational redistribution, whereas a break-even program can only transfer across borrowers in the same cohort.

Our model highlights how, in setting this subsidy, policy makers must consider behavioral responses on two budget relevant margins: maturity choice and refinancing decisions. As the interest rate is lowered, borrowers will on average increase their term choices, which will increase expected costs for the government. This means that the effective size of the subsidy will be larger than the mechanical change. Panel (a) of Figure 1.14, which assumes no extensive margin response, shows that a mechanical 1 pp decrease in the interest rate will actually generate a 1.5 pp increase in expected costs.

However, as the subsidy increases, the average risk of the pool will decrease, since marginally less risky borrowers will no longer refinance. Panel (b) of Figure 1.14 focuses on this extensive margin response, holding term choice constant. This analysis assumes that all individuals with $CV > 0$ will refinance in the private sector. As shown in the previous section, this means that a interest rate that is break-even under full pooling ($g = .066$) will induce low risk individuals to leave and generate a net subsidy from the government. As g is lowered, marginally less risky individuals will remain in the federal portfolio, decreasing the average expected costs. This means that moving the uniform interest rate, for example, from 6.6 to 6% only increases the size of the effective interest rate subsidy by less than .2 percentage points.

1.5 Conclusion

In many consumer credit and insurance markets, from credit card lending to automobile insurance, risk-based pricing algorithms have become widely-used and more fine-grained. On an individual level, as firms become able to identify and accurately price consumer risk, there is the potential for large efficiency gains. Pricing innovations could also have wider implications for market structure. If innovations in risk-based pricing create clear winners and losers, this may change the government's role in ensuring equity and redistribution of surplus.

This paper focuses on the student loan refinancing market, where private firms use rich financial and educational data to underwrite student borrowers. These pricing practices are in contrast with those of the Federal Direct Loan program, which offers borrowers a uniform, break-even interest rate. I use exogenous variation in interest rates and microdata on the market valuation of borrower risk to quantify both the efficiency and equity implications of innovations in risk-based pricing in this setting. My findings highlight how developments in the private sector's ability to price borrower risk will simultaneously i) improve welfare for low risk borrowers by over $2,300 on average and ii) potentially increase sorting of low risk borrowers out the public repayment pool, which under the Direct Loan program's current break-even pricing rule would generate welfare loss for high risk types of $2,100 on average. I also show that the size and distribution of surplus gains change when refinancing frictions are considered, since then only borrowers with the largest gains will select into the private sector.

These findings provide insights into the optimal role government should play in these markets, in particular as a concurrent source of credit or insurance. I show that in a setting where there is no government option, or where the government option is priced at a break-even rate, the changes in surplus from risk-based pricing can be large on average, but heterogenous in sign. This means that innovations in risk-based pricing come at a direct equity cost. I analyze an alternative policy that would address equity concerns: transitioning to a net interest rate subsidy. In other sectors, for example in the mortgage market, the government has responded by providing subsidized insurance or credit to borrowers who would otherwise face high risk-based prices. In the student loan space, several policymakers have already proposed refinancing the Direct Loan portfolio at lower interest rates[33]. Our model quantifies how an interest rate subsidy would impact both borrower welfare and the federal budget – in particular, it shows that low risk, high income borrowers are interest rate sensitive and more likely to decrease term and refinance in response to higher rates. These budget-relevant behavioral responses will change the expected costs of an interest rate subsidy, and thus need to be taken into consideration.

[33]For example, the "RED Act" sponsored by Senators Elizabeth Warren, Patty Murray, Tammy Baldwin, Chuck Schumer in 2016. Hilary Clinton also has proposed a plan to allow borrowers to refinance their student loans at current rates, claiming that it "would help 25 million borrowers across the country, with the typical borrower saving $2,000 over the life of the loan."

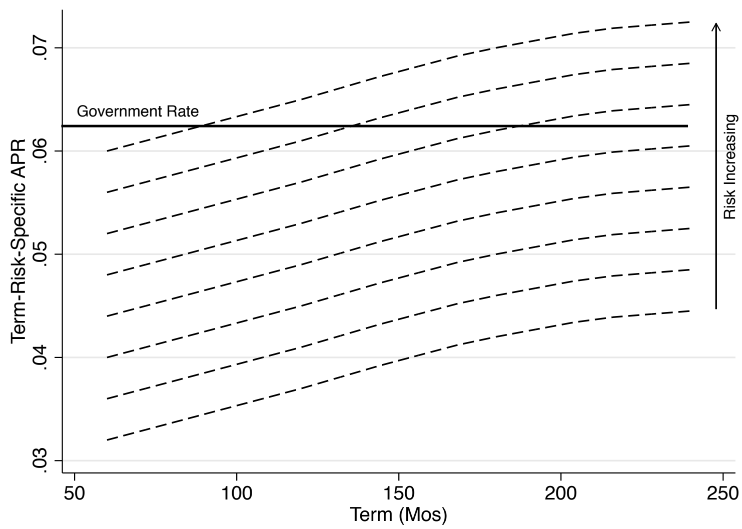
While subsidies are one policy option, in other markets the government has responded by regulating the number of variables that the private sector can price on. Handel, Hendel, and Whinston (2015) examine the welfare effects of regulations in the health insurance market that govern insurers ability to use health status information in pricing; my paper provides a similar framework for thinking about how these types of pricing regulations would impact the student loan market. I show that when firms price can only on FICO score, they reduce the gains to low risk borrowers substantially. However, this restriction did not reduce the extent of selection into the private market, and thus reduced only efficiency, not inequity. A similar exercise could be used to analyze how the amount and distribution of surplus would change if the private sector was restricted to price on, say, school rank or degree type, and would be an interesting area for future study.

Student loans are also unique in that they are non-dischargeable in bankruptcy – this substantially lowers the expected costs, and therefore the interest rates, of student loans relative to other types of uncollateralized lending. Both public and private student loans receive this privileged status, despite facing different pricing constraints. Understanding how the removal, or modification, of the bankruptcy provision for private lenders would impact borrower surplus and market dynamics is a promising area for more research.

In addition to studying these market dynamics, my paper also highlights the importance of maturity choice for student borrowers in repayment – in a setting where loan principal has already been determined when an individual first attended school, the choice of term during repayment is akin to the decision of how much to borrow. I show that allowing for flexible term choice, relative to a traditional setting where borrowers are restricted to fixed terms, improves borrower welfare by allowing them to distribute payments optimally over time given their earnings expectations. This helps both borrowers who prefer to lower monthly payments by extending term, and those who want to lower interest rate payments by decreasing term.

This analysis of pricing developments in the student loan space generates many interesting questions for future work, both at the borrower and market level. My analysis focuses on how interest rates impact repayment decisions, but student borrowers could also respond to interest rate levels at earlier steps in the borrowing process, for example when taking out debt or deciding whether to attend graduate school. Could the availability of better risk-based interest rates change these decisions as well? At the market level, our analysis focuses on risk-based pricing innovations on the primary lending market, with little understanding of how these changes will interact with the secondary asset-backed securities market, where consumer loans are often bundled and sold as securities. If secondary market investors continue to use “traditional” measures of risk like FICO score when comparing and buying bonds, innovations in risk-based pricing could have a limited impact.

Figure 1.1: Federal vs. Private Interest Rates, over Term and Risk



This figure illustrates how federal interest rates compare to private interest rates, which increase both with risk and with term. While all borrowers face the same federal interest rate (solid line), the private sector offers each risk type their own menu of interest rates (dotted lines). This figure shows how individuals with lower risk and lower term preferences have a larger incentive to refinance.

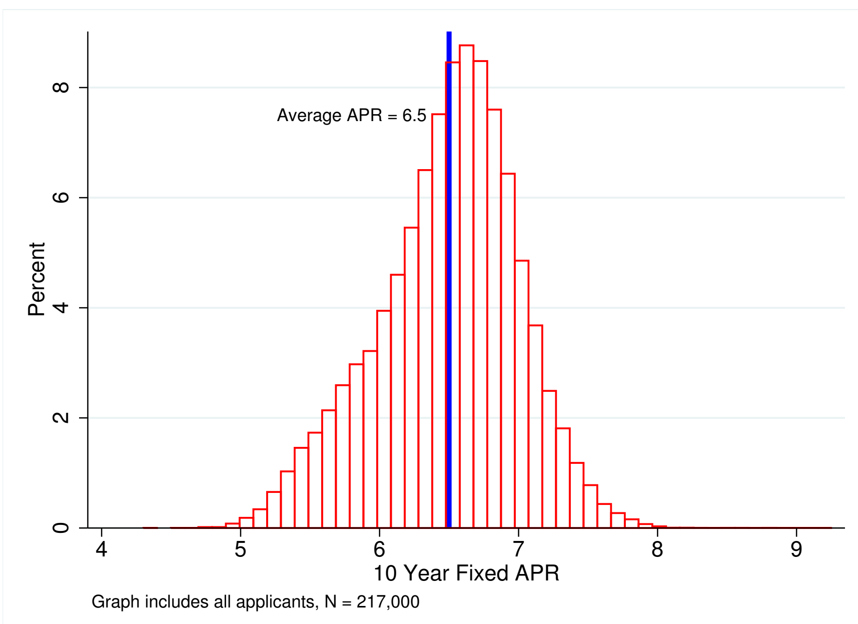


Figure 1.2: Distribution of Potential Refinanced Interest Rates

This figure shows the distribution of potential fixed rate APRs for all applicants, not only individuals who were approved or fully completed the refinancing process. The diversity of market-value interest rates stems both from diversity of risk types and term preferences. The vertical line shows the average APR in this distribution - it is close to the government's "break-even" graduate student rate, which suggests that our sample is at least somewhat representative of the population in the Federal Loan portfolio.

Figure 1.3: Comparison of Applicant Pool to Nationally Representative Sample of Graduate Student Borrowers

This figure compares the student loan amount and income quantiles of my applicant pool to those in a nationally representative sample of graduate student borrowers. The two populations look very similar, which suggests that the individuals I observe are similar to graduate students with Federal Loans.

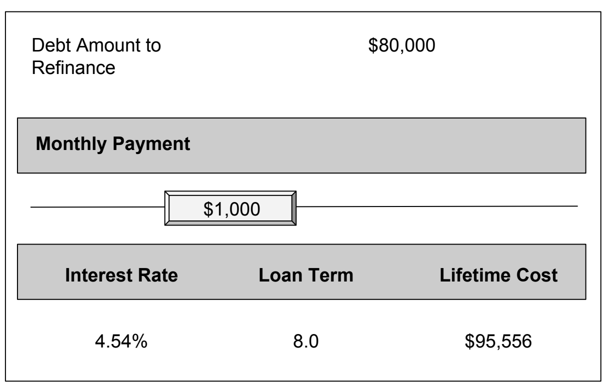
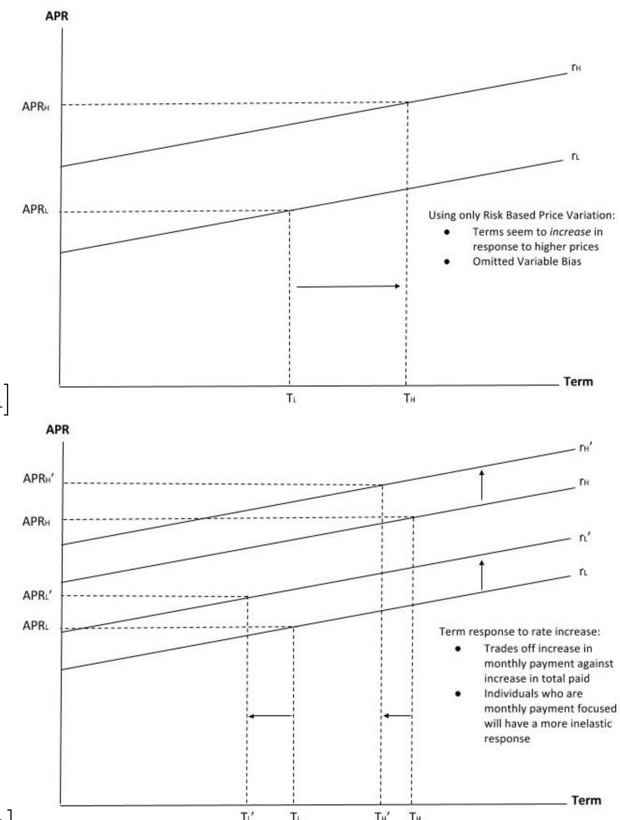


Figure 1.4: User Interface for Term/Monthly Payment Selection

This figure shows the interface individuals use when selecting a loan term. The interface shows them the customized monthly payment, APR, and total paid associated with every possible term choice. This means even less financially-savvy borrowers are well-informed of the implications of their term choices.



Across Risk Price Variation] Within Risk Price Variation]

Figure 1.5: Using Across vs Within Risk Price Variation to Identify Term Elasticities

This figure shows the importance of only using within risk price variation to identify price elasticities. Panel (a) shows the term choices of two different risk types facing risk-based price variation – despite facing higher interest rates, the riskier type chooses a longer loan due to other omitted factors. This makes it seem as though term demand is increasing in interest rates. However, panel (b) shows that when prices increase *within* risk type, term demand actually decreases with interest rate.

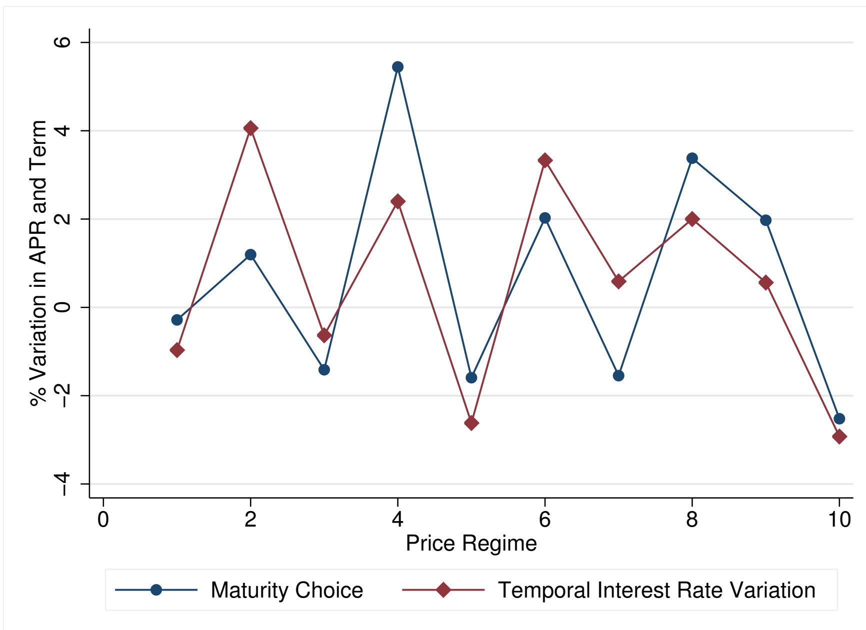


Figure 1.6: Experimental APR Variation and Maturity Responses

This figure shows how interest rates changed over time under various price regimes and the resulting changes in borrower term choices– it plots the average residualized variation in average APR and in borrowers' term over price regimes. To create these residuals, I first regressed average APR on risk score, to remove risk based price variation. I also regressed term choices on risk score and a host of observable characteristics (like debt amount) which could have varied over price regimes and influenced term choice. This plots the residuals from both of those regressions, and shows that term choices are responsive to changes in the level of interest rates.

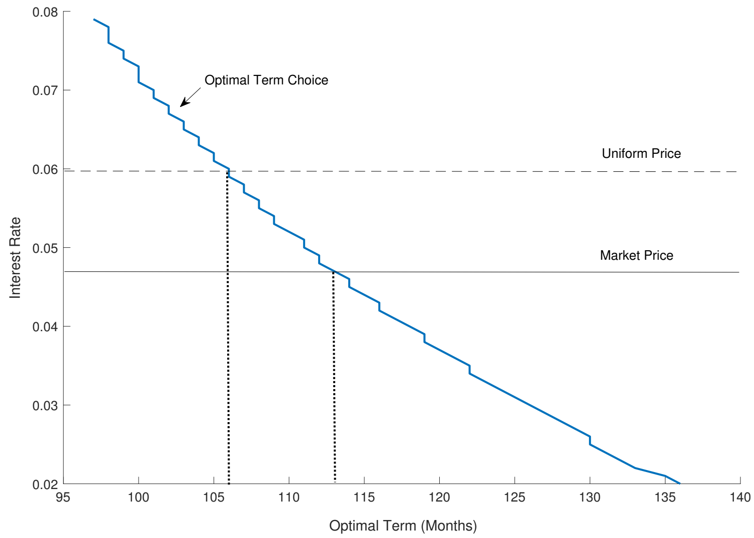
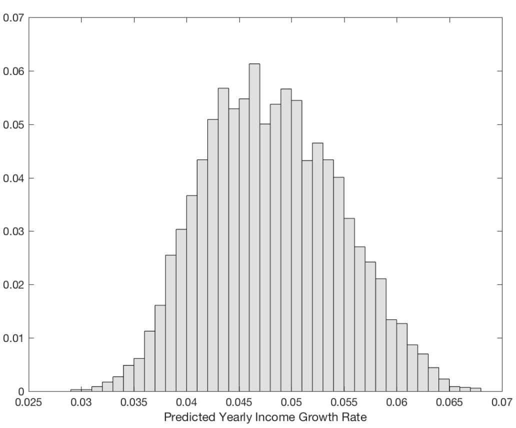


Figure 1.7: Variation in Observable Characteristics over Time

This figure shows how the simulated optimal term choice varies with interest rate levels – as the level of interest rates increases, the optimal maturity choice decreases. This is the interest rate elasticity with respect to savings.



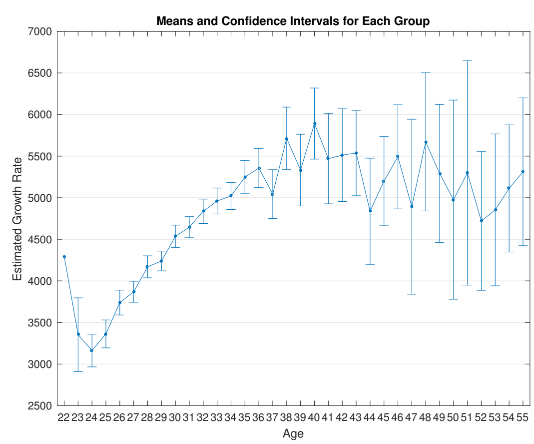


Figure 1.8: Estimated FCF Growth Rates

This figure plots the income growth rates ($X_i'\mu$) estimated with our primary specification. Panel (a) shows the distribution of these estimated growth rates. Panel(b) shows how these income growth rates vary over age.

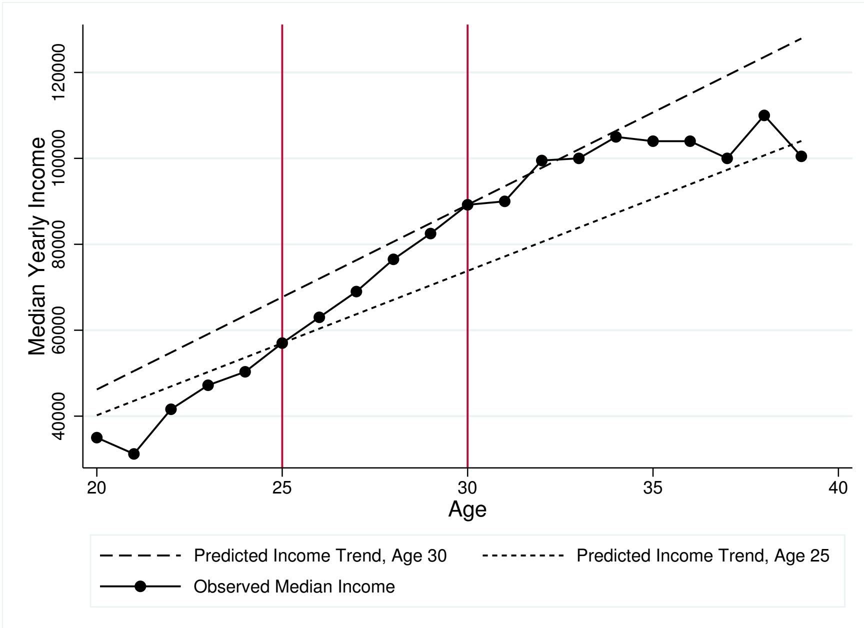


Figure 1.9: Observed vs. Estimated Cross-Sectional Age Earnings Profiles
This figure plots both cross section income trends that are observed, and those that are estimated by our model. The model estimates linear, age-dependent trends. The graph shows how these estimated trends for a 25 and 30 year old compare to actual trends in the data – both the model and the observed data show that income grows more rapidly at age 30 than 25, and the estimated and observed growth rates are quite similar.

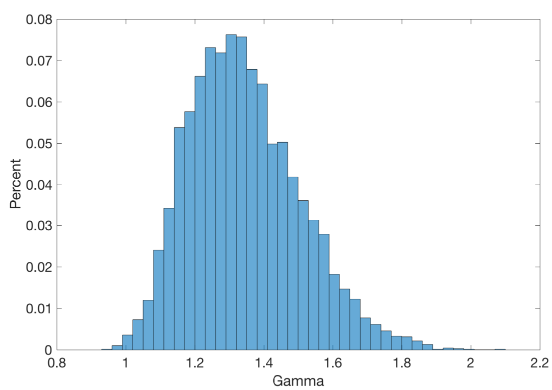


Figure 1.10: Estimated Distribution of γ
This figure shows the distribution of estimated values of γ when income paths are calibrated. Calibration allows us to instead estimate heterogeneity in γ, making it a function of observables like free cash flow and risk type. The impact of these observables on the value of γ are shown in the above table, along with 95% CI.

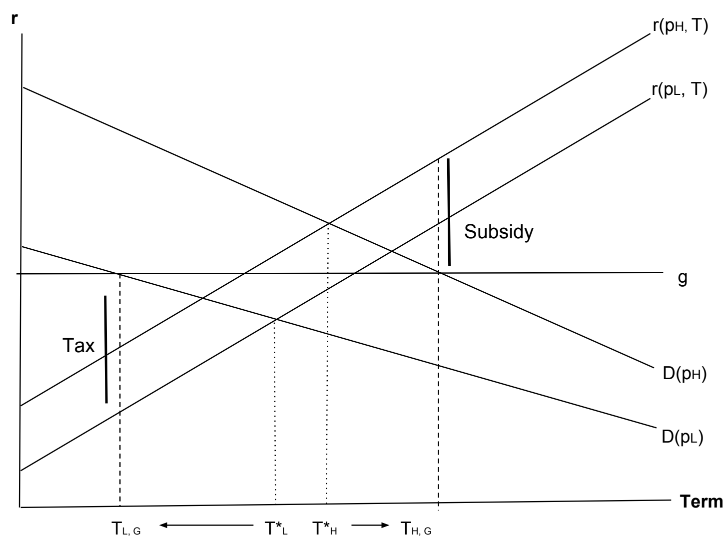


Figure 1.11: Equity and Efficiency Impact of Uniform vs. Risk-based Pricing
This figure depicts the maturity choices a high and low risk borrower would make in the private sector and the public sector. There are two demand curve, the lower one depicting term demand for the low risk borrower, and the higher one depicting term demand for the high risk borrower, which is higher for any given APR. In the private sector they face the two risk specific supply curves, $r(p_H)$ and $r(p_L)$, which demonstrate how APR increases with risk and with term. In the private sector they choose terms T_H^* and T_L^*, which are efficient. In the public sector, they instead both face the same price menu represented by g. In that setting, the low risk type chooses a much shorter term $T_{G,L}$ and the high risk type chooses a much longer term $T_{G,H}$ than in the private sector.

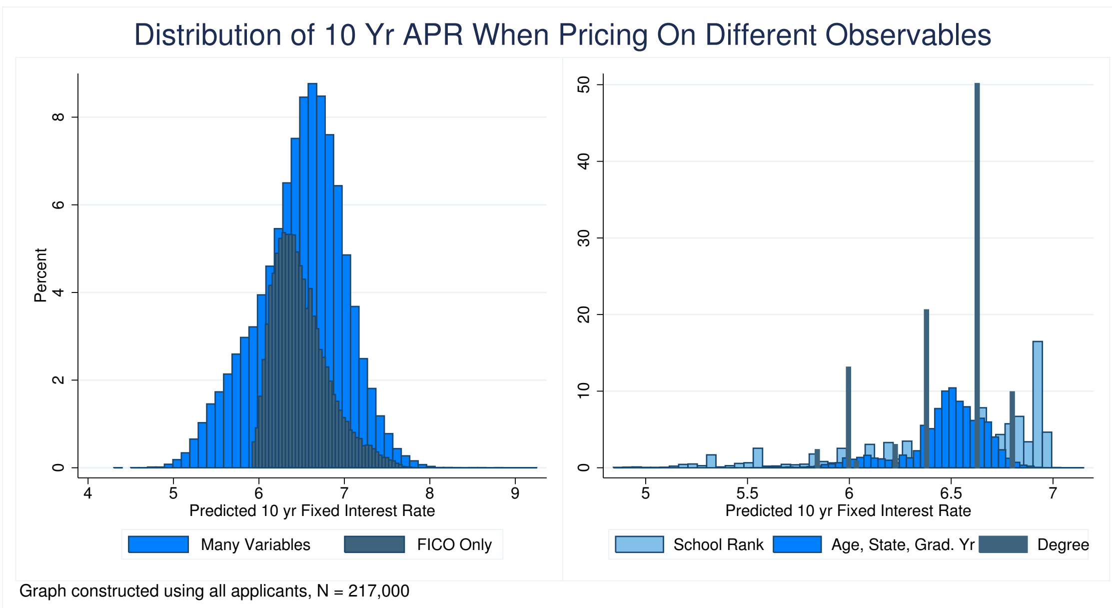


Figure 1.12: Distribution of Interest Rates when Pricing on Different Observables

Panel (a) shows how innovations in risk-based pricing increase the distribution of interest rates charged in the private sector relative to a more coarse measure of borrower quality like FICO score. Here I calculate and plot the 10 year fixed interest rate each borrower would face if the firm could only price on FICO score, as well as the 10 year fixed interest rate each borrower would face if the firm could price on a more comprehensive set of variables including monthly free cash flow, assets, degree type, and occupation. Panel (b) shows how further restricting the set of variables a firm could price on, for example to school rank or degree type, would substantially reduce the spread of the price distribution.

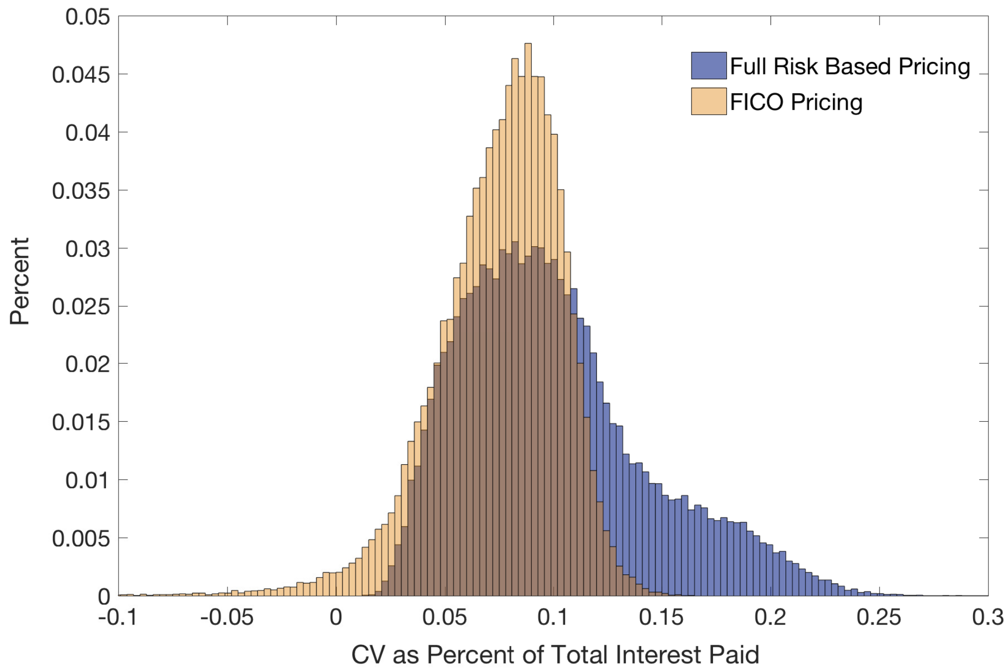


Figure 1.13: Distribution of CV as Percent of Total Interest Paid
This figure plots the distribution of compensating variation as a percentage of total interest paid on the borrowers' loan, assuming a 10 year term at an APR of 6.6%, for individuals with a positive CV under full risk-based pricing. The savings are substantial, over 10% for the average borrower under full risk based pricing and 7.5% under FICO based pricing.

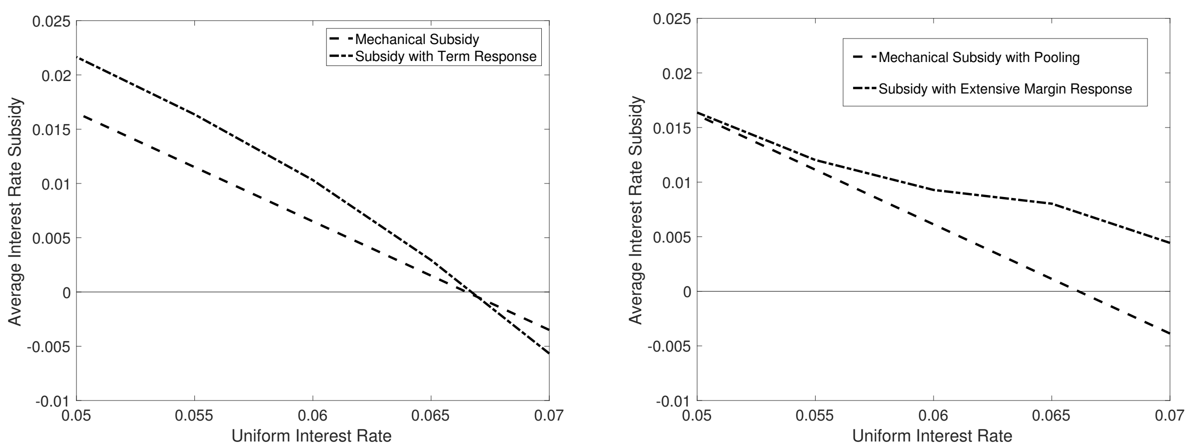


Figure 1.14: Mechanical vs. Effective Interest Rate Subsidies, Accounting for Behavioral Responses
These graphs show how the effective size of an interest rate subsidy, once accounting for the behavioral maturity and refinancing responses, can be large than the mechanical size. Panel (a) illustrates how as the interest rate is lowered, borrowers will increase their term choices, which will increase expected costs for the government. Panel (b) illustrates the extensive margin refinancing effect, showing how as the rate is lowered, marginally less risky individuals will remain in the federal portfolio and decrease the average expected costs.

Table 1.1: 2011-2015 Interest Rates on Federal Direct Student Loans

Loan Type	**Borrower Type**	**Index**	**Add-on**	**Fixed Interest Rate**
Direct Un/Subsidized Loans	Undergraduate	10 Yr Tr	+ 2.05%	3.4-4.66%
Direct Unsubsidized Loans	Graduate/Professional	10 Yr Tr	+ 3.60%	5.41-6.8%
Direct PLUS Loans	Parents& Graduate	10 Yr Tr	+ 4.60%	6.4-7.9%

Source: http://www.ifap.ed.gov/eannouncements/051515DLInterestRates1516.html

Table 1.2: Borrower and Loan Summary Statistics

Borrower Summary Statistics			
	Mean	Median	SD
Income	75,879	68,304	39,799
Loan Amt	67,078	50,656	52,890
FCF	3,636	3,100	2,574
FICO	782	787	36
Mortgages	0.40	0.00	0.60
Graduate	0.70	1.00	0.50
Age	32.60	31.00	6.80
Dependents	0.50	0.00	0.90
Loan Summary Statistics			
Maturity (Months)	106.8	93.0	49.7
Monthly Payment	799.9	600.0	624.8
APR	4.494%	4.64%	1.042%
Variable Rate	0.40	0.00	0.50
N	11663		

Table 1.3: Impact of Debt, Income, APR, and Risk on Term

	(1)	(2)	(3)
ln(Income)	-17.86***	-15.94***	-16.73***
	(2.349)	(2.385)	(2.381)
ln(D)	34.35***	34.05***	34.38***
	(1.083)	(1.083)	(1.084)
Age	1.016***	1.072***	1.039***
	(0.122)	(0.122)	(0.122)
Home Owner	8.294***	8.636***	8.611***
	(1.722)	(1.715)	(1.719)
Variable Rate	2.570	3.162*	2.683*
	(1.632)	(1.626)	(1.627)
Avg. APR	2143.4***	-2355.5***	1217.2
	(198.4)	(847.8)	(916.0)
Risk Score		-15.69**	5.286
		(6.890)	(7.364)
Avg. APR * Risk Score			-306.8***
			(116.1)
Constant	-229.9***	124.6*	-130.9*
	(32.31)	(67.15)	(67.00)
σ			
Constant	58.53***	58.18***	58.36***
	(0.638)	(0.634)	(0.636)
N	11663	11663	11663

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table displays results from a series of Tobit regressions of price and borrower characteristics on term choice (which is truncated at 60 and 240 month). The first specification pools both risk and temporal variation in the Avg. APR variable – because higher risk borrowers (who face higher APRs) prefer longer loans, this regression suffers from omitted variable bias. It seems as though higher APRs drive individuals to increase their term choices. Specification (2) controls directly for risk score and therefore the only remaining variation in avg. APR comes from temporal price changes that were independent of borrower characteristics. Specification (3) allows risk score and price to interact, thereby allowing different risk types to have different price sensitivities.

Table 1.4: Maturity Elasticities

	(1)	(2)
Overall	-0.819***	
	(0.307)	
Highest Risk		-0.226
		(0.352)
Mid Risk		-0.708**
		(0.313)
Lowest Risk		-1.524**
		(0.626)
N	11663	11663

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table converts the coefficients in specifications (2) and (3) from Table 1.3 into elasticities calculated at the mean values of the independent and dependent variables. These values can therefore be interpreted as the percentage change in term in response to a 1% increasing in average APR. The "Overall" elasticity calculates the elasticity for the entire sample, whereas the second column separately calculates the elasticity for the upper, middle, and lowest thirds of the risk distribution.

Table 1.5: Estimated Distribution of γ

Estimated γ Coefficients	
Constant γ	0.59
	(0.225 , 0.95)
log(FCF)	-0.194
	(-0.212 , -0.175)
log(Amt)	0.255
	(0.219 , 0.292)
Home Owner	0.066
	(0.042 , 0.089)
Risk Score	-0.041
	(-0.053 , -0.027)
Dependent	0.087
	(0.073 , 0.102)
N	10075

Table 1.6: Test of Extensive Margin Response and Changes in Borrower Composition

Observables over Price Regimes			
	Avg. APR		
	Coeff.	SE	t
ln(Income)	-.0000367	.0000362	-1.01
ln(Debt)	-9.55e-06	.0000161	-0.59
ln(Savings)	.0000162	.0000168	0.96
Mortgage	1.52e-06	.0000221	0.07
Age	1.24e-06	1.98e-06	0.63
F(5, 11663) = .99			
Response of $\hat{T}$ to APR*			
	$\hat{T}$		
	Coeff.	SE	t
Avg. APR	363.09	341.98	1.06
N	11663		

$\hat{T}$ is the term predicted using all observables *except* price.

Table 1.7: Impact of Refinancing on Borrower Surplus and Break-even Interest Rate

	Without Refinancing Frictions			
	Avg. Refinancer Δ CS	Δg	Avg. Non-refinancer Δ CS	Avg. Δ CS
FICO-based Pricing	1315	0.61	-2040	64
Full Risk-based Pricing	2298	0.63	-2109	480
	With Refinancing Frictions			
	Avg.Refinancer Δ CS	Δg	Avg. Non-refinancer Δ CS	Avg. Δ CS
FICO-based Pricing	1896	0.06	-209	56
Full Risk-based Pricing	3040	0.12	-411	218

Table 1.8: Extensive Margin Elasticities

	Temporal Variation Only
Average APR	-2.363***
	(0.443)
N	76494

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.9: Subsidy Needed to Stop Unraveling

	Without Refinancing Frictions	
	Per. Borrower $ Subsidy	Total Subsidy as % Total Loan
FICO-based Pricing	2027	1.41%
Full Risk-based Pricing	2086	1.61%
	With Refinancing Frictions	
	Per. Borrower $ Subsidy	Total Subsidy as % Total Loan
FICO-based Pricing	266	.44%
Full Risk-based Pricing	499	.76%

1.A Description of Federal Repayment Plans:

Figure 1.15: Federal Loan Repayment Plans

	Eligible Loan Types	Term	Monthly Payment
Standard Fixed Repayment	• Direct Subsidized and Unsubsidized Loans • Subsidized and Unsubsidized Federal Stafford Loans • All PLUS loans	• Up to 10 Years	• Fixed amount, at least $50 per month
Graduated Repayment	• Direct Subsidized and Unsubsidized Loans • Subsidized and Unsubsidized Federal Stafford Loans • All PLUS loans	• Up to 10 Years	• Payments are lower at first and then increase every two years
Extended Repayment	• Direct Subsidized and Unsubsidized Loans • Subsidized and Unsubsidized Federal Stafford Loans • All PLUS loans	• Up to 25 Years	• Payments may be fixed or graduated
Income Based Repayment	• Direct Subsidized and Unsubsidized Loans • Subsidized and Unsubsidized Federal Stafford Loans • All PLUS loans made to students • *Consolidation* Loans (Direct or FFEL) that do not include Direct or FFEL PLUS loans made to parents	• Up to 25 Years • If loan not repaid in full after making the equivalent of 25 years of qualifying monthly payments, any outstanding balance will be forgiven.	• Maximum payment is 15% of discretionary income o Discretionary income is the difference between your AGI and 150% of the poverty guideline • Payments change as income changes • Individual needs to demonstrate partial financial hardship to qualify
Pay as You Earn Repayment	• Direct Subsidized and Unsubsidized Loans • Direct PLUS loans made to students • Direct Consolidation Loans that do not include (Direct or FFEL) PLUS loans made to parents • Must be a new borrower on or after Oct. 1, 2007, and must have received a *disbursement* of a Direct Loan on or after Oct. 1, 2011.	• Up to 20 Years • If loan not repaid in full after making the equivalent of 20 years of qualifying monthly payments, any outstanding balance will be forgiven.	• Maximum payment is 10% of discretionary income • Payments change as income changes • Individual needs to demonstrate partial financial hardship to qualify

Source: https://studentaid.ed.gov/repay-loans/understand/plans

This table describes the various repayment plans available for Federal Direct Loans as of 2015.

1.B Derivation of Analytical First Order Condition:

Analytical Estimation:
When choosing a term, individuals chose T to maximize the discounted stream of yearly utility, which lead to the first order condition:

$$E[\sum_1^T \beta^t \frac{\partial d}{\partial T}(w_{it} - d_i)^{-\gamma}] = E[\beta^{T+1}(-d_i)(w_{iT})^{-\gamma}]$$

$$s.t.\ d_i = T * D_i * \frac{r(T, p_i)}{(1 - (1 + r(T, p_i))^{-T})}$$

d_i represents the yearly payment for individual i at term T, and $r(T, p_i)$ is the risk, term specific interest rate faced by individual i at term T.

I assume that log income follows the unit root process:

$$ln(w_{it}) = ln(w_{it-1}) + (X_i'\mu) + u_{it}$$

where $X_i'\mu$ is a yearly growth rate specific to observable characteristics and

$$u_{it} \sim N(0, \sigma_u^2)$$
$$\sigma_u^2 = (\omega - \upsilon * p_i)^2$$

is a individual-specific yearly income shock that is allowed to be a function of observable risk type p_i.

We observe starting income levels w_{i0}. This means that we can express income at time t as:

$$ln(w_{it}) = ln(w_{i0}) + t * (X_i'\mu) + \sum_1^t u_{it}$$
$$w_{it} = w_{i0} * e^{t*(X_i'\mu)+\sum_1^t u_{it}}$$

If we return to the uncertain portions of the right hand side of our first order condition, $E[(w_{it} - d_i)^{-\gamma}]$, note that we can rewrite the expected marginal utility as the marginal utility of a certainty equivalent given by:

$$E[(w_{i0} * e^{t*(X_i'\mu)} * e^{\sum_1^t u_{it}} - d_i)^{-\gamma}] = (w_{i0} * e^{t*(X_i'\mu)} * e^{\pi_{it}} - d_i)^{-\gamma}$$

where π_{it} is the certain amount an individual would have to be given in that period to make their certain utility equivalent to the expected utility. Specifically:

$$\pi_{it} = \frac{1}{2} * t * \sigma^2[1 - (1+\gamma)\frac{w_{i0} * e^{t*(X_i'\mu)}}{w_{i0} * e^{t*(X_i'\mu)} - d_i}] \qquad \text{for t<T+1}$$
$$\pi_{it} = \frac{1}{2} * t * \sigma^2(-\gamma) \qquad \text{for t≥T+1}$$

To derive π_{it}, note that we can write:

$$E[u'(w_{it})] = E[u'(w_{i0} * e^{t*(X_i'\mu)} * e^{\sum_1^t \sigma\epsilon_{it}})]$$

where $\epsilon_{it} \sim N(0,1)$. We want to find the value of $\pi(\sigma)$ that allows us to write:

$$E[u'(w_{i0} * e^{t*(X_i'\mu)} * e^{\sum_1^t \sigma\epsilon_{it}} - d_i)] = u'(w_{i0} * e^{t*(X_i'\mu)} * e^{\pi(\sigma)} - d_i)$$

For simplicity, start with the case of no income growth in period 1.

$$E[u'(w_{i0} * e^{\sigma\epsilon_{i1}} - d_i)] = u'(w_{i0} * e^{\pi(\sigma)} - d_i)$$

We first take the derivative of this expression w.r.t. σ:

$$E[w_{i0} * \epsilon * e^{\sigma\epsilon_{i1}} u''(w_{i0} * e^{\sigma\epsilon_{i1}} - d_i)] = \pi'(\sigma) w_{i0} * e^{\pi(\sigma)} u''(w_{i0} * e^{\pi(\sigma)} - d_i)$$

At $\sigma = 0$ this becomes zero since $E[\sigma\epsilon] = 0$ and thus $\pi'(0) = 0$.
We next take the second derivative of this expression w.r.t. σ, and evaluate it at $\sigma = 0$:

$$E[\epsilon^2 u''(w_{i0} - d_i) + \epsilon^2 w_{i0} u'''(w_{i0} - d_i)] = \pi''(0) u''(w_{i0} - d_i)$$
$$\pi''(0) = [1 + w_{i0}\frac{u'''(w_{i0} - d_i)}{u''(w_{i0} - d_i)}]$$

Under the assumption of CRRA utility, this becomes:

$$\begin{aligned}\pi''(0) =& [1 + w_{i0}\frac{u'''(w_{i0} - d_i)}{u''(w_{i0} - d_i)}] \\ =& [1 - (1+\gamma)\frac{w_{i0}}{w_{i0} - d_i}]\end{aligned}$$

We now have a value for $\pi''(0)$. This is helpful when evaluating a Taylor expansion of $\pi(\sigma)$:

$$\pi(\sigma) \approx \pi(0) + \pi'(0)\sigma + \frac{1}{2}\sigma^2\pi''(0)$$
$$\pi(\sigma) \approx \frac{1}{2}\sigma^2[1 - (1+\gamma)\frac{w_{i0}}{w_{i0} - d_i}]$$

Therefore our analytical estimating moment becomes:

$$g_i(\theta) = \sum_{1}^{T} \beta^t \frac{\partial d}{\partial T} (w_{i0} * e^{t*(X_i'\mu)} * e^{\pi_{it}} - d_i)^{-\gamma} - \beta^{T+1}(-d_i)(w_{i0} * e^{(T+1)*(X_i'\mu)} * e^{\pi_{i(T+1)}})^{-\gamma}$$

To estimate the model, I use nonlinear least squares, choosing the parameters that minimize the quadratic form:

$$b = \arg\min_{\theta} g_i(\theta)' g_i(\theta)$$

1.C Flexible Term Choice in the Government Sector

Thus far our analysis has assumed that individuals repaying in the Federal market have the flexibility to pay over any term. However, the true extent of this flexibility requires awareness and diligence on the borrower's part, and has been changing over time – for instance, the option to extend repayment for up to 20-25 years under an income based plan was introduced in only the last 10 years. The majority of borrowers are initially defaulted into a standard 10 year term when they enter repayment. Thus to pay the loan off faster, they must actively increase their monthly payment each month above that which they are billed, and to increase term they must enroll in a extended term plan. This means that the majority of borrowers remain in 10 year fixed payment plans – the latest aggregate figures from the Direct Loan Portfolio show that 51% of borrowers remain in 10 year fixed plans. Most private student loans also offer only a few term options (for example 7, 10, or 15 years).

Our model allows us to quantify the gains to consumers from flexible term choice relative to fixed maturity. Figure 1.16 plots the CV borrowers would require to be as well off under a 10 year repayment plan at 7% APR, as under a flexible repayment plan at 7% APR and at market rates. The gains are small – .005% of loan principal, or $240 for the average borrower in my sample – but significant for individuals who would optimally choose a shorter or longer loan at 7% APR, and grow as the optimal term choice diverges from 10 years. Flexible term choice leaves all borrowers better off under uniform pricing, whereas a combination of term choice and market pricing will leave short term borrowers (who also tend to be low risk) much better off and long term borrowers (who also tend to be high risk) worse off.

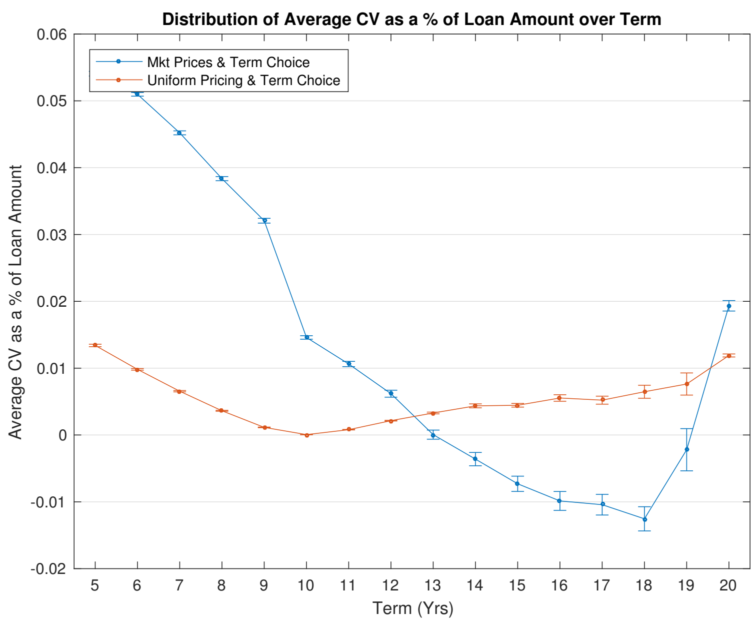


Figure 1.16: CV for Uniform 7% APR, 10 Year Fixed Plan relative to 7% with Term Flexibility and Market Prices with Term Flexibility

1.D Alternative Estimation Routines:

Simulated Method of Moments:
In addition to using the analytical moment condition, I also estimating the model using simulated method of moments with the moment condition:

$$g_n(\theta) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{S}\sum_{1}^{S}(\sum_{1}^{T_i}\beta^t\frac{dd_i}{dT_i}(w_{ist} - d(T_i)_i)^{-\gamma} - \beta^{T_i+1}(-d(T_i))(w_{ist})^{-\gamma})$$

choosing the parameters that minimize the quadratic form:

$$b_n = \arg\min_{\theta} g_n(\theta)' g_n(\theta)$$

where S is number of simulated income path draws. I use *simulated* moments because one cannot analytically express the first order condition in terms of the error term u_{it}. The estimation routine also accounts for truncation of term choices at 60 and 240 years by estimating a parameter that corrects the moment condition for observations at each endpoint. The exact estimation routine is described below. Let:

- N = number of individuals
- $T_{max} = 20$, the maximum length of a loan contract
- T = an individuals' chosen length of loan
- S = number of draws for simulation

Then:

1. Draw a $NxT_{max}xS$ matrix of normal random errors. This is the per-period shock that happens to income each period. This draw stays constant throughout estimation routine.
2. Start estimation with candidate vector of parameters: ($\theta = \{\gamma,\ \alpha_{60},\ \alpha_{240},\ \mu,\ \upsilon,\ \omega\}$).
3. For each individual, scale the random error by $\omega - \upsilon * p_i$. This allows the variance of income shocks to flexibly differ over risk score.
4. For each individual, calculate fixed consumption trend $X_i'\mu$, which is a linear function of income, debt amount, home ownership, FICO, family size, age, and degree type (dummies) and corresponding coefficients. This is the fixed amount that consumption (income minus fixed expenses) grows by each year.
5. For each individual and each simulation draw, calculate the path of future consumption. This is a AR(1) process given by: $y_{ist} = X_i'\mu + y_{is(t-1)} + u_{ist}$, where $X_i'\mu$ is the fixed consumption trend, $y_{is(t-1)}$ is the previous periods consumption, and u_{ist} is the current period shock. This creates S possible income paths for each individual.

6. Calculate the first order moment condition for each of these income paths using the current candidate set of parameters.

7. Calculate the mean of the S first order conditions for each individual - this provides the simulated moment for each individual that should in expectation be equal to zero.

8. If an individual has chosen the minimum term (60 months) or the maximum term, add the term α_{60} or α_{240} to their simulated moment. This accounts for the truncation of the choice set at these two points.

9. Sum the squares of the simulated moments across all individuals - this is the value to be minimized. Update set of candidate vectors accordingly.

10. Repeat steps 3-9 until at a minimum for sum of squared errors – these estimates are called $\hat{\theta}$.

The results from the simulated estimation are shown in column 5 of the structural estimates table, and are very similar to the analytical results. This table also includes several alternative specifications – in specification 2, I allow consumption paths to be deterministic, rather than uncertain. This equates to allowing income to grow at a constant rate that is specific to observable characteristics, and now includes risk score as one of these coefficients rather than as a determinant of the variance of the error term. This structural assumption has little impact on the estimates – γ has a very similar magnitude, and the coefficients relating other observable characteristics to free cash flow growth are of a similar size and sign to those in the primary specification. Higher risk types are estimated to have higher rates of income growth, when in the context of our model would rationalize their choices of longer terms, all else constant.

1.E Modeling Borrower Delinquency

In reality, borrowers are not able to default on their debt – student loans are not dischargeable in bankruptcy even when they are refinanced in the private sector, which removes the possibility of "strategic default". However, borrowers do sometimes become delinquent on their loans, which means they are late on their payments. This generates costs for the lender, and also has negative consequences for the borrower who receives a worse credit score and must repay the missed portion in the future with additional interest.

In our model, borrower income risk primarily drives decisions because of the possibility that a low income draw minus a large debt payment will generate a very high marginal utility. We could also specifically model the impact of delinquency as follows – note that it will not change optimal term choice.

Assume that in expectation, an individual chose term T^* as their optimal term. By definition, this means that $E[U(T^*)] > E[U(T^*+1)]$ and $E[U(T^*)] > E[U(T^*-1)]$.

I now introduce a threshold subsistence consumption level x: if $w_{it} - d_i$ falls below x, I assume an individual will not be able to make a full debt payment that period. Instead they will pay up to x, and have to pay the remaining amount plus interest, as well as incur a penalty, at the end of their current loan term:

$$\begin{aligned}
&\text{If } w_{it} - d_i < x \\
&\qquad\qquad \text{Pay} \quad n = w_{it} - x && \text{in period } t \\
&\qquad\qquad \text{Pay} \quad (d_i - n)(1+r)^{T^*+1-t} + R_{it} && \text{in period } T^*+1
\end{aligned}$$

It is more likely for individuals with low w_{it} and high income variance p_i that $w_{it} - d_i = x$ binds.

Assume that for individual i in period t the consumption condition is binding. This will change total utility as it it currently modeled ($E[U(T^*)]$) to $E[U(T^*)']$as follows:

$$\begin{aligned}
E[U(T^*)'] =& E[U(T^*)] - \underbrace{\beta^t(u(w_{it} - d_i) - u(w_{it} - n))}_{\text{Additional utility in period } t \text{ when constraint binds}} \\
&- \underbrace{\beta^{T^*+1}(u(w_{iT+1}) - u(w_{iT+1} - ((d_i - n)(1+r)^{T^*+1-t} + R_{it})))}_{\text{Negative utility in period } T+1 \text{ when pay additional interest \& penalty}}
\end{aligned}$$

If the penalty R_{it} is set such that:

$$\beta^t(u(w_{it} - d_i) - u(w_{it} - n)) = \beta^{T^*+1}(u(w_{iT+1}) - u(w_{iT+1} - ((d_i - n)(1+r)^{T^*+1-t} + R_{it}))) - \epsilon$$
$$\epsilon > 0$$

then:

$$\begin{aligned}
& & E[U(T^*)'] =& E[U(T^*)] - \epsilon & \\
&\text{and} & E[U(t)'] =& E[U(t)] - \epsilon & \forall t \neq T^*
\end{aligned}$$

This means that:

$$
\begin{aligned}
& E[U(T^*)'] < E[U(T^*)] \\
\text{and} \qquad & E[U(T^*)'] > E[U(t)'] \qquad \forall t \neq T^*
\end{aligned}
$$

This rules out strategic default (a borrower will never want to default on their loan), and also means that the optimal term choice will not be impacted by the possibility of delinquency.

1.F Model Estimates

Table 1.10: Choice Model Parameter Estimates

Parameter	Log Income Process	Non-Log Process	Constant Income	Simulated MM
γ	1.119	1.289	1.848	1.024
	(1.0833 , 1.1546)	(1.2027 , 1.3735)	(1.5228 , 2.1735)	
μ - Constant	7.49E-02	9399.965	–	2732.100
	(0.055561 , 0.094261)	(7837.9 , 11101)	–	
μ - Starting Income	7.48E-03	0.901	-0.169	0.910
	(0.0060717 , 0.0088882)	(0.79995 , 1.0042)	(-0.4772 , 0.13858)	
μ - log(Debt)	-9.35E-03	-611.649	-0.386	-303.060
	(-0.010121 , -0.0085806)	(-692.5 , -532.87)	(-0.63297 , -0.13809)	
μ - Home Owner	-8.21E-05	8.793	-0.203	16.577
	(-0.00086079 , 0.00069657)	(-96.163 , 113.53)	(-0.41776 , 0.012249)	
μ - # Dependents	5.31E-04	163.838	-0.446	276.350
	(0.00011908 , 0.00094268)	(96.153 , 237.02)	(-0.57852 , -0.31353)	
μ - Age	1.44E-04	108.083	-0.047	103.920
	(-0.00033136 , 0.00061954)	(36.004, 185.86)	(-0.18916 , 0.095549)	
μ - Age^2	-2.51E-06	-1.119	0.000	-1.790
	(-8.739e-06 , 3.7159e-06)	(- 2.2437 , -0.067359)	(-0.0018908 , 0.0019279)	
μ - Degree MD	4.05E-04	-391.035	0.449	-601.340
	(-0.0026784 , 0.0034888)	(-791.4 , 12.713)	(-0.39542 , 1.2935)	
μ - Degree JD	-4.42E-03	-607.457	0.586	-992.690
	(-0.0056486 , -0.003201)	(-798.38 , -419.51)	(0.25414 , 0.9177)	
μ - Degree Masters	-2.11E-03	-297.765	0.239	-506.180
	(-0.0029002 , -0.0013099)	(-418.72 , -182.43)	(0.025228 , 0.45221)	
ω	4.45E-04	8.018	–	440.390
	(-0.00080498 , 0.0016949)	(-1.782 , 17.818)	–	
υ - Risk Score	-2.22E-04	-33.211	0.571	-81.936
	(-0.0038 , -0.0035)	(-63.787 , -2.635)	(0.47069 , 0.67065)	
υ - Age	–	–	–	
	–	–	–	
α_{60}	-0.89	0.247	2.038	-0.127
	(-.1.1462 , -.55437)	(0.15438 , 0.3401)	(1.7933 , 2.2825)	
α_{240}	1.23	1.92	-12.69	3.33
	(0.65173 , 1.45392)	(1.3815 , 2.4539)	(-13.897 , -11.474)	
N	11585	11585	11585	11585

95 CI in parentheses

This table presents results from our structural model. The first column estimates come from our preferred specification, which directly models log income as a unit root process with a growth rate specific to a host of observable characteristics, including age and age squared, degree type, loan amount, and starting income. Column 2 uses the same specification but lets income (not log income) grow as a unit root process. Column 3 assumes constant income over time and controls for a vector of observable characteristics. Column 4 uses method of simulated moments, rather than nonlinear least squares.

1.7 Additional Figures and Tables

Table 1.11: Budget Lifetime Default Rates

Year Loan Enters Repayment	**2007**	**2008**	**2009**	**2010**	**2011**
2-Year Non-Profit/Public	32.10%	31.60%	31.10%	31.40%	33.80%
2-Year Proprietary	47.30%	48.60%	49.00%	48.40%	49.40%
4-Year Freshmen & Sophomores	24.70%	24.00%	23.60%	24.20%	25.40%
4-Year Juniors & Seniors	12.40%	12.30%	12.10%	11.90%	13.00%
Graduate Students	6.20%	6.20%	6.10%	6.10%	6.40%
Weighted Average	*15.90%*	*16.50%*	*17.30%*	*17.60%*	*18.40%*

Source: U.S. Department of Education (based on figures published in fiscal year 2014)

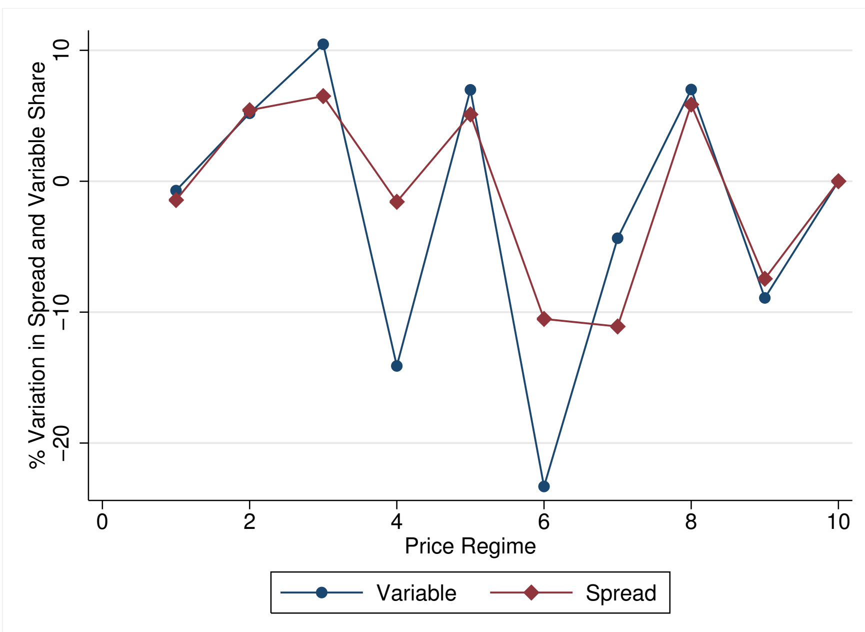


Figure 1.17: Response of Share Choosing Variable to Exogenous Variation in the Fixed Variable Spread:

This figure shows how interest rates changed over time under various price regimes and the resulting changes in borrower rate choices– it plots the average residualized price variation and rate choices of all borrowers over price regimes. To create these residuals, variable rate choice and the fixed variable spread were first regressed on risk score, to remove risk based price variation. Rate choices are highly responsive to changes in the level or spread of interest rates.

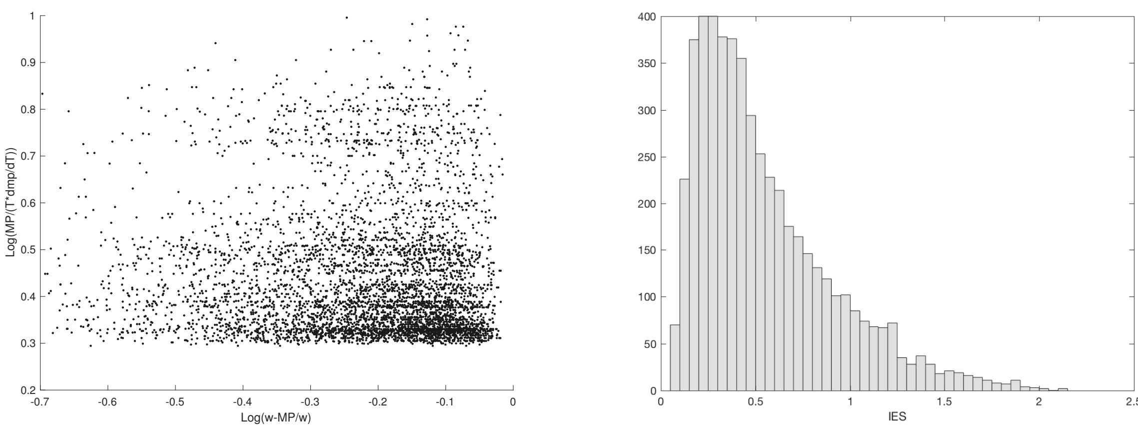


Figure 1.18: Distribution of Estimated Values of γ using Baseline Model

These plots show the value of γ that would be necessary to rationalize the term choice of each individual under the assumption of constant income and $\beta = 1$.

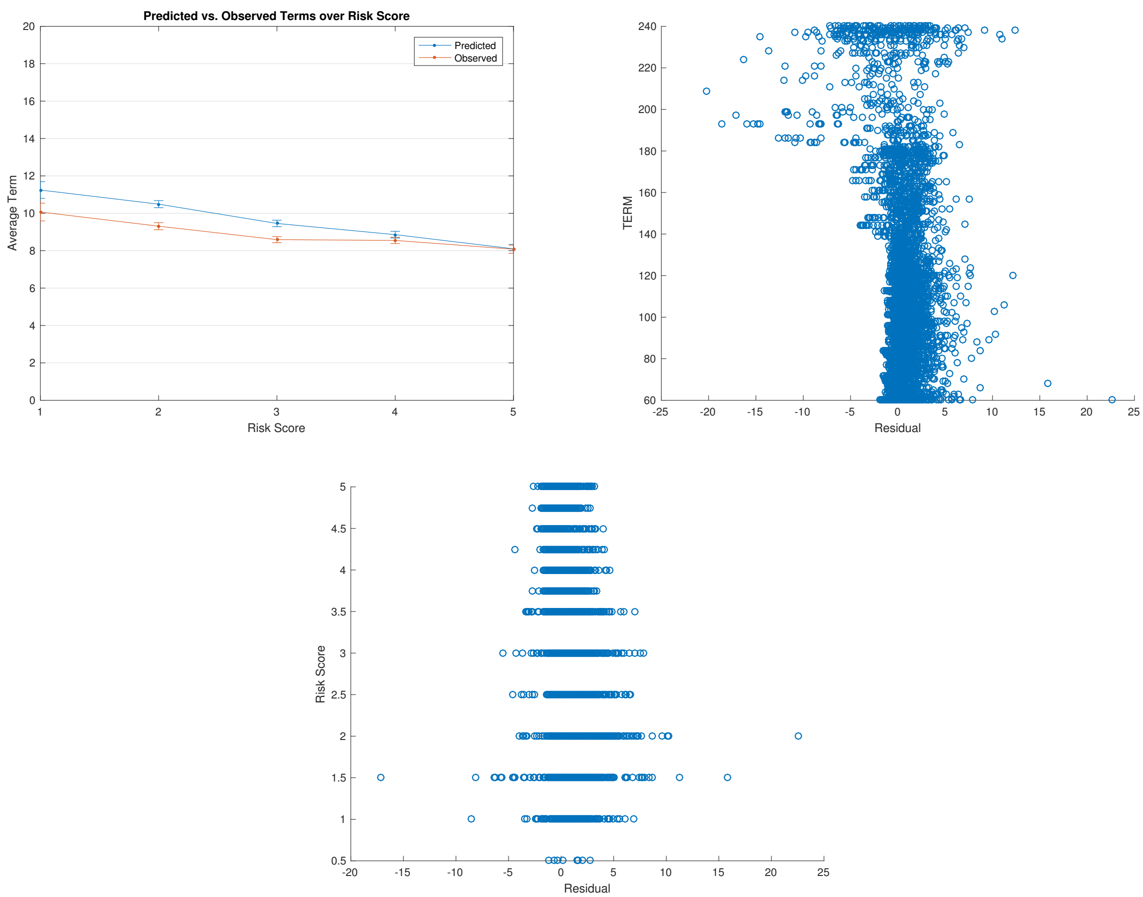


Figure 1.19: Model Fit:Predicted vs. Observed Term Choices, Over Risk, Residual over Term Choices, and Residual over Risk Score.

These figures analyze the model fit, comparing observed and predicted term choices, as well as the model residual over term and risk score. They show that in general the model slightly over predicts terms, but otherwise seems to perform well. All counterfactual exercises use these predicted term choices as a comparison point.

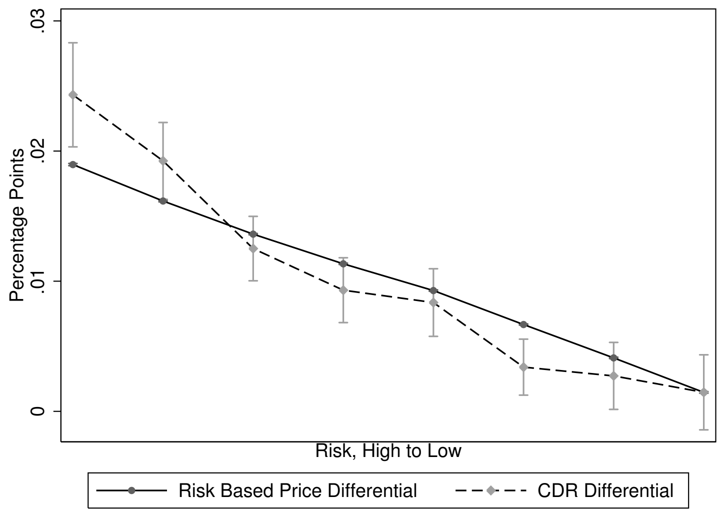


Figure 1.20: Cost Differential Relative to Lowest Risk Rating

The CDR is calculated and published by the Federal government at the school level, and reflects the student loan default rate of a cohort of students from that school after 3 years of completion. It is a much cited measure of expected costs used by Federal loan program. This figure compares the difference in the CDR between the highest and lowest risk types in my sample (which is roughly 2 percentage points) to the spread in their risk-based interest rates, and shows that private sector risk scores are highly correlated with the CDR.

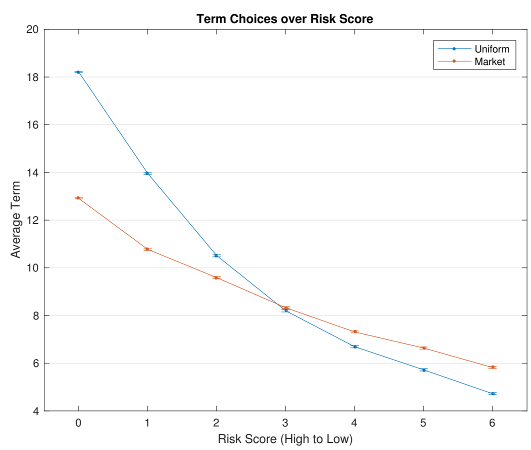


Figure 1.21: Predicted Term Choices over Risk Score under Uniform 6.6% vs. Market Prices

This figure shows the average term choices predicted by our model for borrowers in different risk bins under market pricing and uniform pricing – as shown in the theoretical graph above, lower risk borrowers decrease their terms, while higher risk borrowers increase their terms. When these term responses are translated into term elasticities, one can see that the elasticities generated by our structural model are very close to the reduced form elasticities.

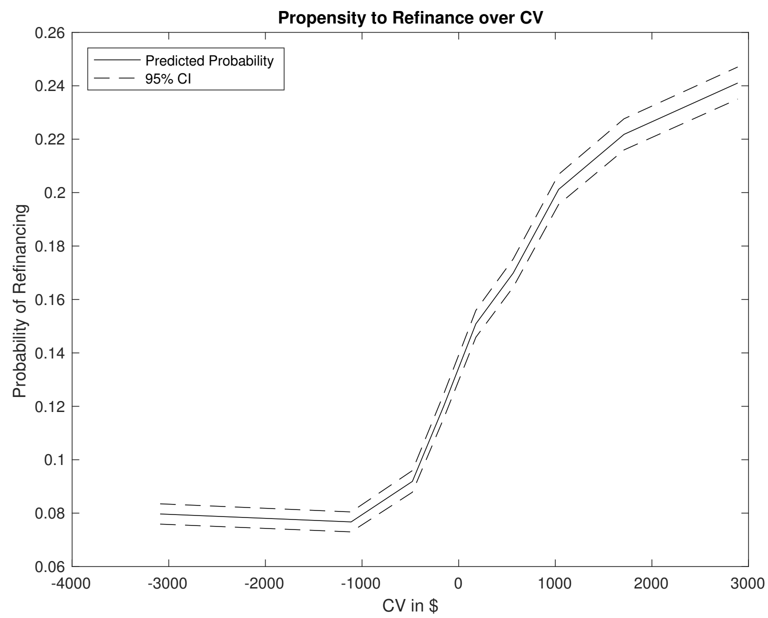


Figure 1.22: Propensity to Refinance over CV

This figure plots the observed probability that an individual in our sample attempts to refinance (i.e. they fully fill out an application after seeing a risk-specific interest rate quote) against their estimated CV, and shows the strong relationship between this extensive margin response and the welfare loss experienced under uniform pricing. Borrowers who require a $1000 CV instead of a $500 CV to be as well off under uniform pricing are 25% more likely to try to refinance into the private sector. It is also interesting to note that the propensity to refinance increases gradually with the associated financial incentives and that not all applicants who could benefit from refinancing necessarily apply. This suggests that there are potentially switching costs associated with refinancing, and/or that some individuals may value certain aspects of the Federal repayment program (like income based repayment) more than the interest rate savings they could achieve by refinancing.

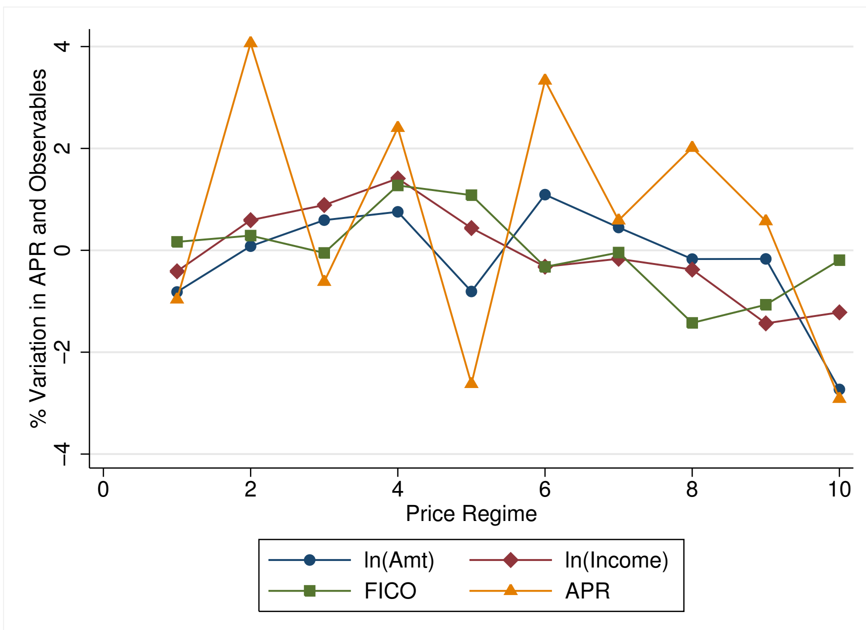


Figure 1.23: Variation in Observable Characteristics over Time

This graph shows changes in three important observable characteristics, income, debt amount, and FICO score, over 10 price regimes. While there are differences across price regimes, it is comforting to note that there are no obvious monotonic trends in these three variables and that they are not correlated with the experimental price shifts.

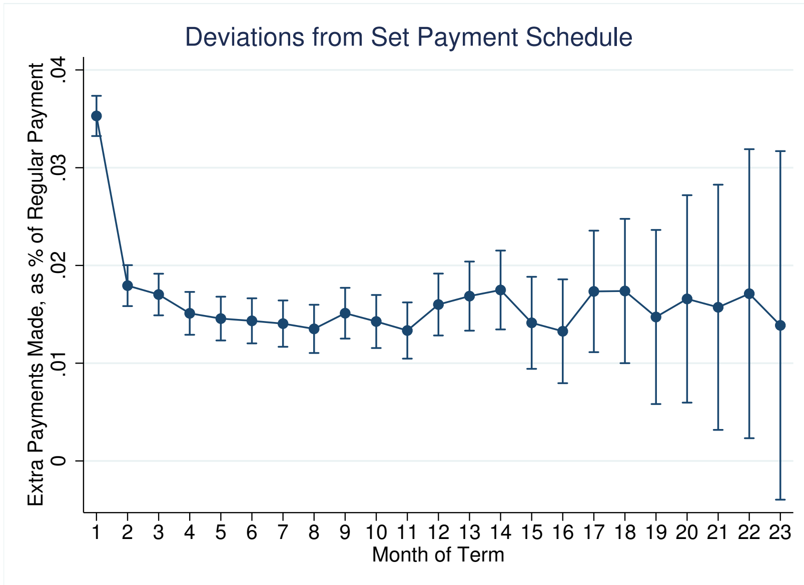
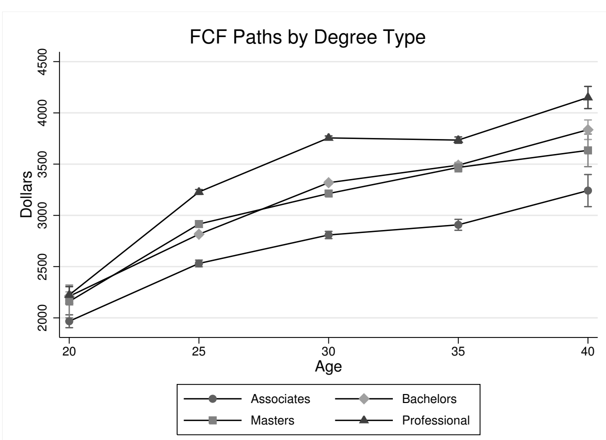


Figure 1.24: Average Size of Extra Payments Over Time Made by Borrowers

I look at payment patterns over time within my sample of refinancers – in other words, do any individuals change their payment level over time permanently, or do they systematically make higher or lower payments on their debt. I find that there are some extra payments in the data, but they are small and do not vary systematically over time. This supports our model's assumption that borrowers make a term choice in year 1 to maximize expected utility over the life of the loan and are not in fact choosing a monthly payment to fit their *current* income level, with the intent to refinance and change term yet again in the future when their income level changes.



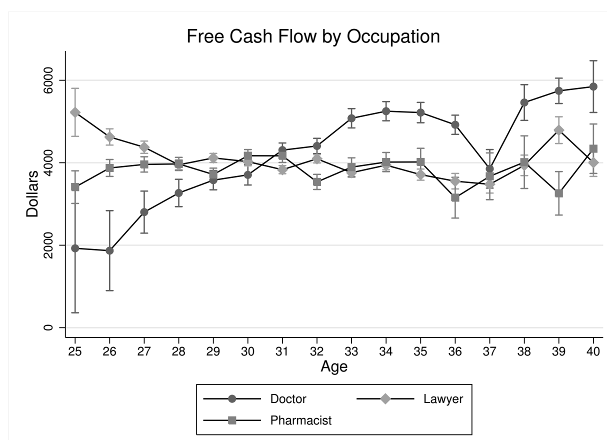


Figure 1.25: Cross-Sectional Age Earnings Profiles, by Degree and Occupation

Here I plot free cash flow paths for individuals with different degree types and occupations. There are notable differences in both the level and the changes in free cash flow over the lifetime for each of these groups.

	Initial Monthly Payment			
	Initial Auto	*Initial Real Estate*	*Initial Non-Real Estate*	*Initial Credit Card*
Mean	449.32	1927.72	1104.31	94.37
Median	386.00	1698.00	882.00	52.00
IQR (25,75)	246.50	1210.00	897.00	84.00
	Change in Monthly Payment			
	Auto Change	*Real Estate Change*	*Non-Real Estate Change*	*Credit Card Change*
Mean	27.15	63.64	-181.06	-12.74
Median	0.00	0.00	-75.00	0.00
IQR (25,75)	0.00	37.00	490.00	52.00

Figure 1.26: Levels and Changes in Other Monthly Payments Before and After Refinancing
My model assumes that borrowers are not readjusting on other financial margins when refinancing. In other words, contemporaneous savings and debt decisions are assumed to be exogenous, predetermined, and unaffected by maturity and refinancing decisions. Here I test this assumption by looking at borrowers' other monthly payments before and after refinancing, and measuring whether they adjust immediately during refinancing. This table describes changes in other monthly payments (mortgages, auto loans, credit cards, etc) before vs. after refinancing for individuals who had positive monthly payments to begin with, and shows that for the vast majority of borrowers these stayed constant. This makes sense, since many of these payments are fixed installments, and it would take active work on the borrower's part to readjust.

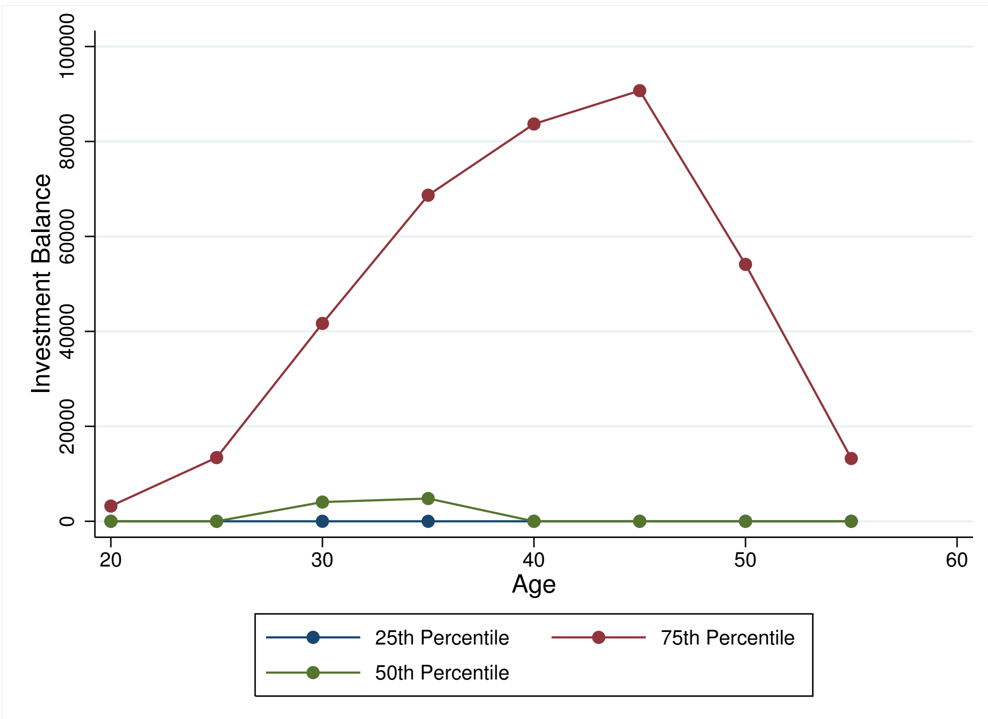


Figure 1.27: Investment Balances over Lifetime

My model defines yearly consumption as income minus the student debt payment; in reality individuals may also be making savings decisions that could impact their maturity choices. I can observe the savings and investment behavior of borrowers in my sample: because individuals in my sample are young, they have relatively low levels of savings to begin with. Slightly under 40% have a formal retirement savings account – for example 25% have a 401k, with a median balance of $24,000. The number of individuals with investment holdings increases with age. This figure shows that while the median borrower continues to not have substantial savings through age 60, the 75th percentile has accumulated over $80,000 by age 50. However, 90% of my borrowers are under 40 years old, and therefore even the most active savers have investment holdings that are much smaller than their student debt amount.

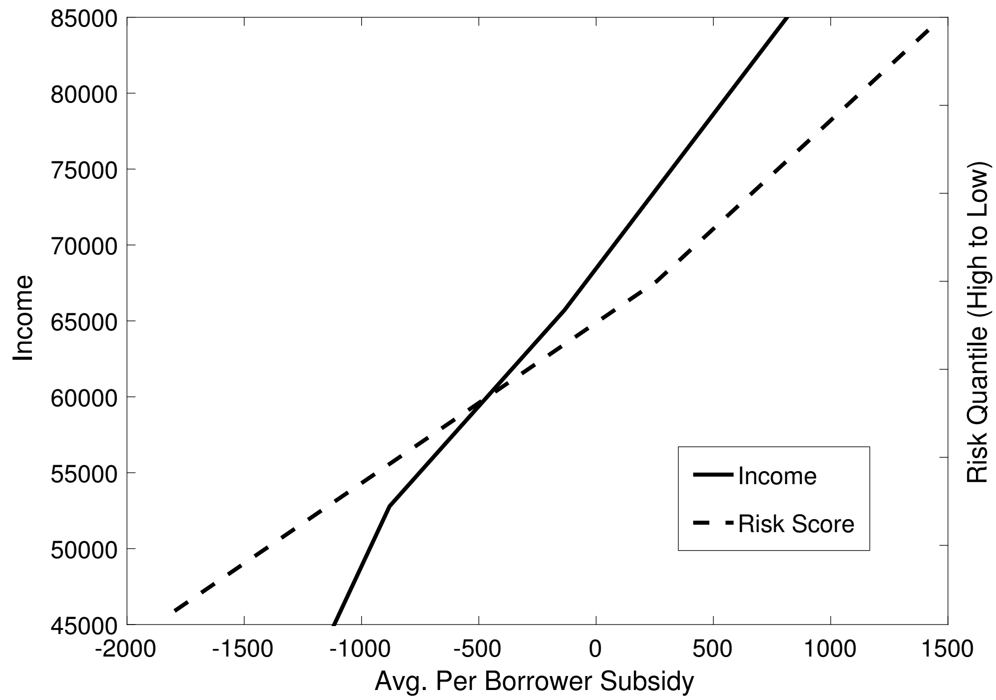


Figure 1.28: Average Government Subsidy over Borrower Risk Type and Borrower Income
This figure plots the average subsidy given to each borrower under uniform pricing over both borrower risk type and borrower income. The lowest income borrowers get a subsidy of slightly more than $1000, while the riskiest borrowers get an average subsidy of almost $2,000. This is because income is not perfectly correlated with risk type or maturity preferences (the two dimensions that differentiate costs and thus directly generate redistribution) the uniform rate is an imperfect instrument for achieving redistribution over income.

Chapter 2

The Influence of Risk Aversion and Switching Costs on Households' Financial Choices

2.1 Introduction

Online lenders often differentiate themselves from traditional banks by offering financial products that are more flexible and personalized – for example, loan products that allow individuals to customize their monthly payment down to the dollar, or robo-advisors that tailor investment portfolios. This movement towards a more complete contract space would seem to be welfare improving, at least in a frictionless setting with symmetric information.

However, consumers could make sub-optimal decisions in the presence of misinformation and/or high switching costs, and these frictions can be exacerbated in settings where consumers face more complex choices. In both the health insurance and mortgage markets, research has shown that consumers often make incorrect initial decisions, and fail to ever update their initial choices dynamically, even as the associated risk and costs evolve. Work has also shown that these frictions are often positively correlated with other welfare relevant consumer characteristics, like risk type. This research suggests that this offering a wider, more flexible set of financial products online may have ambiguous, and importantly heterogeneous, impacts across different types of consumers. If firms are aware that these frictions exist for certain borrowers, they may also have implications for how products are priced.

In this paper, I empirically measure the frictions, namely switching costs, present for borrowers making fixed/variable rate decisions in an online lending context, and analyze how these switching costs are correlated with borrower liquidity constraints. I use an individual-level panel dataset of borrowers making active and inactive interest rate decisions (a fixed or adjustable rate) on large personal loans. Importantly there are no explicit costs for updating the rate on the loan once every six months. Over the sample time period, the available interest rates on the fixed and variable rate loans changed considerably. Comparing the active and inactive choices in these different interest rate environments allows me to separately identify switching costs from risk preferences. I present both reduced form evidence on the determinants of rate decisions and the presence of switching costs, and also estimate a structural model that maps these findings to the risk aversion coefficient and a switching cost parameter.

I observe a wide set of socioeconomic and risk-related variables about each borrowers in addition to their rate choices, including their risk score, income, degree type, occupation, employment history, location, age, credit score and report, assets, liabilities, and savings behavior. This information importantly allows me to identify borrowers who are more likely to be liquidity constrained using the level of interest rates that they face, debt balances, and credit histories. Past work in the mortgage literature has shown that liquidity constraints can be a determinant of rate choices, with constrained borrowers making decisions that are more influenced by current, rather than expected future, interest rates and monthly payments. In my setting, liquidity constraints appear to again be a strong determinant of rate decisions – borrowers who are high risk, and therefore face higher interest rate levels, are much more likely to choose a variable rate all else constant. They are also the least likely to ever switch from a variable to a fixed rate contract.

I also present reduced form evidence on how active and passive rate choices are influenced by temporal variation in interest rates. I show that for *active* initial choices, the future LIBOR rate is a strong predictor of today's fixed rate choice – namely, a 10 bp increase in the future rate increases the share choosing a fixed rate loan by 2.7 percentage points (about 4%). The future rate is also a much stronger predictor than the current LIBOR rate, which serves as the index for the variable rate loan, for these initial choices. Large jumps in the current LIBOR predict spikes in the rate of switching to a fixed rate amongst existing variable rate loans. However, the *level* of switching in the data remains very low, suggesting that there are large costs to changing rates.

I ultimately use this reduced form evidence to estimate a structural model of interest rate choice and switching costs. The model captures the initial comparison borrowers make between the certain utility they derive from a fixed rate loan vs. the uncertain expected utility they derive from a variable rate loan whose interest rate could change over the next year. All borrowers face a lower monthly payment today under the variable rate contract, but if the expected fixed rate is much higher in the next period then they may still prefer to lock in the current fixed interest rate. In this initial active choice, there are no switching costs involved. I model the subsequent years' rate choices as a comparison of these two expected utility quantities plus an additional switching cost. A key assumption of this modeling framework is that individuals are naive - they make their initial rate choice assuming that they will re-optimize the following year without a switching cost.

Using a methodology from the health insurance literature, I identify individuals' risk preferences by observing their *initial* rate choice when they first refinance their loans. I call this an active choice, because switching costs are not present. For the entire population I estimate a coefficient of absolute risk aversion of .0564, which implies a high level of risk aversion: specifically, the value X that would make an individual with our estimated risk preferences indifferent between inaction and a gamble with a 50 percent chance of gaining $100 and a 50 percent chance of losing $X is $12. I then use the subsequent decisions of the same borrowers to update or stay with their current rate to estimate a lower and upper bound on switching costs. I estimate an average lower bound (identified off of individuals who should switch but do not) of $166, and an average upper bound (identified off of individuals who do switch) of $1,185. Each period I observe both new and old borrowers making these rate decisions under the new prevailing interest rate environment. This allows for an additional identification opportunity, since one can compare the active and inactive rate decisions cross-sectionally over borrowers at any given point in time, as well as within borrowers over time.

My analysis relates to two literatures: those analyzing households' choice of fixed and variable rate financial contracts, and those measuring switching costs in consumer choice settings. The decision between fixed and variable interest rates has been well studied in the mortgage literature - this is at least partially motivated by the popularity of adjustable rate mortgages (ARMs) in the 2000s, and their potential contribution to the sub-prime mortgage crisis. Some papers attribute variable rate decisions to forward looking behavior that seeks to minimize costs, whereas others (Campbell and Cocco) also allow current borrowing constraints to impact choices. These latter set of papers emphasize the fact that adjustable rate mortgages often appeal to high risk, liquidity constrained households (Johnson Li 2011) who seek immediately lower interest rates and monthly payments. However, if the variable rates on these loans increase, this constrained population is also the least able to adjust to increased monthly payments. My paper speaks empirically to this point – I also show that households where borrowing constraints are more likely to bind are much more likely to choose a variable rate (with a currently lower rate), whereas more liquid households are more sensitive to movements in the expected future interest rate. Campbell et al 2015 conducts a related panel analysis of how interest rates impact the fixed vs. adjustable rate mortgage decision in an international setting, allowing the spread and future interest rate expectations to impact choices. They also find that household seem to anticipate and take into consideration movements in the variable rate for one year after origination.

There has been extensive analysis and modeling of switching costs in other settings involving risk and complex products. In a study of health insurance choices, Handel estimates that the average individual forgoes $2,032 annually due to inertia. He also shows that reducing this inertia would be welfare decreasing in equilibrium, since it reduces the extent of adverse selection in the market. In our setting, reducing frictions could also have a complex equilibrium outcome if it impacts how firms set prices. In the household finance setting, Andersen, Campbell, and Nielsen present evidence of substantial inertia during refinancing decisions in the Danish mortgage market, where there are also low barriers (aka no penalties) to refinancing. They show that this inertia decreases with "income and housing wealth but increases with financial wealth".

It is important to note that "financial literacy" (or understanding of financial contracts) may also play a role in the fixed variable choice; for example, Bucks and Pence (2008) show that borrowers often underestimate the size of interest rate or monthly payment changes under an ARM contract, and are unfamiliar with the details of their mortgage terms. My data does not contain information on financial literacy, which is important for interpretation of my findings. For example, if literacy is correlated with borrower liquidity, I will attribute its impact to this observable characteristic in both the reduced form and structural evidence.

2.2 Setting and Data:

I use a proprietary dataset from an online lending firm that observes the choices and choice sets offered to borrowers refinancing large, uncollateralized loans (the median loan size is \$51,526, the mean \$71,483). The dataset has a panel structure, and follows all borrowers refinancing with the firm from July 2015 through February 2017; it links background financial information (debt amount, income, assets, credit score) about borrowers with the menu of interest rates they faced and the ultimate rate type and maturity choices they made when refinancing with the firm.

When borrowers first comes to refinance their loans, they are forced to make an active choice between the fixed and variable rate; there is no "default" option. In all subsequent months they remain defaulted into their initial chosen rate type, with the option to costlessly change the rate type up to once a year. Therefore, we observe all borrowers both in a setting where switching costs are not present (they are forced to make an active choice), and where they are present (they are defaulted into their existing rate type).

In the months following many borrowers' active choices, the index for the variable rate rose significantly. This makes this an ideal setting to measure switching costs, since the increase in interest rates made it optimal for many borrowers to switch their rate type from variable to fixed in the absence of switching costs. I therefore merge the borrower dataset with information on the prevailing interest rate environment at each point in time – namely data on the current interest rate, and expectations about future interest rates. The sections below describe the data used in further detail.

Borrower Summary Statistics and Panel Description

The population of borrowers are high earners, high debt, and highly educated. The majority (70%) hold higher than a bachelor degree, are in their early to mid-30s (IQR = (29, 35)), and earn a post-tax median income of \$67,500. Given this graduate student profile, and the fact that they have attended many years of schooling, it is not surprising that they also hold large amounts of personal student debt. The average monthly payment on refinanced debt, assuming a 10 year fixed rate repayment term, is \$600 per month.

The richness of our data allows for a more thorough description of these borrowers beyond monthly income and debt amount. Over 40% are home owners, they spend a median of \$1,300 on housing each month, they have an interquartile range of FICO scores ranging from 760 to 800. In terms of assets and liabilities, the median borrower holds \$38,000 in assets, \$0 in investments (the 75th percentile has \$15,000 in investments), owes \$89,000 in liabilities, and has a median monthly free cash flow (post tax income minus student debt and other monthly payments) of \$3,100. Borrowers hold a host of degrees and occupations; JDs (lawyers) make up 13% of the sample, MBAs are 17%, MDs 5%, pharmacists 6%, and dentists 4%.

The dataset is structured such that I observe new cohorts of individuals refinancing each month, and then follow the repayment history of individuals in each cohort over the following months. The summary statistics for the panel can be found in Table 2.1. Approximately 800 new individuals refinance (and make an "active" rate choice) every month, and they are followed on average in the dataset for 10 months, although individuals who signed earlier are followed for a longer period of time.

Fixed vs. Variable Rate Decision

When refinancing their loans, borrowers are offered a menu of loan terms at interest rates that are specific to their risk type. From this menu, they select a maturity for their loan (which ranges from 5 to 20 years) and a rate type. These choices determine the APR on their loan, and the monthly payment, and the total paid over the life of the loan. I focus here on the choice of rate type.

There are several key differences between the fixed and variable rate loans: in terms of APR, the fixed rate is priced at a premium to the variable rate due to the fact that is carries no interest rate risk. Holding risk score and maturity constant, there is on average a 1.1 percentage point rate spread between fixed and variable rates. The interest rate on the fixed rate loan will remain constant over the life of the loan, where as the interest rate on a variable rate loan can rise or fall over the life of a loan with the three-month LIBOR index. Table 2.2 shows examples of how the variable rate could vary with the LIBOR, using examples of historical LIBOR rates from the past two decades. The three-month LIBOR rate at the time of this analysis was very low, but has reached much higher levels in the past.

Due to the difference between the fixed and variable interest rates, there is also a difference in monthly payments and total payments under either contract. At least initially in my setting, variable rate monthly payments are lower than fixed rate monthly payments – Table 2.2 shows that for a \$50,000 loan being paid off over 10 years, the initial monthly payment would be almost \$50 lower under a variable contract. However, monthly payments will fluctuate with the interest rate on a variable rate loan. If interest rates increase, the monthly payment will as well.

After borrowers initially select a rate type (fixed or variable), they can actively switch from a fixed to a variable rate, or vice versa, up to once a year; if they do not make the active decision to switch rate types, they will remain with their current rate type. However, if a borrower decides to switch, they will be offered the *current* prevailing interest rates – for example, if interest rates begin to rise and they switch from a variable to a fixed rate, they will be offered a higher fixed rate then they would have received initially. Therefore, the option to easily switch between rates each year decreases the risk associated with choosing a variable rate, but does not remove it entirely.

My dataset captures the fixed/variable rate choices and interest rates faced by several cohorts of incoming borrowers over an almost two year period. Figure 2.4 shows the 1-month LIBOR rate (which is used as the index for the variable rate contract) in each month of the sample time period, the initial share of new borrowers choosing a fixed rate each month, and the share of variable rate borrowers each month that switched to a fixed rate. The summary statistics suggest that while *initial* active choices are very sensitive to the prevailing interest rate, the switching costs to updating this rate choice are very high, since few individuals exercise the option despite large increases in the LIBOR rate.

While prevailing interests definitely impact borrower rate choices, the decision also depends on borrower specific characteristics.[1] Table 2.3 explores the fixed and variable populations on several of these key dimensions. The population that initially chose a fixed rate does seem to have a slightly lower monthly free cash flow and higher median maturity. One strong, interesting pattern that emerges in the data is that individuals with higher risk scores systematically are more likely to choose variable interest rates. Figure 2.1 shows this strong relationship between risk score and rate choices – individuals who are considered by the firm to be more risky are offered higher levels of interest rates, which is shown on the horizontal access as the offered five year fixed rate. The graph shows that individuals offered higher *levels* of interests rates, all else constant, are more likely to choose a variable rate loan.

There are several reasons why this relationship could exist in the data. For one, some individuals might have lower levels of risk aversion and thus prefer a riskier contract. These riskier preferences may also be why they are in a higher risk category to begin with. Or there may be differences in financial literacy or understanding across risk types. Some may not understand completely the risks of a variable rate, and this knowledge could be correlated with risk type. Finally, credit constraints could also play a role: individuals classified as high risk will face higher interest rates everywhere. While they may prefer a fixed rate contract in an unconstrained setting, they might choose a variable rate loan, which is priced at a discount, to lock in a lower interest rate.

Data on Actual and Expected Changes in Future Interest Rates

Expectations and realizations of interest rates will be an important determinant of fixed rate decisions: if rates seem likely to increase in the future or there is a great deal of uncertainty around a rate change, borrowers may feel it is a good time to "lock in" a low rate with a fixed rate loan. Similarly, if the realized risk premium on a fixed interest rate falls, or the rate on a variable loan increases, a borrower might choose a fixed rate.[2] Therefore, in addition to the data on borrower characteristics, I also use data on: the current interest rate on the borrower's loan, the current fixed and variable interest rate offered to borrowers at any point in time, market expectations about the mean change the next period's interest rate, and market expectations about the variance of the change in the next period's interest rate.

[1]Our later structural model suggests that namely, the likelihood of choosing a variable rate should decrease with the debt to income ratio and loan maturity.

[2]The role of these factors is formalized in the model section of this paper.

I observe changes in two important *realized* interest rates over the sample period: the index rate and the fixed variable spread. The 1-month LIBOR is the variable rate index, meaning if a borrower chooses a variable rate, their interest rate will increase or decrease with movements in the 1-month LIBOR. Over the sample period the 1-month LIBOR increased by about 60 basis points, which is a 15% increase off the average variable interest rate in the sample (3.97%).

The fixed-variable spread refers to the difference in the fixed and variable APR for a given risk type and loan maturity; if a borrower wants to choose or switch to a fixed rate loan, they will pay a premium in comparison to the variable rate. The spread covaries negatively, but not perfectly, with movements in the 1-month LIBOR. This spread should incorporate the market's expectations about where interest rates will be in the future. As the expected future change in the LIBOR becomes smaller, the risk premium for the fixed rate will also shrink. While the fixed-variable spread should imply something about where the market thinks rates are going in the future, the model I estimate specifically asks for separate measures of the mean expected change and variance of this change. In addition, the spread I observe is specific to the firm I study, since over the sample period the firm changed this spread several time for experimental, non-market reasons. Therefore, it is not a perfect proxy for market expectations.

I instead use historical data on the Eurodollar 90 day continuous future contract price and implied volatility to measure where the market (and thus the borrowers) believe interest rates will be in the future.[3][4] In Figure 2.2 I plot the current 1-month LIBOR rate, the Eurodollar future rate, and the share of my sample choosing a fixed rate. I am using a 90 day future rate, therefore it is at a higher level than the 1-month LIBOR rate[5]; however changes in the futures rate covary closely with the next period's realized rate, and, more importantly, is a very strong predictor of the *current* period's share choosing a fixed rate loan. This suggests that expectations about future interest rate movements do matter for borrowers' initial decisions.

In Figure 2.3 I plot the future rate along with the bounds that represent the market's implied volatility measure. The implied volatility represents how far the market believes the rate could change over one standard deviation; for example, a $10 stock with a 20 percent implied volatility has a 68 percent chance to be priced between $8 and $12 one year from now. If there is more uncertainty about the path of interest rates, there will be a higher implied volatility. This volatility plays an important role in the structural model.

[3]Historical price and the implied volatility for ED1 come from Bloomberg.

[4]Several studies (Piazzesi and Swanson 2004; Gurkaynak, Sack, and Swanson 2006) have shown that futures contracts on short-term interest rates are decent predictors of the future path of interest rates at least in the short term, but that they tend to generate excess positive returns due perhaps to a risk premia incorporated in the price.I do not current adjust for this risk premium, but currently leave this for future work.

[5]I adjust for the spread between these two rates in the structural estimation.

2.3 Preliminary Analysis:

In this section I further explore the empirical relationship between interest rates and variable rate choices. Panel (a) of Figure 2.4 shows the movement of the 1-month LIBOR, the share of new loans that are fixed rate (borrowers making active rate choices), and the share of old loans that are fixed rate over the sample period. The graph makes it clear that there is a very strong correlation between the share of incoming borrowers choosing a fixed rate and the prevailing interest rate, but considerably smaller share of individuals updating their loan term to reflect the changing interest rate environment.

Panel (b) plots the share of eligible borrowers (individuals with a variable rate loan) that moved to a fixed rate over the same time period. This "switching" share also covaries closely with the LIBOR rate – it jumps after large movements in the LIBOR, specifically in Dec. 2015 and February 2017. However, the overall share ever switching stays at a very low level, below .9% of eligible borrowers. Table 2.4 describes individuals who originally chose a variable rate by their eventual switching status. Individuals who switch tend to have larger loan amounts, significantly longer maturities, and be lower risk (as measured by a higher "risk score").

How does switching impact the interest rate that an individual faces on a loan? It is important to note that when individuals first switch from a variable to fixed rate, they move to a higher interest rate, but one that will not change over the remaining life of the loan. Panel (a) of figure 2.5 plots the average change in APR that borrowers experience in the sample when they switch rates, which is slightly over .7 pp, or 18% off a mean variable APR of 3.9%. Panel (b) plots the more gradual change in interest rate experienced by variable rate borrowers who have *not* switched – for some of the earliest borrowers, their variable rate has also increased substantially, by almost .6 pp or 15.3%.

The graphs above suggest that while new cohorts of borrowers forced to make *active* rate decisions are very interest rate sensitive, the cost of updating the rate once it is locked in is high. However, this analysis does not account for the fact that other relevant characteristics, like risk score or debt to income ratio, may vary over cohorts. I next explicitly control for these confounding factors using a linear regression framework. To analyze borrowers' initial rate decisions, I restrict the sample to individuals in the first period of refinancing, and run the following logistic regression:

$$F_i = \alpha + \beta^1 spread_i + \beta^2 p_i + \beta^3 LIBOR_i + \beta^4 LIBOR\_F_i + \beta^5 LIBOR\_Vol_i + \mu' X_i + \epsilon_i$$

where F_i is a dummy variable that is equal to one if the borrower chooses a fixed rate, $spread_i$ is the fixed variable spread for the 10 yr maturity, p_i is the individuals' risk score, $LIBOR_i$ is the prevailing 1-month LIBOR rate the individual faced when they refinanced, $LIBOR\_F_i$ is the 90-day LIBOR future price when the individual refinanced, $LIBOR\_Vol_i$ is the implied volatility when they refinanced, and X_i is a vector of observable characteristics including monthly free cash flow (FCF), log debt amount, age, and chosen loan maturity.

Several things are apparent from the results of this regression, shown in Table 2.5. For one, it is clear that in their initial choice individuals are sensitive to changes in the future expected LIBOR rate. Specifically, a 8.3 bp (10%) increase in the future rate increases the share choosing a fixed rate loan by 1.63 percentage points (about 2.4%). Borrowers are also sensitive to the risk premium charged for the fixed rate loan – when the spread increases by 16 bps (10%), making the fixed rate loan relatively more expensive, the share of individuals choosing a fixed rate contract falls by .76 percentage points (.113%). These results also again show that riskier borrowers as measured by their "risk score", who face higher levels of interest rates, are more likely to choose a variable rate contract all else constant.

We can use this regression to imagine what share of individuals each month would choose a fixed rate contract if there were *no* switching costs – this predicted share, and the observed share, are plotted in Figure 2.6. Using this predicted share controls for the compositional changes potentially occurring over cohorts, and thus is preferred to comparing new vs. old borrowers. The wedge between the two lines is what our model attributes to switching costs – by the end of the sample, this wedge is over 10 pp or one third of the observed variable share. In the next section I move to a structural utility-based framework in order to quantify the size of this switching cost in dollar terms.

2.4 Modeling the Fixed-Variable Rate Decision

To provide economic intuition for the choices I observe in the data, and link these decisions to parameters capturing risk aversion and switching costs, I provide a simple model of borrowers' choice of a fixed or variable interest rate. This decision captures the preferences of borrowers smoothing consumption over states of uncertainty. By locking in the current prevailing interest rate, the fixed rate provides "insurance" to the borrower against future volatility in interest rates. It is therefore priced at a premium to the variable rate, which is pegged to the prevailing market rate and thus can change over time. A more risk averse individual would be willing to pay a higher fixed rate premium to insure against this uncertainty, just as in the term choice scenario they would be willing to pay a higher term premium to smooth payments over time.

I model this choice by analyzing the marginal tradeoff that a borrower faces in choosing a fixed vs. variable rate loan. By choosing a variable rate, an individual has a lower immediate monthly payment $MP_t^V < MP_t^F$. However, if rates increase and they want to switch to a fixed rate loan in the next period, they will have to lock in a higher prevailing fixed rate (with a high monthly payment $MP_{t+1}^F > MP_t^F$) at that time and pay that higher rate over the remaining loan maturity. If an individual instead chooses a fixed rate, they loose the immediate utility of a lower monthly payment today, but they gain the insurance value of having the current fixed rate over the remaining life of the loan.

This marginal tradeoff becomes my key estimating equation – an individual will choose a fixed rate if:

$$U_t^F = T * u(w - MP_t^F) > u(w - MP_t^V) + (T-1)E[u(w - MP^F{}_{t+1})] = E[U_t^V]$$
$$u(w - MP_t^V) - u(w - MP_t^F) < (T-1) * (u(w - MP_t^F) - E[u(w - MP^F{}_{t+1})])$$

In my setting, individuals are aware that they will be able to change this rate (from fixed to variable, or from variable to fixed) once every year with no monetary cost. However, while there may not be a fee for switching the rate on the loan, the borrower might well incur a nonpecuniary "switching" cost from having to go online and update the terms of their loan contract. I therefore model the *initial* rate choice the borrower makes as a comparison between the certain utility they derive from a fixed rate loan vs. the uncertain expected utility they derive from a variable rate loan whose interest rate could change over the next year. I model the subsequent years' rate choices as a comparison of these two expected utility quantities *plus* an additional switching cost. A key assumption of this modeling framework is that individuals are naive – they make their initial rate choice assuming that they will reoptimize the following year without a switching cost.

To model the volatility of payments under a variable rate contract, I assume the monthly payment is tied to the market interest rate r_t, which changes with uncertainty over the following year. I model this process as:

$$ln(r_{t+1}) = ln(r_t) + e_t$$

where e_t represents a normally distributed error term with mean κ_t and $Var(e_t) = \sigma_t^2$. This means that in expectation, rates the next year will change by κ_t, and the uncertainty surrounding this expected change is parametrized by σ_t^2. I allow κ_t and σ_t^2 to vary over time, since I will ultimately try to calibrate them using empirical measures of market-based expectations at time t.

An individual's available fixed rate monthly payment will permanently increase or decrease when the interest rate changes by e_t. Specifically, the monthly payment will change as follows:

$$MP_{t+1}^F = MP_t^F + m_t$$

where $m_t \sim N(\bar{m}_t, v_t^2)$ is a transformation of e_t that represents the change in monthly payment that occurs when the interest rate changes. I approximate these values as $\bar{m}_t \approx \kappa_t * \frac{dMP^F}{dln(r)}$ and $v_t^2 \approx \sigma_r^2 * \frac{dMP^F}{dln(r)}^2$. This means market-wide interest rate shocks will actually have an "individual specific" impact on monthly payments, in the sense that the same variation in r will cause larger changes in MP^F for individuals who have greater debt amounts and higher base interest rates. This link makes the choice of term and rate type correlated, since individuals choosing longer terms will have higher base interest rates and lower monthly payments.

Initial, Active Rate Decision:

When first choosing between the two rates (conditional on term), an individual compares the known difference in utility they will derive under the two rate types *today* against the expected difference in utility they will face under the two rate types tomorrow over the remaining life of the loan if interest rates change. All borrowers in my sample face a lower monthly payment today under the variable rate contract, but if the expected fixed rate is much higher in the next period then they may still prefer to lock in the current fixed interest rate. In this initial active choice, there are no switching costs involved.[6] I assume that borrowers have a CARA utility function, $u(w) = -e^{-\gamma w}$, that their consumption is given by their income, w, minus the monthly payment on this debt, and that they have chosen a maturity of T on their loan. An individual chooses a fixed rate if:

$$u(w - MP_t^V) - u(w - MP_t^F) < (T-1)(u(w - MP_t^F) - E[u(w - MP^F{}_{t+1})])$$

The key simplifying assumptions here, that later lend to empirical tractability, are that:

- Income is not changing over time.
- There is no discounting
- Borrowers are naive and believe they will switch in the next period if they need to (i.e. they believe the switching costs are zero).
- We can rewrite expected utility in terms of its certainty equivalent.

The assumption that $m_t \sim N(\bar{m}_t, v_t^2)$ allows me to express expected utility from the variable rate contract $E[u(w - MP^F{}_{t+1})]$ in terms of its certainty equivalent, $u(w - M\bar{P}_{t+1}^F)$, which under the assumption of CARA utility means that:

$$\begin{aligned} M\bar{P}_{t+1}^F &= MP_t^F + \bar{m}_t + \frac{v_t^2}{2} * \frac{u''(c_t)}{u'(c_t)} \\ &= MP_t^F + \bar{m}_t + \frac{\gamma * v_t^2}{2} \end{aligned}$$

This value shows that even if the fixed rate does not increase at all in expectation ($\bar{m}_t = 0$) , the uncertainty related to the future rate lowers the expected utility by the risk premium $\frac{\gamma * v_t^2}{2}$. The certainty equivalent decreases as an individual becomes more risk averse (γ increases), or as the expected variance of the monthly payment increases (v_t^2).

This means we can express the fixed rate choice as holding if the difference in utility from having a lower variable monthly payment *today* is smaller than the difference in expected utility over the remaining term of the loan from having to potentially lock in a higher fixed rate tomorrow:

[6] Since an individual can change rate types once per year, the uncertainty of the variable rate contract only encompasses the potential changes in market interest rates that can happen over the next year.

$$u(w - MP_t^V) - u(w - MP_t^F) < (T-1)(u(w - MP_t^F) - u(w - M\bar{P}_{t+1}^F))$$

To simplify this equation further, and easily see how this choice relates to borrower and market specific variables, it is helpful to approximate the difference in utility on either side of the inequality as:

$$u(x_1) - u(x_2) \approx (x_1 - x_2)u'(x_1)$$

This makes the inequality:

$$\begin{aligned}(MP_t^F - MP_t^V)u'(w - MP_t^F) &< (T-1)(M\bar{P}_{t+1}^F - MP_t^F)u'(w - MP_t^F) \\ \rightarrow (MP_t^F - MP_t^V) &< (T-1)(MP_t^F + \bar{m}_t + \frac{\gamma * v_t^2}{2} - MP_t^F) \\ \rightarrow (MP_t^F - MP_t^V) &< (T-1) * (\bar{m}_t + \frac{\gamma * v_t^2}{2})\end{aligned}$$

This equation shows that the fixed variable choice depends on various factors:

- As the fixed variable rate spread increases, $(MP_t^F - MP_t^V)$ will increase, making the individual less likely to choose a fixed rate.
- As an individual's level of risk aversion increases (γ), the inequality will be more likely to hold, making an individual more likely to choose a fixed rate loan.
- Expectations that rates will increase or an increase in the variance of this expectation ($\bar{m}_t$ and v_t^2) will make an individual more likely to choose a to choose a fixed rate loan.
- An increase in the maturity of the loan will make an individual more likely to choose a fixed rate

Many of these predictions were confirmed empirically in our reduced form analysis, for example a negative coefficient on the fixed variable spread, and a positive coefficient on the eurodollar future rate and implied volatility.

Inactive Rate Decisions: Modeling Switching Costs

Once they have made their initial choice, the borrower faces this same problem again in the next period, but with a new switching cost μ. I include μ within the utility function, which means that the switching cost will be estimated in dollar terms. Now the borrower will choose a fixed rate if:

$$u(w - MP^V_{t+1}) - u(w - MP^F_{t+1} - \mu) < (T-2)(u(w - MP^F_{t+1}) - u(w - M\bar{P}^F_{t+2}))$$
$$\text{if Choice}_t = \text{Variable}$$
$$u(w - MP^V_{t+1} - \mu) - u(w - MP^F_t) < (T-2)(u(w - MP^F_{t+1}) - u(w - M\bar{P}^F_{t+2}))$$
$$\text{if Choice}_t = \text{Fixed}$$

If we make the same linear approximation as above, this becomes:

$$MP^F_{t+1} + \mu - MP^V_{t+1} < (T-2) * (\bar{m}_{t+1} + \frac{\gamma * v^2_{t+1}}{2})$$
$$\text{if Choice}_t = \text{Variable}$$
$$MP^F_t - MP^V_{t+1} - \mu < (T-2) * (\bar{m}_{t+1} + \frac{\gamma * v^2_{t+1}}{2})$$
$$\text{if Choice}_t = \text{Fixed}$$

Empirically, this parameter μ can be estimated by the number of individuals observed switching in the data as interest rates and interest rate expectations change.

2.5 Structural Estimation

In this section, I use the reduced form elasticity estimated in the previous section to calibrate the parameters laid out in the structural model. Note that we can rewrite the main decision rule as:

$$(MP^F_t - MP^V_t) < (T-1) * (\bar{m}_t + \frac{\gamma * v^2_t}{2})$$
$$\rightarrow \gamma > \underbrace{(\frac{(MP^F_t - MP^V_t)}{T-1} - \bar{m}_t)\frac{2}{v^2_t}}_{\tilde{\gamma}_{it}}$$

If we assume that γ is lognormally distributed and that the distribution does not change over time, then we get the expression that an individual will choose a fixed rate in period 1 if:

$$ln(\gamma) + \epsilon_i > ln(\tilde{\gamma}_{i1})$$

This inequality easily translates into a estimating equation that we can calibrate using the share of individuals we observe choosing a fixed rate in the sample over time:

$$Pr(Fixed_{i1}) = \Phi(ln(\gamma) - ln(\tilde{\gamma}_{i1}))$$

Let $\tilde{\gamma}_{i2}^F$ represent the probability of choosing fixed in period two given your last period choice was fixed, and let $\tilde{\gamma}_{i2}^V$ represent the probability of choosing variable in period two given your last period choice was fixed. Then:

$$\begin{aligned} Pr(Fixed_{i2}|F_{i1}) &= \Phi(ln(\gamma) - ln(\tilde{\gamma}_{i1}^F)) && \text{if initial choice = Fixed} \\ Pr(Fixed_{i2}|V_{i1}) &= \Phi(ln(\gamma) - ln(\tilde{\gamma}_{i1}^V)) && \text{if initial choice = Variable} \end{aligned}$$

Then for each chain of observed choices we have the probabilities given by:

$$\begin{aligned} Pr(F_{i1}, F_{i2}) &= Pr(F_{i1}) * Pr(F_{i2}|F_{i1}) \\ Pr(F_{i1}, V_{i2}) &= Pr(F_{i1}) * Pr(V_{i2}|F_{i1}) \\ Pr(V_{i1}, V_{i2}) &= Pr(V_{i1}) * Pr(V_{i2}|V_{i1}) \\ Pr(V_{i1}, F_{i2}) &= Pr(V_{i1}) * Pr(F_{i2}|V_{i1}) \end{aligned}$$

I estimate the distribution of risk aversion parameters that explains the fixed variable choices observed in the data, as well as the distribution of switching costs that explains the number of rate changes observed over the sample period. The intuition behind the identification of these parameters is as follows: by calibrating beliefs about interest rate risk (using market based probabilities), the risk aversion coefficient remains the only "free" parameter in the model that explains the fixed or variable decision at any given point in time. We would be able to estimate γ using a static cross section of fixed variable choices. Once the risk aversion parameter is nailed down with a single cross sectional observation, μ is then identified in the next period when interest rates have changed. This price change will make a new rate type optimal for some agents. If we do not observe them switching rate types (i.e. reoptimizing), this will be loaded onto the switching cost parameter.

Estimation of Risk Preferences using Initial Choices:

I initially estimate the model using only individuals' initial, active choices ($t = 0$) over the entire sample period. Individuals refinancing at different points in time faced different interest rate levels and expectations, but did not face switching costs since they were all making active decisions. Therefore, this exercise only provides an estimate of the risk aversion parameter, not the switching cost parameter.

Recall that if someone has a risk aversion parameter γ that is larger than $\tilde{\gamma}$, they will choose a fixed rate. Therefore as γ increases, this inequality should hold for more individuals, and we should see more people choosing a fixed rate. Figure 2.7 illustrates this logic for our sample – the solid line plots the increase in the predicted share choosing fixed in our sample as γ increases. The dashed line plots the observed share choosing a fixed rate, .67. The value of $\hat{\gamma}$ that best fits the observed share is .0564. I also estimate γ as a function of observable characteristics, namely loan amount, monthly free cash flow, and risk score. The results (shown in the second column of Table 2.6) imply that as individuals become more wealthy, have larger loans, and are better risk scores, they make more risk averse decisions.

Since it is difficult to interpret the numeric value of the coefficient of risk aversion, one convention is to express the level of risk aversion in terms of the value X that would make an individual with our estimated risk preferences indifferent between (i) inaction and (ii) accepting a gamble with a 50 percent chance of gaining \$100 and a 50 percent chance of losing \$X. As an individual becomes more risk averse, the value of X will go to zero. In this setting, $X = 12$ (see Table 2.6) which is very small[7] and suggests that individuals in this scenario are making very risk averse decisions.

Our model also predicts that the value of $\tilde{\gamma}$ should change over time, with movements in realized and expected interest rates. If we assume that the true value of γ stays constant, than this explains why we see the share of fixed rate loans changing over time for new refinancers: the inequality $\tilde{\gamma} < \gamma$ will hold for more or fewer individuals. The graphs on the left hand side of Figure 2.8 plot the average value of $\tilde{\gamma}$ against different values of $\bar{m}_t$ and v_t^2 over the sample period. The risk aversion threshold falls as the expected change in interest rates increases, and as the expected volatility of interest rates increases. The right hand graph plots the predicted share (in blue) choosing fixed rates using the estimated $\hat{\gamma}$. This of course moves in close tandem with the average value of $\tilde{\gamma}$, and also moves closely with the observed share choosing fixed (in red), which is a comforting check of the model fit.

Calculation of Switching Costs:

Now that we have estimated the value of γ that rationalizes the active choices of all incoming borrowers, we can use this model to specify who should have eventually switched rate types as the interest rate environment changed, and the size of the switching costs implied by the behavior of individuals who should have switched rates but did not. I use the subset of individuals (N = 9,779) who I observe at least twice, once making an active decision and once six months later. On *average*, the LIBOR rate increase by 18 basis points for this group over six months. However, since these individuals refinanced at different points in time, they individually experienced anywhere from a 2 to 35 basis point change in the actual LIBOR rate, and a 3 to 40 bp change in the future LIBOR rate.

Using the value of $\hat{\gamma}$ that I estimated on individuals' initial choices, I calculate the new optimal rate choice for each individual given the change in interest rates they experienced over this time period. Table 2.7 shows that the model predicts that about 36% of the individuals who originally chose a variable rate should have switched to a fixed rate, but that only .16% actually switched. For individuals who should have, but did not switch, we can calculate a lower bound on the switching cost implied by this sticky behavior. Specifically, the lower bound on the cost μ required for individuals originally in a variable rate to stay in a variable

[7] For reference, Handel 2013 finds X is \$94.6.

rate is given by:

$$\mu > (T-2) * (\bar{m}_t + \frac{\gamma * v_t^2}{2}) - (MP_t^F - MP_t^V)$$

On average, these switching costs are about \$166 per individual (the median is \$84), which equates to about 4.7% of the payments that the individual is making on the loan during that period. These are relatively moderate switching costs, but when interpreted as a *lower bound* they are substantial. If rates continue to increase, and individuals still do not switch rates, then this will increase the lower bound. As we eventually observe more individuals switching empirically in the data this will allow us to actually estimate a point estimate, rather than a lower bound, of μ. Currently, using the few switchers I do observe in the data, I calculate the average "upper bound" on μ implied by their behavior. Specifically, if an individual switches from variable to fixed rates it means that:

$$\mu < (T-2) * (\bar{m}_t + \frac{\gamma * v_t^2}{2}) - (MP_t^F - MP_t^V)$$

For the few individuals I do observe switching, I calculate an average upper bound on μ of \$1,185 (which equates to 25% of that period's payment), and a median upper bound of \$202.

2.6 Conclusion

Previous work on health insurance choice, pension plan decisions, and the mortgage market has shown that consumers rarely update their initial contract choices to reflect changes in health risk or prevailing market conditions. I extend this literature to the online personal loan market, where borrowers are given the option to switch rate types (fixed or variable) at any point, yet exhibit considerable inertia in updating their contract choices despite changes in interest rates. This analysis is especially relevant for this online lending context, since one key differentiating factor for fin-tech firms is offering more flexible and personalized financial products; however, little is known about how consumers will interact with these more complete, but also more complex, contracts.

I use a panel of borrowers making both active and inactive interest rate decisions over a time period when the market index interest rate changed considerably. An analysis of fixed and variable rate choices shows that incoming borrowers making active rate decisions are very sensitive to the prevailing interest rate environment. A belief that rates will rise in the near future, or uncertainty surrounding these beliefs, causes a large increase in the share of borrowers choosing a fixed rate loan. However, despite the absence of any refinancing penalties or change fees, existing borrowers who have already chosen a variable rate are very unlikely to update their rate choice, which suggests that there are large switching costs. The analysis also reveals interesting heterogeneity in preferences across borrowers – namely, borrowers who are lower income and have higher observable risk types are more likely to choose variable rates and less likely to switch rates.

I next write down and estimate a tractable economic model that relates how both initial and inactive rate decisions relate to factors like risk aversion and interest rate risk, and allow me to estimate the dollar switching cost implied by borrower inaction. In the model, borrowers have CARA utility, a given loan amount, loan maturity, and face a risk-score dependent fixed and variable interest rates. When choosing between the variable or fixed rate, borrowers tradeoff the certain monthly payment savings they will experience today under the variable rate contract against the uncertain possibility that next period their fixed rate option will increase. When estimated, this models interprets the sensitivity of borrowers' initial rate choices as a very high level of risk aversion; Specifically, I estimate a risk aversion parameter of .0564, which implies borrowers' are indifferent between (i) inaction and (ii) accepting a gamble with a 50 percent chance of gaining $100 and a 50 percent chance of losing $X is $12. When the CARA is estimated as a function of observable characteristics, the model interprets the tendency of higher income, lower risk borrowers to choose a fixed rate loan as them being more risk averse. I then use the subsequent decisions of the same borrowers to update or stay with their current rate to lock down a lower and upper bound on switching costs. I estimate an average lower bound (identified off of individuals who should switch but do not) of $166, and an average upper bound (identified off of individuals who do switch) of $1,185. Each period I observe both new and old borrowers making these rate decisions under the new prevailing interest rate environment. This allows for an additional identification opportunity, since one can compare the active and inactive rate decisions cross-sectionally over borrowers at any given point in time, as well as within borrowers over time.

In future work, it would be interesting to understand how switching costs not only impact consumer surplus, but also producer surplus. Do lenders anticipate inertia on the borrower's part and is it priced into their offered contracts? As more individuals eventually switch rate types, it will allow us to also estimate μ as a function of observable characteristics, and explore potential correlations between the prices set by firms and estimated switching costs.

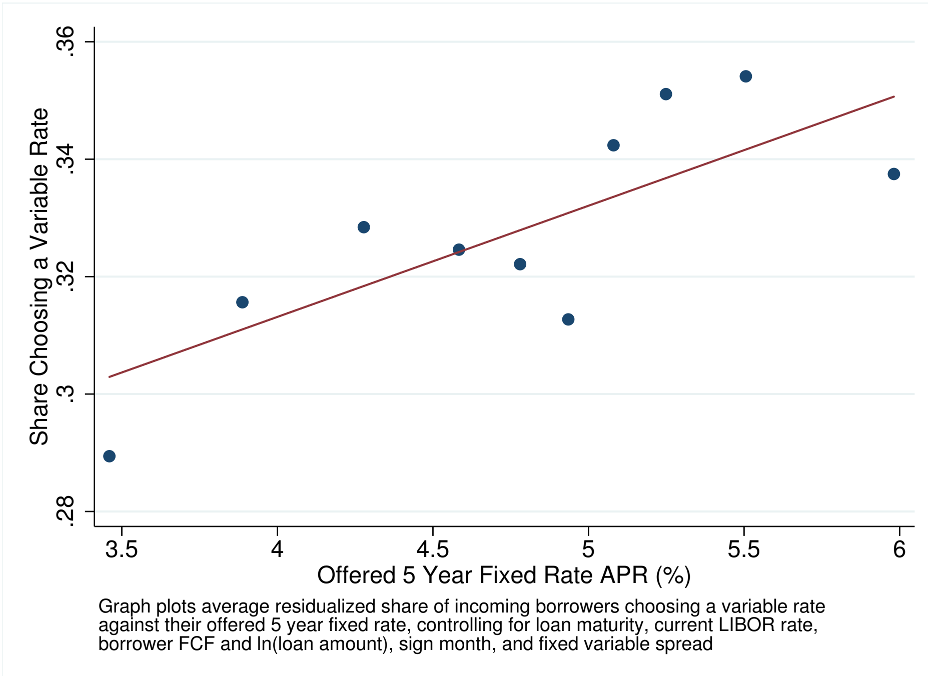


Figure 2.1: Relationship Between Variable Choice and Risk Rating

The choice of rate type is highly correlated with an individual's "risk rating" - those who have a higher risk, and thus face higher prices, are much more likely to choose variable. This graph illustrates this relationship by plotting the offered five year fixed rate against the share of individuals choosing a variable rate loan. Individuals with a higher risk score are offered higher interest rates, and are more likely to choose a variable rate.

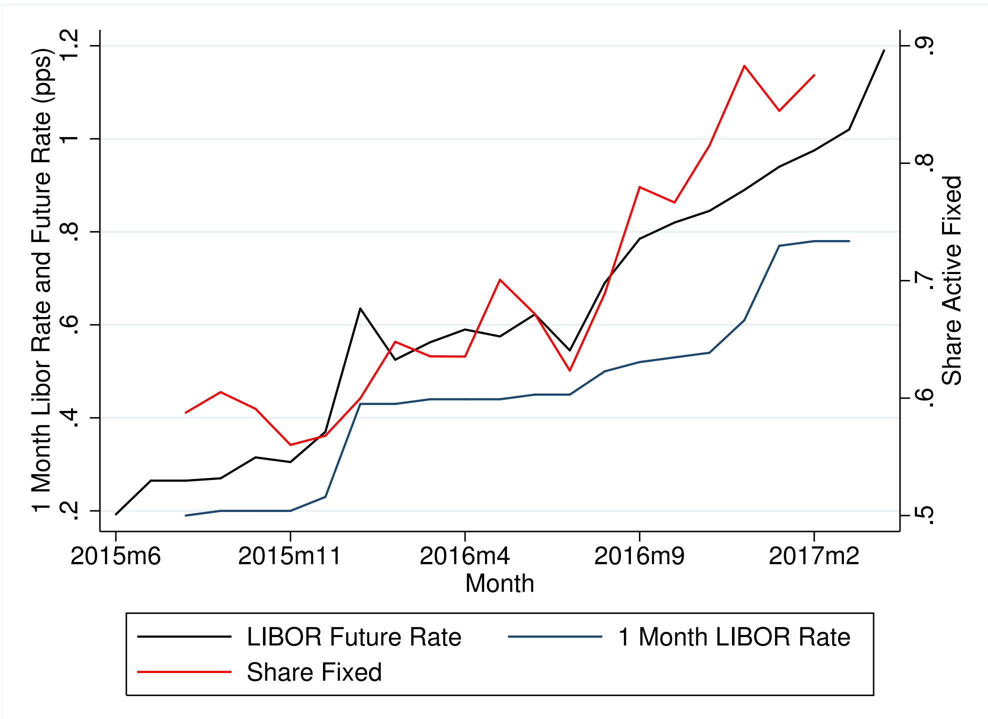


Figure 2.2: Changes in Expected Interest Rates: 90 Day LIBOR Future Rate and Share choosing Fixed Rate

The Eurodollar futures rate (plotted in black) is highly predictive of the *current* period's share choosing a fixed rate loan (plotted in red). This suggests that expectations about future interest rate movements do matter for borrowers' initial rate decisions.

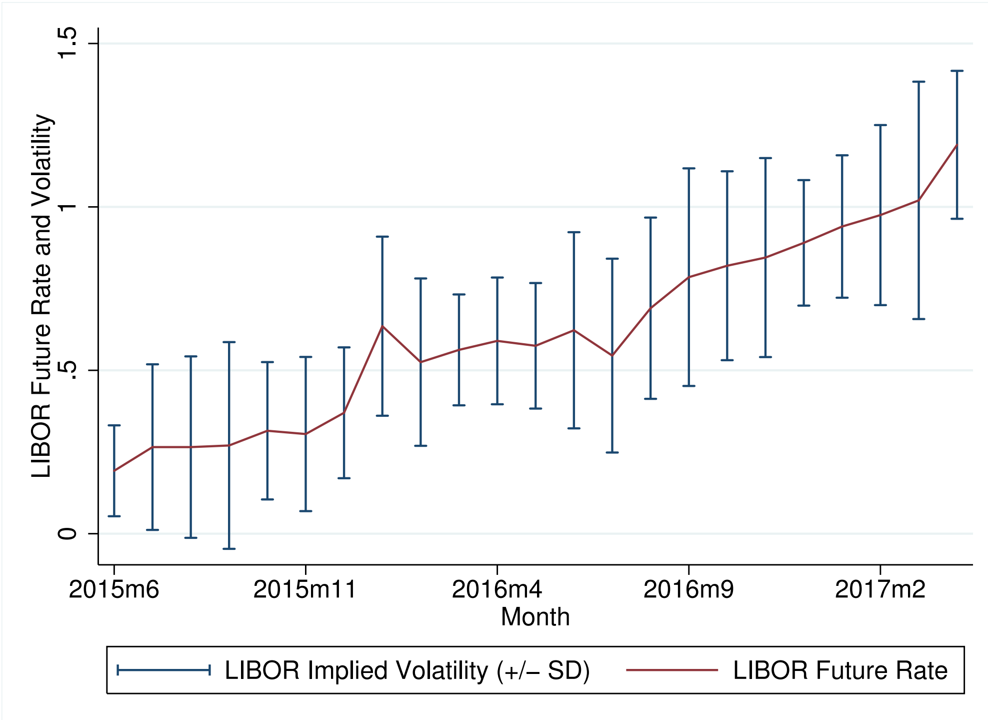


Figure 2.3: 90 Day LIBOR Future Rate and Implied Volatility

The 90-day Eurodollar future rate is plotted here over the sample period along with the bounds that represent the market's implied volatility measure. The implied volatility represents how far the market believes the rate could change over one standard deviation; for example, a $10 stock with a 20 percent implied volatility has a 68 percent chance to be priced between $8 and $12 one year from now. If there is more uncertainty about the path of interest rates, there will be a higher implied volatility.

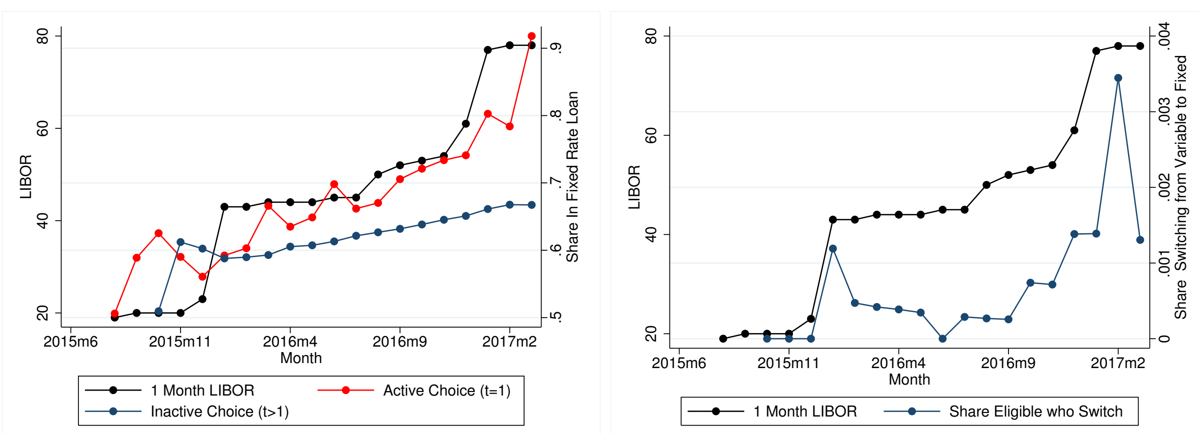


Figure 2.4: Fixed Rate Choices and LIBOR rate over the Sample

Panel (a) shows the movement of the 1-month LIBOR, the share of new loans that are fixed rate (borrowers making active rate choices), and the share of old loans that are fixed rate over the sample period. Panel (b) plots the share of eligible borrowers (individuals with a variable rate loan) that moved to a fixed rate over the same time period.

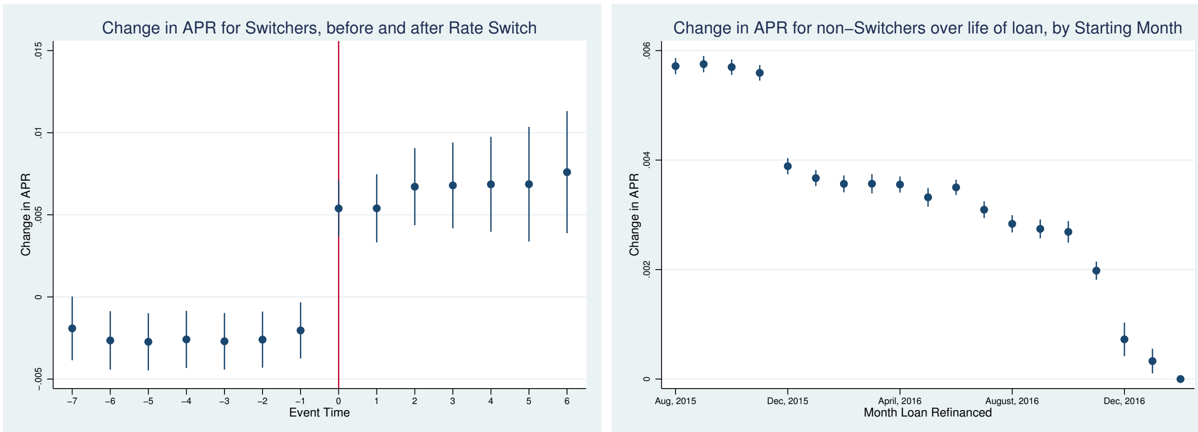


Figure 2.5: Change in APR for Switchers and Non-Switchers
Panel (a) shows the average change in APR that borrowers face when switching from a variable to a fixed rate loan. The "event" illustrated in the plot is the time of switching rates. Panel (b) the average change in APR that borrowers who did *not* switch have experienced over the life of their loan. Their interest rate instead has gradually increased with the index rate (LIBOR 1-month) on their loan.

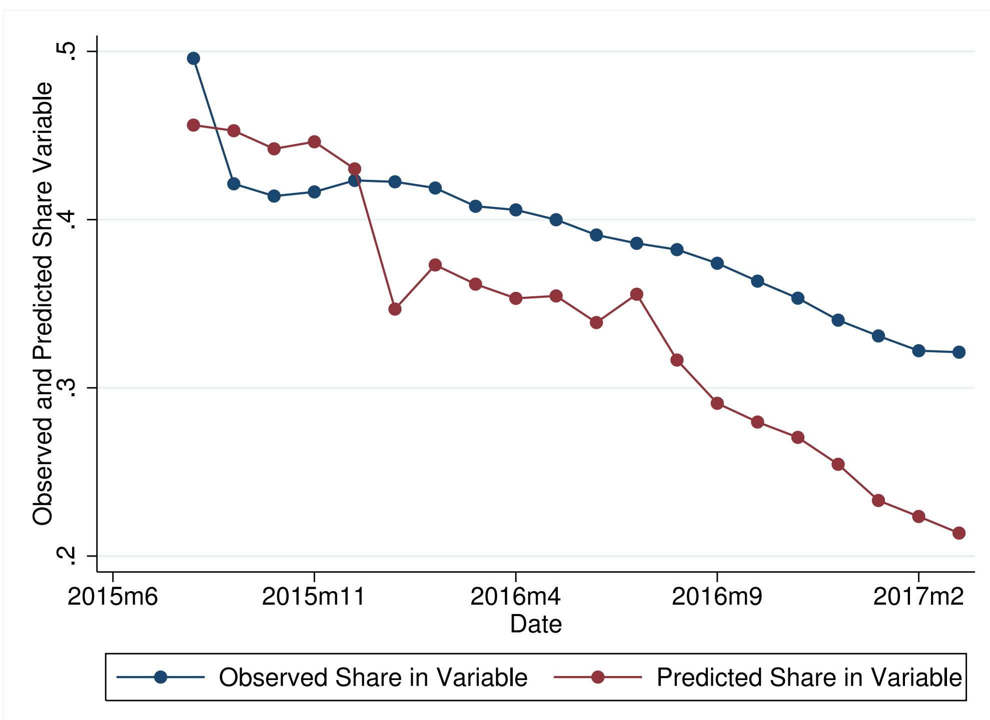


Figure 2.6: Observed vs. Predicted Variable Share, over the Sample Period

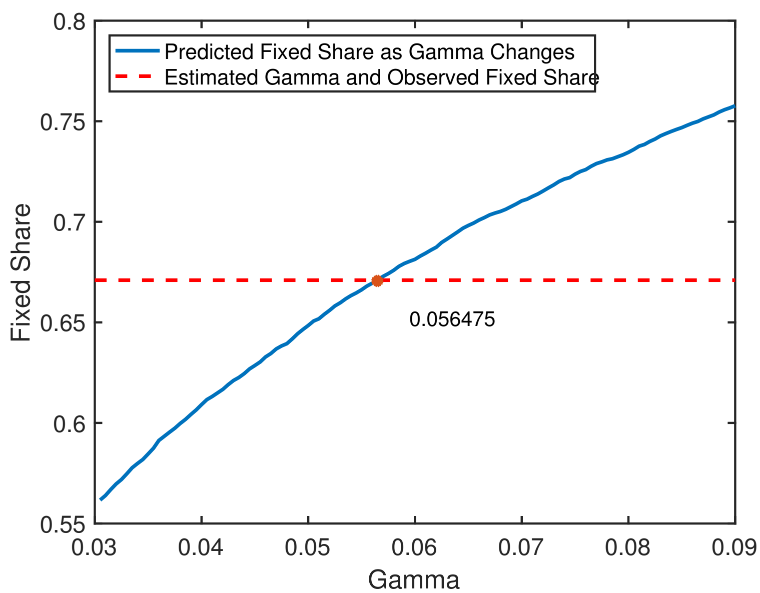


Figure 2.7: γ Fixed Share Relationship, and estimated $\hat{\gamma}$

This graph plots the logic behind estimation of $\hat{\gamma}$. As $\hat{\gamma}$ increases, the model predicts more individuals should choose a fixed rate (solid line). Therefore, the estimation procedure chooses the value of $\hat{\gamma}$ to match the predicted to the observed share.

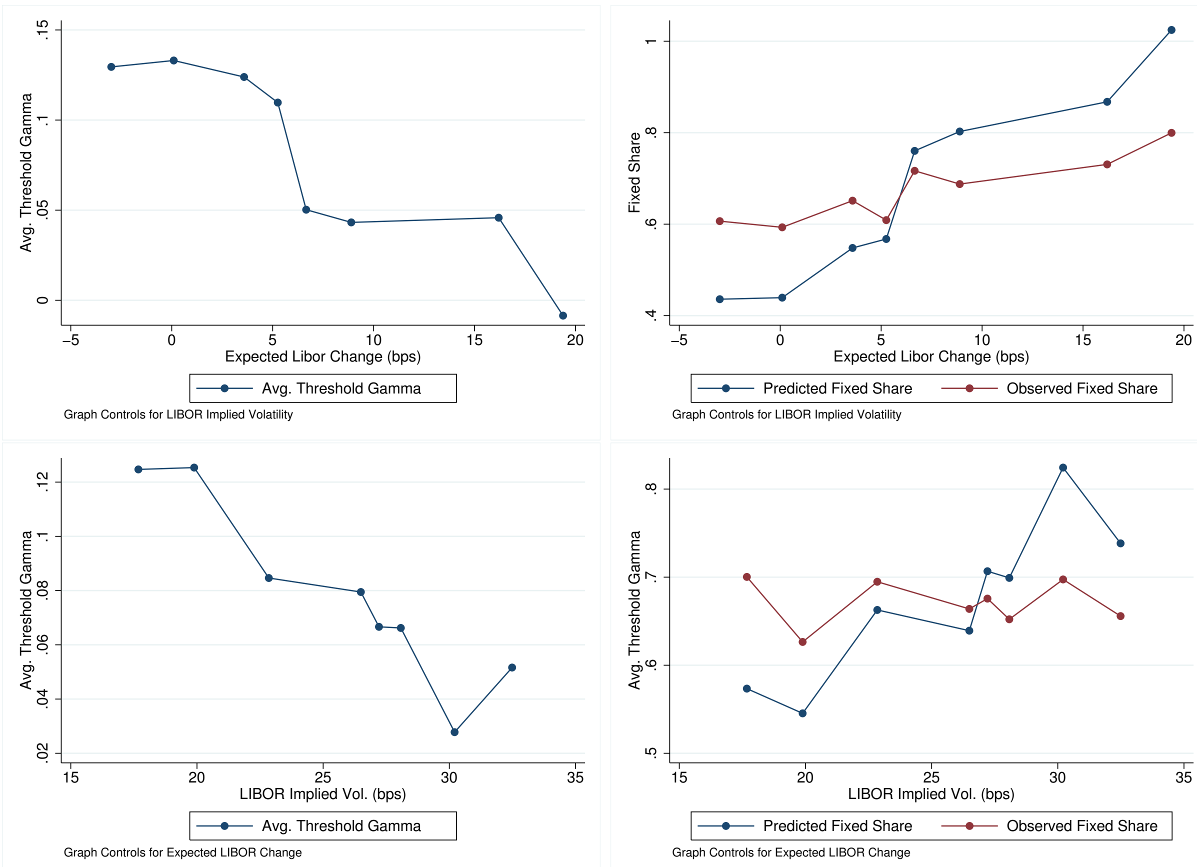


Figure 2.8: Threshold γ and Predicted vs. Observed Shares
The left hand graphs plot the changes in the average value of $\tilde{\gamma}$ against expected changes and volatility of the future LIBOR rate. If the population has a true value of γ, than as $\tilde{\gamma}$ increases, it will be less likely that $\gamma > \tilde{\gamma}$, and therefore less likely that we observe individuals choosing a fixed rate loan. The right hand graph plots the predicted share of individuals choosing a fixed rate using this rule – it moves in almost perfect tandem with the average value of $\tilde{\gamma}$.

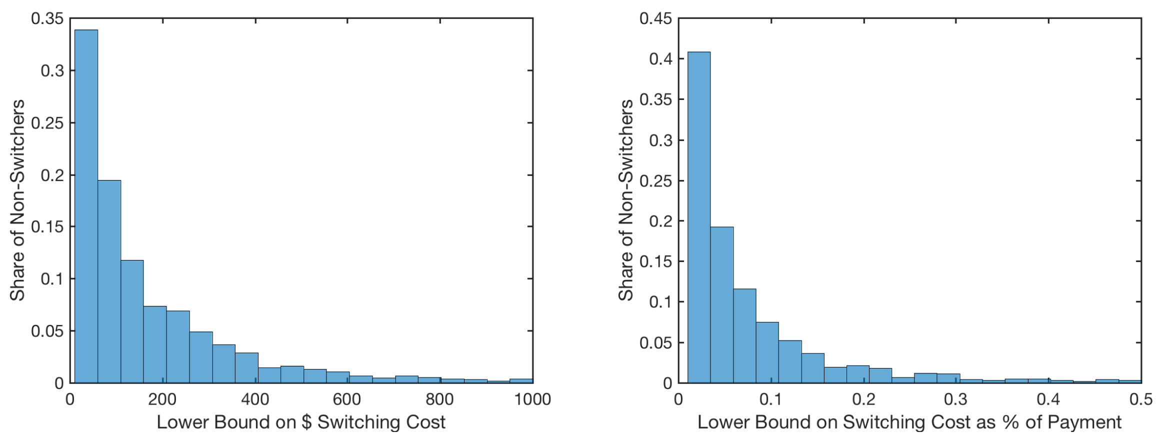


Figure 2.9: Distribution of Lower Bound on Switching Costs

These histograms show the distribution of the *lower bound* on implied switching costs for individuals who the model suggests should have switched from a variable to a fixed rate loan, but did not. On the left, the costs are given in dollar terms, and on the right they are expressed as a percentage of the loan payment over the same period.

Table 2.1: Summary Statistics on Dataset Structure

Cohort Start Date	Cohort Size	Avg. Panel Length	Initial LIBOR	Δ LIBOR over panel
Aug-15	895	17.2	19.9	54.7
Sep-15	767	16.6	20.2	55.6
Oct-15	954	15.4	20.3	54.7
Nov-15	889	14.7	23.1	52.7
Dec-15	681	13.7	40.4	35.4
Jan-16	888	12.7	43	33.2
Feb-16	813	11.9	44	33
Mar-16	493	11.1	44.1	32.7
Apr-16	950	10	44	32.9
May-16	635	8.9	44.9	32
Jun-16	1308	8.1	45.5	31.8
Jul-16	700	7	49.6	27.8
Aug-16	814	6.1	51.9	25.6
Sep-16	898	5.2	53.2	24.6
Oct-16	527	4.1	54	23.8
Nov-16	1192	3.3	61.6	16.3
Dec-16	332	2.4	74.4	3.6
Jan-17	559	1.2	77.9	0.1
Feb-17	8	1.2	78	0
Total	14303	9.8	43.5	33.3

Table 2.2: Sample Monthly Payments for a 10-year, $50,000 loan at variable and fixed rates

	+ Current LIBOR rate (.20%)	+ 2005 LIBOR rate (3.39%)	+ 1995 LIBOR rate (5.97%)
Variable rate	2.68%	5.87%	8.45%
Variable monthly payment	$475	$552	$619
Fixed rate	4.61%	4.61%	4.61%
Fixed monthly payment	$521	$521	$521

Table 2.3: Descriptive Statistics by Loan Rate Type

Rate Type	FICO	Loan Amount	Monthly FCF	Scheduled MP	Maturity	APR
Fixed *(Mean)*	784.07	71403.15	2271.74	836.50	108.80	5.07
(SE)	1.27	639.43	34.64	8.06	0.53	0.01
(Median)	786.00	51272.96	2506.55	592.16	95.00	5.08
Variable *(Mean)*	778.54	71645.98	3097.52	807.35	109.06	3.97
(SE)	0.53	886.58	48.99	10.95	0.77	0.01
(Median)	782.00	52099.88	2882.22	566.59	91.00	3.95
Total *(Mean)*	782.23	71483.69	2545.76	826.83	108.89	4.70
(SE)	0.87	518.73	28.47	6.49	0.44	0.01
(Median)	785.00	51526.50	2630.40	582.32	93.00	4.80

Table 2.4: Descriptive Statistics: Variable Rate Borrowers by Switching Status

Ever Switch?		Loan Amount	FCF	Maturity	Risk Score	Age
No	*(Mean)*	71,637.54	3,101.45	108.83	5.23	33.26
	(SE)	893.61	49.36	0.78	0.02	0.10
	(Median)	52,135.44	2,883.69	91.00	5.21	32.00
Yes	*(Mean)*	78,826.98	3,010.55	125.84	5.99	34.17
	(SE)	7,332.94	404.93	7.23	0.13	0.92
	(Median)	54,346.19	2,918.98	109.00	6.16	33.00
Overall	*(Mean)*	71,734.51	3,100.22	109.06	5.24	33.27
	(SE)	887.08	49.00	0.77	0.02	0.10
	(Median)	52,142.60	2,883.69	91.00	5.22	32.00

Table 2.5: Logit Model for Determinants of Fixed/Variable Rate Decision

	(1) Fixed Rate, Active Choice		(2) Ever Switch Var to Fixed	
	β / SE	Mfx %x/%y	β / SE	Mfx %x/%y
Maturity (Mo)	−0.001***	−0.036***	0.005	0.524
	(0.000)		(0.003)	
ln(Income)	−0.344***	−1.266***	−0.574***	−6.586***
	(0.051)		(0.209)	
ln(Amount)	0.092***	0.320***	0.106	1.148
	(0.027)		(0.271)	
Current LIBOR	−0.000	−0.006	−0.109***	−4.393***
	(0.005)		(0.014)	
LIBOR future	1.920***	0.288***		
	(0.488)			
FV Spread	−0.003***	−0.159***	−0.061***	−9.827***
	(0.001)		(0.008)	
Risk Score	0.081***	0.137***	0.393**	2.056**
	(0.020)		(0.188)	
Age	0.004	0.041	0.028	0.943
	(0.003)		(0.028)	
# Mortgages	0.232***	0.030***	−0.092	−0.033
	(0.035)		(0.219)	
Observations	13618		4471	
Pseudo R^2	0.032		0.390	
Baseline predicted probability	0.701		0.002	

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

Table 2.6: Estimate Values of γ and Monetary Interpretation

	Single γ	γ as Function of Obs.
Avg. $\hat{\gamma}$	0.056475	0.05803
Cons.	-	0.0067789
Risk Score	-	0.0021655
log(FCF)	-	0.0014064
log(Loan Amt)	-	0.0026458
\$X	12.24	-
$\$X_{Max}$	-	15.74
$\$X_{Min}$	-	10.01

Table 2.7: Predicted Share of Switchers and Implied Switching Costs

	Predicted, No Inertia	Observed, With Inertia
Initial Fixed % (t=0)	51.59	61.97
Inactive Fixed % (t=6)	69.22	62.03
Share of Var to Fixed %	36.42	0.16
Avg. LB Switching Cost $	0	166.9
Avg. UB Switching Cost $	0	1185.6

Bibliography

[1] William Adams, Liran Einav, and Jonathan Levin. "Liquidity constraints and imperfect information in subprime lending". In: *The American Economic Review* 99.1 (2009), pp. 49–84.

[2] Sumit Agarwal, Itzhak Ben-David, and Vincent Yao. "Systematic mistakes in the mortgage market and lack of financial sophistication". In: *Journal of Financial Economics* 123.1 (2017), pp. 42–58.

[3] Sumit Agarwal, John C Driscoll, and David I Laibson. "Optimal Mortgage Refinancing: A Closed-Form Solution". In: *Journal of Money, Credit and Banking* 45.4 (2013), pp. 591–622.

[4] Sumit Agarwal, Richard J Rosen, and Vincent Yao. "Why Do Borrowers Make Mortgage Refinancing Mistakes?" In: *Management Science* 62.12 (2015), pp. 3494–3509.

[5] Sumit Agarwal et al. *Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinancing Program*. Tech. rep. National Bureau of Economic Research, 2015.

[6] Steffen Andersen et al. *Inattention and inertia in household finance: Evidence from the Danish mortgage market*. Tech. rep. National Bureau of Economic Research, 2015.

[7] Orazio P Attanasio, Pinelopi Koujianou Goldberg, and Ekaterini Kyriazidou. "Credit constraints in the market for consumer durables: Evidence from micro data on car loans". In: *International Economic Review* 49.2 (2008), pp. 401–436.

[8] Orazio P Attanasio and Guglielmo Weber. "Consumption and saving: models of intertemporal allocation and their implications for public policy". In: *Journal of Economic literature* 48.3 (2010), pp. 693–751.

[9] Christopher Avery and Sarah Turner. "Student loans: Do college students borrow too much?or not enough?" In: *The Journal of Economic Perspectives* 26.1 (2012), pp. 165–192.

[10] Cristian Badarinza, John Y Campbell, and Tarun Ramadorai. "What calls to ARMs? International evidence on interest rates and the choice of adjustable-rate mortgages". In: *Management Science* (2017).

[11] Michael Carlos Best et al. *Interest Rates, Debt and Intertemporal Allocation: Evidence From Notched Mortgage Contracts in the UK*. Tech. rep. Working Paper, 2015.

[12] Harald Beyer et al. "Connecting Student Loans to Labor Market Outcomes: Policy Lessons from Chile". In: *American Economic Review* 105.5 (2015), pp. 508–13. DOI: 10.1257/aer.p20151026.

[13] Brian Bucks and Karen Pence. "Do borrowers know their mortgage terms?" In: *Journal of urban Economics* 64.2 (2008), pp. 218–233.

[14] M Kate Bundorf, Jonathan Levin, and Neale Mahoney. "Pricing and welfare in health plan choice". In: *The American Economic Review* 102.7 (2012), pp. 3214–3248.

[15] Paul S Calem and Loretta J Mester. "Consumer behavior and the stickiness of credit-card interest rates". In: *The American Economic Review* 85.5 (1995), pp. 1327–1336.

[16] John Y Campbell and Joao F Cocco. "Household risk management and optimal mortgage choice". In: *The Quarterly Journal of Economics* 118.4 (2003), pp. 1449–1494.

[17] Caroline Carlin and Robert Town. "Adverse selection, welfare and optimal pricing of employer-sponsored health plans". In: *U. Minnesota Working Paper* (2009).

[18] Raj Chetty and Amy Finkelstein. *Social insurance: Connecting theory to data*. Tech. rep. National Bureau of Economic Research, 2012.

[19] Alma Cohen and Liran Einav. "Estimating risk preferences from deductible choice". In: *The American economic review* 97.3 (2007), pp. 745–788.

[20] James C Cooper et al. "Does price discrimination intensify competition? Implications for antitrust". In: *Antitrust Law Journal* 72.2 (2005), pp. 327–373.

[21] Susan Dynarski and Judith Scott-Clayton. *Financial aid policy: Lessons from research*. Tech. rep. National Bureau of Economic Research, 2013.

[22] Wendy Edelberg. "Risk-based pricing of interest rates for consumer loans". In: *Journal of Monetary Economics* 53.8 (2006), pp. 2283–2298.

[23] Liran Einav and Amy Finkelstein. "Selection in insurance markets: Theory and empirics in pictures". In: *The Journal of Economic Perspectives* 25.1 (2011), pp. 115–138.

[24] Liran Einav, Amy Finkelstein, and Mark R Cullen. *Estimating welfare in insurance markets using variation in prices*. Tech. rep. National Bureau of Economic Research, 2008.

[25] Liran Einav, Amy Finkelstein, and Jonathan Levin. *Beyond testing: Empirical models of insurance markets*. Tech. rep. National Bureau of Economic Research, 2009.

[26] Liran Einav, Amy Finkelstein, and Paul Schrimpf. "Optimal mandates and the welfare cost of asymmetric information: Evidence from the uk annuity market". In: *Econometrica* 78.3 (2010), pp. 1031–1092.

[27] Liran Einav, Mark Jenkins, and Jonathan Levin. "Contract pricing in consumer credit markets". In: *Econometrica* 80.4 (2012), pp. 1387–1432.

[28] Liran Einav, Mark Jenkins, and Jonathan Levin. "The impact of credit scoring on consumer lending". In: *The RAND Journal of Economics* 44.2 (2013), pp. 249–274.

[29] Amy Finkelstein and James Poterba. *Testing for Adverse Selection with" unused Observables"*. National Bureau of Economic Research, 2006.

[30] David B Gross and Nicholas S Souleles. *Do liquidity constraints and interest rates matter for consumer behavior? Evidence from credit card data*. Tech. rep. National bureau of economic research, 2001.

[31] Jonathan Gruber. "A tax-based estimate of the elasticity of intertemporal substitution". In: *The Quarterly Journal of Finance* 3.01 (2013), p. 1350001.

[32] Ben Handel, Igal Hendel, and Michael D Whinston. "Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk". In: *Econometrica* 83.4 (2015), pp. 1261–1313.

[33] Benjamin R Handel. "Adverse selection and inertia in health insurance markets: When nudging hurts". In: *The American Economic Review* 103.7 (2013), pp. 2643–2682.

[34] Andrew Hertzberg, Andres Liberman, and Daniel Paravisini. *Adverse Selection and Maturity Choice in Consumer Credit Markets: Evidence from an Online Lender*. Tech. rep. Working Paper, 2015.

[35] Erik Hurst et al. *Regional redistribution through the US mortgage market*. Tech. rep. National Bureau of Economic Research, 2015.

[36] Matteo Iacoviello. "House prices, borrowing constraints, and monetary policy in the business cycle". In: *The American economic review* 95.3 (2005), pp. 739–764.

[37] Gastón Illanes. *Switching Costs in Pension Plan Choice*. Tech. rep. Working Paper, 2016.

[38] David S Johnson, Jonathan A Parker, and Nicholas S Souleles. "Household expenditure and the income tax rebates of 2001". In: *The American Economic Review* 96.5 (2006), pp. 1589–1610.

[39] Kathleen W Johnson and Geng Li. "Are Adjustable-Rate Mortgage Borrowers Borrowing Constrained?" In: *Real Estate Economics* 42.2 (2014), pp. 457–471.

[40] Dean Karlan and Jonathan Zinman. "Expanding credit access: Using randomized supply decisions to estimate the impacts". In: *Review of Financial studies* (2009), hhp092.

[41] Dean Karlan and Jonathan Zinman. *Long-run price elasticities of demand for credit: evidence from a countrywide field experiment in Mexico*. Tech. rep. National Bureau of Economic Research, 2013.

[42] Dean S Karlan and Jonathan Zinman. "Elasticities of demand for consumer credit". In: *Yale University Economic Growth Center Discussion Paper* 926 (2005).

[43] Benjamin J Keys, Devin G Pope, and Jaren C Pope. "Failure to refinance". In: *Journal of Financial Economics* 122.3 (2016), pp. 482–499.

[44] Benjamin J Keys et al. *Mortgage rates, household balance sheets, and the real economy.* Tech. rep. National Bureau of Economic Research, 2014.

[45] Lance Lochner and Alexander Monge-Naranjo. *Credit constraints in education.* Tech. rep. National Bureau of Economic Research, 2011.

[46] Lance Lochner and Alexander Monge-Naranjo. *Student Loans and Repayment: Theory, Evidence and Policy.* Tech. rep. National Bureau of Economic Research, 2015.

[47] Lance J Lochner and Alexander Monge-Naranjo. "The nature of credit constraints and human capital". In: *The American Economic Review* 101.6 (2011), pp. 2487–2529.

[48] Adam Looney and Constantine Yannelis. "A Crisis in Student Loans?: How Changes in the Characteristics of Borrowers and in the Institutions They Attended Contributed to Rising Loan Defaults". In: *Brookings Papers on Economic Activity* 2015.2 (2015), pp. 1–89.

[49] Deborah Lucas. *Credit policy as fiscal policy.* 2012.

[50] Deborah Lucas. "Introduction to" Measuring and Managing Federal Financial Risk"". In: *Measuring and Managing Federal Financial Risk.* University of Chicago Press, 2010, pp. 1–12.

[51] Deborah Lucas. "Valuation of government policies and projects". In: *Annu. Rev. Financ. Econ.* 4.1 (2012), pp. 39–58.

[52] Deborah Lucas and Damien Moore. "Guaranteed versus direct lending: the case of student loans". In: *Measuring and managing federal financial risk.* University of Chicago Press, 2007, pp. 163–205.

[53] Nuno C Martins and Ernesto Villanueva. "The impact of mortgage interest-rate subsidies on household borrowing". In: *Journal of Public Economics* 90.8 (2006), pp. 1601–1623.

[54] Daniel Paravisini and Antoinette Schoar. "The Incentive Effect of IT: Randomized Evidence from Credit Committees". In: *NBER Working Paper Series* (2013), p. 19303.

[55] Jesse Rothstein and Cecilia Elena Rouse. "Constrained after college: Student loans and early-career occupational choices". In: *Journal of Public Economics* 95.1 (2011), pp. 149–163.

[56] J Brad Schwartz. "Student financial aid and the college enrollment decision: The effects of public and private grants and interest subsidies". In: *Economics of Education Review* 4.2 (1985), pp. 129–144.

[57] Donald J Smith. "The borrower's choice between fixed and adjustable rate loan contracts". In: *Real Estate Economics* 15.2 (1987), pp. 110–116.

[58] Nicholas Wonder, Wendy Wilhelm, and David Fewings. "The financial rationality of consumer loan choices: Revealed preferences concerning interest rates, down payments, contract length, and rebates". In: *Journal of Consumer Affairs* 42.2 (2008), pp. 243–270.

[59] Jonathan Zinman. *Household debt: Facts, puzzles, theories, and policies*. Tech. rep. National Bureau of Economic Research, 2014.