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

[Job Market Paper Draft. Latest Version Available at nmcox.weebly.com.]

Natalie Cox *

January 12, 2017

Abstract

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 this paper, 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 changes 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 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.

*Department of Economics, UC Berkeley, Evans Hall, Berkeley, CA 94720 (email: nmcox@berkeley.edu). I thank my dissertation committee chairs Ben Handel and David Sraer for their guidance on this project, and my committee members, Ulrike Malmendier and Emmanuel Saez, who provided invaluable advice throughout the research process. This work has benefited particularly from the comments of Alan Auerbach, Pierre Bachas, Zarek Brot-Goldberg, Claire Celerier, Kei Kawai, Patrick Kline, Jon Kolstad, Jennifer Kwok, Julien Lafortune, Waldo Ojeda, Jon Schellenberg, Avner Schlain, Katalin Springel, Boris Vallée, Danny Yagan, Gabriel Zucman and seminar participants at UC Berkeley. I also thank the patient and helpful employees at the firm whose data I analyze.
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 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 of the remaining federal borrowers will rise, forcing the government to either raise its uniform rate or subsidize the remaining pool.

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 changes, 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,

---

\(^1\)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

\(^3\)Maturity is the relevant response for student borrowers in repayment, given that their loan principal is already fixed.
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. 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 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.

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"
I leverage the exogenous 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

5I 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 almost twenty private
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. 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

firms and 10 state-run refinancing programs
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.
Setting and Data:

2.1 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), 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 complete 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 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).

6The percentage of loans in delinquency 3 years after entering repayment
2.1.1 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. A description of the federal repayment plan options currently available is provided in the Appendix. 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.7

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 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

---

7I discuss the welfare implications of limiting term contracts (for example to only a 10 year fixed term) in the Appendix.
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.

2.1.2 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. The risk scores in my dataset are typical of “data-driven” underwriting, and while a function of thousands of data points, are primarily driven by free cash flow, degree, income, savings, FICO, and debt.

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). 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.

<table>
<thead>
<tr>
<th>Loan Type</th>
<th>Borrower Type</th>
<th>Index</th>
<th>Add-on</th>
<th>Fixed Interest Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Un/Subsidized Loans</td>
<td>Undergraduate</td>
<td>10 Yr Tr</td>
<td>+ 2.05%</td>
<td>3.4-4.66%</td>
</tr>
<tr>
<td>Direct Unsubsidized Loans</td>
<td>Graduate/Professional</td>
<td>10 Yr Tr</td>
<td>+ 3.60%</td>
<td>5.41-6.8%</td>
</tr>
<tr>
<td>Direct PLUS Loans</td>
<td>Parents&amp; Graduate</td>
<td>10 Yr Tr</td>
<td>+ 4.60%</td>
<td>6.4-7.9%</td>
</tr>
</tbody>
</table>


Table 1: 2011-2015 Interest Rates on Federal Direct Student Loans

2.2 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

---

8Legally 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.

9These variables explain over 85% of the risk score.
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.

2.2.1 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 median borrower owing over $50,000\(^{10}\). 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 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 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 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.

\(^{10}\) 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.
### Borrower Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>75,879</td>
<td>68,304</td>
<td>39,799</td>
</tr>
<tr>
<td>Loan Amt</td>
<td>67,078</td>
<td>50,656</td>
<td>52,890</td>
</tr>
<tr>
<td>FCF</td>
<td>3,636</td>
<td>3,100</td>
<td>2,574</td>
</tr>
<tr>
<td>FICO</td>
<td>782</td>
<td>787</td>
<td>36</td>
</tr>
<tr>
<td>Mortgages</td>
<td>0.40</td>
<td>0.00</td>
<td>0.60</td>
</tr>
<tr>
<td>Graduate</td>
<td>0.70</td>
<td>1.00</td>
<td>0.50</td>
</tr>
<tr>
<td>Age</td>
<td>32.60</td>
<td>31.00</td>
<td>6.80</td>
</tr>
<tr>
<td>Dependents</td>
<td>0.50</td>
<td>0.00</td>
<td>0.90</td>
</tr>
</tbody>
</table>

### Loan Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturity (Months)</td>
<td>106.8</td>
<td>93.0</td>
<td>49.7</td>
</tr>
<tr>
<td>Monthly Payment</td>
<td>799.9</td>
<td>600.0</td>
<td>624.8</td>
</tr>
<tr>
<td>APR</td>
<td>4.494%</td>
<td>4.64%</td>
<td>1.042%</td>
</tr>
<tr>
<td>Variable Rate</td>
<td>0.40</td>
<td>0.00</td>
<td>0.50</td>
</tr>
<tr>
<td>N</td>
<td>11663</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Borrower and Loan Summary Statistics

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 2: Distribution of Potential Refinanced Interest Rates
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 3: Comparison of Applicant Pool to Nationally Representative Sample of Graduate Student Borrowers

**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 than 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 monthly payment, APR, and total paid for the maturity they 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 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. Other borrowers seem to have specific

---

11This may also be somewhat of an “inertia” or “default” effect, since most borrowers are refinancing from
Debt Amount to Refinance: $80,000

Monthly Payment: $1,000

Interest Rate: 4.54%

Loan Term: 8.0

Lifetime Cost: $95,556

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.

Figure 4: User Interface for Term/Monthly Payment Selection

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.

2.2.2 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 within-risk. 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 unrelated to the characteristics of any given borrower. Figure 5 provides a graphical explanation of why within risk score price variation is necessary for identification, and how it can be used.

The exogenous 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 settings that offer only traditional term choices like 10 years, and may prefer to stay with their current term.
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.

Figure 5: Using Across vs Within Risk Price Variation to Identify Term Elasticities

To demonstrate the importance of using only the exogenous 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\(^\text{12}\) maturity choice on \( P_i \) (and observables), pooling both sources of variation:

\[
T_i = \alpha + \beta P_i + X_i'\mu + \epsilon_i
\]

The results (see column 1 of Table 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.

\(^{12}\)I use a tobit specification to account for the truncation of the choice set at 60 and 240 months.
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 3: Impact of Debt, Income, APR, and Risk on Term
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 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 4: Maturity Elasticities

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.\(^{13}\)

The elasticity that corresponds with this coefficient (see Table 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 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 than their low-risk counterparts’. Table 4 expresses these results in terms of elasticities – these show that the highest risk individuals are essentially inelastic to price changes, whereas

\(^{13}\)A higher score means an individual is less risky and faces a lower level of risk based prices.
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 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 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 6 shows that while the exogenous 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.

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 (note that I have plotted the negative maturity response in order to make the correlation between the two variables more clear visually). 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 6: Exogenous APR Variation and Maturity Responses

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

---

14I 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.
loans), they could confound our estimates of the impact of APR. Figure 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 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 exogenous 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 5 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, $T_i$, 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 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.

<table>
<thead>
<tr>
<th>Observables over Price Regimes</th>
<th>Avg. APR</th>
<th>Coeff.</th>
<th>SE</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Income)</td>
<td>-.0000367</td>
<td>.0000362</td>
<td>-1.01</td>
<td></td>
</tr>
<tr>
<td>ln(Debt)</td>
<td>-9.55e-06</td>
<td>.0000161</td>
<td>-0.59</td>
<td></td>
</tr>
<tr>
<td>ln(Savings)</td>
<td>.0000162</td>
<td>.0000168</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>Mortgage</td>
<td>1.52e-06</td>
<td>.0000221</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.24e-06</td>
<td>1.98e-06</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>F(5, 11663) = .99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response of $T$ to APR*</th>
<th>$T$</th>
<th>Coeff.</th>
<th>SE</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. APR</td>
<td>363.09</td>
<td>341.98</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>11663</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$T$ is the term predicted using all observables except price.

Table 5: Test of Extensive Margin Response and Changes in Borrower Composition

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.

3.1 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:

$$
\ln(w_{it}) = \ln(w_{i,t-1}) + u_{it}
$$

$$
u_{it} \sim N(\sigma_i^2)
$$

where $g_i$ is a yearly growth rate and $\sigma_i^2$ 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, $\beta$. 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.
3.2 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)\(^{15}\). Specifically, they solve:

$$\max_T E\left[\sum_{t=1}^T \beta^t u(w_t - d_t) + \sum_{T+1}^{20} \beta^t u(w_t)\right]$$

s.t. $d_t = D_t \cdot \frac{r(T,p_i)}{(1 - (1 + r(T,p_i))^{-T})}$

where $d_t$ 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:

$$0 = -E\left[\sum_0^T \beta^t \frac{\partial d}{\partial T} u'(w_t - d_t) + \beta^{T+1} u(w_{i(T+1)} - \frac{d_{i+1}}{2}) - \beta^{T+1} u(w_{i(T+1)})\right]$$

which can be rewritten as:

$$E\left[\sum_0^T \beta^t \frac{\partial d}{\partial T} u'(w_t - d_t)\right] = E[\beta^{T+1}(-d_t)u'(w_{i(T+1)} - \frac{d_{i+1}}{2})]$$

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.\(^{16}\)

3.2.1 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 $\frac{dT}{dT}$ does not have an analytical solution, but Figure 7 shows how simulated optimal term choices vary over interest rate levels – as the level of interest rates increases, the optimal maturity choice decreases.

\(^{15}\)In the public sector the interest rate $r(T,p_i)$ is replaced with $g$

\(^{16}\)To make this condition empirically tractable, I approximate the second term, $\beta^{T+1} u(w_{i(T+1)} - d_{i+1}) - \beta^{T+1} u(w_{i(T+1)})$, with the expression:

$$\beta^{T+1} u(w_{i(T+1)} - d_t) - \beta^{T+1} u(w_{i(T+1)}) \approx \beta^{T+1}(-d_t)u'(w_{i(T+1)} - \frac{d_{i+1}}{2})$$
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 7: Variation in Observable Characteristics over Time

3.2.2 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 $\gamma$ 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 $\gamma$ (ie a low intertemporal elasticity of substitution) also means that the term choices of individuals will be less responsive to price changes. Thus $\gamma$ is essential for understanding how a borrower’s decisions, and utility, would respond to changes in price, the
3.2.3 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.

3.2.4 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.

### 3.2.5 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 26 in the appendix describes changes in other monthly payments (mortgages, auto loans, credit cards, etc) before vs. after refinancing, 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 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.
3.2.6 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 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 an 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.

3.3 Estimation

Our reduced form evidence expressed maturity choice as a linear function of observables, risk type, and interest rate ($X_i$, $p_i$, and $r(T,p)_i$), 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 $r(T,p)_i$ derived from the borrower’s first order condition, and the price shifts now serve to identify $\gamma$. 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, $\gamma$, 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 $\gamma$, 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 $T$.

3.3.1 Empirical Framework:

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

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

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

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 $T_i$, $\frac{d_d}{dT_i}$, 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\(^\text{17}\) 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

$$u_{it} \sim N(0, \sigma_u^2)$$

$$\sigma_u^2 = (\omega - v \times p_i)^2$$

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 exogenous 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

\(^{17}\)In the robustness checks I relax this assumption and try other specifications.
estimation, the model would wrongly attribute differences in maturity choices across risk type
to differences in offered APR, and our estimate of $\gamma$ 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:

$$E[(w_{i0} * e^{(X_i^\mu)} * e^{\Sigma_i u_i})^{-\gamma}] = (w_{i0} * e^{(X_i^\mu)} * e^{\pi_i})^{-\gamma}$$

where $\pi_{it}$ 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\(^{18}\):

$$\pi_{it} = \frac{1}{2} * t * \sigma^2[1 - (1 + \gamma) \frac{w_{i0} * e^{(X_i^\mu)}}{w_{i0} * e^{(X_i^\mu)} - d_i}]$$

for $t < T+1$

$$\pi_{it} = \frac{1}{2} * t * \sigma^2(-\gamma)$$

for $t \geq T+1$

One can see that as income volatility, risk aversion, and the debt to income ratio ($w_{i0} * e^{(X_i^\mu)} - d_i$) ratio increases, the certainty equivalent becomes more negative.

Using this expression, our analytical estimating moment becomes:

$$g_i(\theta) = \sum_{1}^{T} \beta_i \frac{\partial d}{\partial T}(w_{i0} * e^{(X_i^\mu)} * e^{\pi_i})^{-\gamma} - \beta^{T+1}(-d_i)(w_{i0} * e^{(T+1)} * e^{\pi_i(T+1)})^{-\gamma}$$

This makes the first order condition a nonlinear function of observable variables, ($r, T, w_{i0}, D_i, X_i, p_i$),
and unobservable parameters, $\theta = \{\gamma, \mu, v, \omega\}$, that we need to estimate\(^{19}\). To estimate the
model, I use nonlinear least squares, choosing the parameters $\theta = \{\gamma, \mu, \omega, v\}$ that minimize the
quadratic form:

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

3.3.2 Identification:

Ideally, to identify a parameter like $\gamma$ 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

\(^{18}\)for derivation of $\pi_{it}$ see Appendix

\(^{19}\)I calibrate $\beta = .98$
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 ($\mu$ and $v$) are identified off of how the static level of term choice varies across these observable characteristics: we can identify the coefficients in $\mu$ 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 $\mu$ and $v$ help to control for how observable heterogeneity amongst borrowers, that are possible changing over price regimes, could impact maturity choice.

In contrast, $\gamma$, 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 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 changes. 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.

### 3.4 Results

The results from the structural estimation are shown in Table 11. 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 $\gamma$ 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 $\gamma$ 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 $\gamma$ upwards.
Individuals actively refinancing their loans are also likely more cognizant of the interest rate tradeoffs they are making than 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 $\gamma$. 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 $\gamma$, 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 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 $\gamma$, 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. 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, acting above

---

20These coefficients essentially restate the mean $T$ conditional on $X_i$ in terms of income growth.
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 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 $\gamma$, 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

\footnote{Omitted degree type is BA/BS.}
risk score and 92.5% of expected income for the lowest risk score.

Figures 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.

3.4.1 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 $\gamma$. 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 $\gamma$ 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 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 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 $\gamma$, but no longer estimate $\nu$ and $\mu$. 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
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 9: Observed vs. Estimated Cross-Sectional Age Earnings Profiles

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. 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 \( \gamma \) 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 \( \gamma \) 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 \( \gamma \): Calibrating, rather than estimating, these income paths also allows us to instead use the variation in term choices to estimate heterogeneity in \( \gamma \) (rather than heterogeneity in implied income growth). When I estimate \( \gamma \) as a function of observables (see 22I do not observe the same risk score variable in this second dataset, and therefore use FICO instead.)
Table 6: Estimated \( \gamma \) Coefficients

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>( \gamma )</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.59</td>
<td>(0.225, 0.95)</td>
</tr>
<tr>
<td>log(FCF)</td>
<td>-0.194</td>
<td>(-0.212, -0.175)</td>
</tr>
<tr>
<td>log(Amt)</td>
<td>0.255</td>
<td>(0.219, 0.292)</td>
</tr>
<tr>
<td>Home Owner</td>
<td>0.066</td>
<td>(0.042, 0.089)</td>
</tr>
<tr>
<td>Risk Score</td>
<td>-0.041</td>
<td>(-0.053, -0.027)</td>
</tr>
<tr>
<td>Dependent</td>
<td>0.087</td>
<td>(0.073, 0.102)</td>
</tr>
<tr>
<td>N</td>
<td>10075</td>
<td></td>
</tr>
</tbody>
</table>

Figure 10: Estimated Distribution of \( \gamma \)

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

Table 6, 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 \( \gamma \) (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 \( \gamma \) range from 1 to 1.9.

**Alternative Specifications:** Our primary specification models log income growth as an autoregressive process. Table 11 presents results from several alternative specifications which test 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:

\[ w_{it} = X'_{it} \mu + w_{it-1} + u_{it} \]

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 \( \gamma \). 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...
The results from this specification give a slightly higher estimate of \( \gamma \). 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 \( \gamma \), 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 \( \gamma \)), 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 \( \gamma \). I estimate \( \gamma \) as both a constant parameter and a function of observables using maximum likelihood estimation; the results are shown in Table 10. The average value of \( \gamma \) 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 \( \gamma \), perhaps due to the fact that they are less liquidity constrained, and thus less likely to choose a variable rate for interest rate savings.
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 refinancees. 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 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.

4.1 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.
4.1.1 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:

\[
\begin{align*}
\text{CV} & = u(w_1 - d) + E\left[\sum_{t=2}^{T} \beta^t u(w_t - d) + \sum_{t=1}^{T\text{max}} \beta^t u(w_t)\right] - \\
& - u(w_1 - d' + \text{CV}) + E\left[\sum_{t=2}^{T'} \beta^t u(w_t - d') + \sum_{t=T'+1}^{T\text{max}} \beta^t u(w_t)\right]
\end{align*}
\]

I express the \(CV\) both in absolute dollar terms, and also as a percent of the each borrower’s total interest payments.23,24

4.1.2 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 repre-

---

23 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:

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

24 I use a certainty equivalent approach when calculating expected utility, where:

\[
E[u(w_t - d)] \approx u(w_{i0} \cdot e^{r(X^t | \mu)} e^{\pi_{it}} - d)
\]

where now

\[
\begin{align*}
\pi_{it} & = \frac{1}{2} * t * \sigma^2 \left[1 - \gamma \frac{w_{i0} \cdot e^{r(X^t | \mu)}}{w_{i0} \cdot e^{r(X^t | \mu) - d_t}}\right] \\
& \quad \text{for } t < T+1 \\
\pi_{it} & = \frac{1}{2} * t * \sigma^2 (1 - \gamma) \\
& \quad \text{for } t \geq T+1
\end{align*}
\]
sent 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 \( r(T, p_i) \).

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 \( \alpha \) 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 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.\(^{26}\) One concern when using private sector costs as the basis for \( g \) is that the fixed costs of lending, \( \alpha \), in the private market are different then those in the public sector. In the private sector \( \alpha \) 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.

\(^{25}\)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.

\(^{26}\)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 20). This means any risk-based cost differential would have to stem from differences in the recovery rates across sectors.
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:

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

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:

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

4.2 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 Federal Interest Rates for graduate students\textsuperscript{27}. 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 11 compares the uniform benchmark to the opposite efficient extreme: a fully risk-based setting\textsuperscript{28}. 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 \( g \),

\textsuperscript{27}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%.

\textsuperscript{28}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, \( \{ T : E[\sum_{t=0}^{T} \frac{d^t}{r^t} (\frac{w_{i+1} - d(T)}{w_{i}(T+1)})^{-\gamma}] - \frac{d(T)}{r} = 0 \} \), 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 \( \frac{\partial d}{\partial r} \).
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.

Figure 11: Equity and Efficiency Impact of Uniform vs. Risk-based Pricing

Using our model, we can calculate the quantities represented here graphically: 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. The net change in surplus under the two schemes is equal an average of $1006 per borrower, or 5% of

---

29To 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 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 -0.9%, 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.

30The 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 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.
interest paid.

While uniform pricing redistributes surplus over risk type, it achieves more modest redistribution over another equity-relevant variable: income. Figure 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.

4.3 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 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 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 13 plots the distribution of compensating variation as a percentage of total interest paid on the borrowers’ loan, assuming a
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 12: Distribution of Interest Rates when Pricing on Different Observables

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 refinance, 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.
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 13: Distribution of CV as Percent of Total Interest Paid

4.3.1 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 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.

<table>
<thead>
<tr>
<th>Without Refinancing Frictions</th>
<th>Avg. Refinancer $\Delta$ CS</th>
<th>$\Delta g$</th>
<th>Avg. Non-refinancer $\Delta$ CS</th>
<th>Avg. $\Delta$ CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FICO-based Pricing</td>
<td>1315</td>
<td>0.61</td>
<td>-2040</td>
<td>64</td>
</tr>
<tr>
<td>Full Risk-based Pricing</td>
<td>2298</td>
<td>0.63</td>
<td>-2109</td>
<td>480</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>With Refinancing Frictions</th>
<th>Avg. Refinancer $\Delta$ CS</th>
<th>$\Delta g$</th>
<th>Avg. Non-refinancer $\Delta$ CS</th>
<th>Avg. $\Delta$ CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FICO-based Pricing</td>
<td>1896</td>
<td>0.06</td>
<td>-209</td>
<td>56</td>
</tr>
<tr>
<td>Full Risk-based Pricing</td>
<td>3040</td>
<td>0.12</td>
<td>-411</td>
<td>218</td>
</tr>
</tbody>
</table>

Table 7: Impact of Refinancing on Borrower Surplus and Break-even Interest Rate
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 changes. 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:

\[
U_{ij} = X_i'\mu - \alpha r(T,p)_{ij} + v \star p_i + \epsilon_{ij} \\
= V_{ij} + \epsilon_{ij}
\]

be the utility that individual \( i \) gets from refinancing with firm \( j \), where \( \epsilon_{ij} \) 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 exogenous 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 price changes 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:

\[
Pr(R_{ij} = 1) = \frac{\exp(V_{ij})}{\exp(V_{ij}) + \exp(V_{ij})}
\]

I estimate the model using a logistic regression framework. The resulting estimates (Table 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 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.

### 4.4 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 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

---

<table>
<thead>
<tr>
<th>Temporal Variation Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average APR</td>
</tr>
<tr>
<td>-2.363*** (0.443)</td>
</tr>
<tr>
<td>( N )</td>
</tr>
<tr>
<td>76494</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* \( p < 0.10 \), ** \( p < 0.05 \), *** \( p < 0.01 \)

Table 8: Extensive Margin Elasticities
<table>
<thead>
<tr>
<th>Without Refinancing Frictions</th>
<th>Per. Borrower $ Subsidy</th>
<th>Total Subsidy as % Total Interest Payments</th>
</tr>
</thead>
<tbody>
<tr>
<td>FICO-based Pricing</td>
<td>2027</td>
<td>10.5%</td>
</tr>
<tr>
<td>Full Risk-based Pricing</td>
<td>2086</td>
<td>10.9%</td>
</tr>
<tr>
<td>With Refinancing Frictions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FICO-based Pricing</td>
<td>266</td>
<td>1.4%</td>
</tr>
<tr>
<td>Full Risk-based Pricing</td>
<td>499</td>
<td>2.6%</td>
</tr>
</tbody>
</table>

Table 9: Subsidy Needed to Stop Unraveling

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 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 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.

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
These graphs show how the effective size of an interest rate subsidy, once accounting for the behavioral maturity and refinancing responses, can be larger 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.

Figure 14: Mechanical vs. Effective Interest Rate Subsidies, Accounting for Behavioral Responses

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 heterogeneous 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\textsuperscript{31}. 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.

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

\textsuperscript{31}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."
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.
References


## A Description of Federal Repayment Plans:

<table>
<thead>
<tr>
<th>Eligible Loan Types</th>
<th>Term</th>
<th>Monthly Payment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard Fixed Repayment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Direct Subsidized and Unsubsidized Loans</td>
<td>Up to 10 Years</td>
<td>Fixed amount, at least $50 per month</td>
</tr>
<tr>
<td>• Subsidized and Unsubsidized Federal Stafford Loans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• All PLUS loans</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Graduated Repayment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Direct Subsidized and Unsubsidized Loans</td>
<td>Up to 10 Years</td>
<td>Payments are lower at first and then increase every two years</td>
</tr>
<tr>
<td>• Subsidized and Unsubsidized Federal Stafford Loans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• All PLUS loans</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Extended Repayment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Direct Subsidized and Unsubsidized Loans</td>
<td>Up to 25 Years</td>
<td>Payments may be fixed or graduated</td>
</tr>
<tr>
<td>• Subsidized and Unsubsidized Federal Stafford Loans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• All PLUS loans</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Income Based Repayment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Direct Subsidized and Unsubsidized Loans</td>
<td>Up to 25 Years</td>
<td>Maximum payment is 15% of discretionary income</td>
</tr>
<tr>
<td>• Subsidized and Unsubsidized Federal Stafford Loans</td>
<td></td>
<td>o Discretionary income is the difference between your AGI and 150% of the</td>
</tr>
<tr>
<td>• All PLUS loans</td>
<td></td>
<td>poverty guideline</td>
</tr>
<tr>
<td>• Consolidation Loans (Direct or FFEL) that do not include Direct or FFEL PLUS loans</td>
<td></td>
<td>Payments change as income changes</td>
</tr>
<tr>
<td>made to parents</td>
<td></td>
<td>Individual needs to demonstrate partial financial hardship to qualify</td>
</tr>
<tr>
<td>• Consolidation Loans (Direct or FFEL) that do not include Direct or FFEL PLUS loans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>made to students</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Consolidation Loans (Direct or FFEL) that do not include Direct or FFEL PLUS loans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>made to students and must be a new borrower on or after Oct. 1, 2007, and must have</td>
<td></td>
<td></td>
</tr>
<tr>
<td>received a disbursement of a Direct Loan on or after Oct. 1, 2011.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Consolidation Loans (Direct or FFEL) that do not include Direct or FFEL PLUS loans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>made to parents and must be a new borrower on or after Oct. 1, 2007, and must have</td>
<td></td>
<td></td>
</tr>
<tr>
<td>received a disbursement of a Direct Loan on or after Oct. 1, 2011.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Direct Consolidation Loans that do not include (Direct or FFEL) PLUS loans</td>
<td>Up to 20 Years</td>
<td>Maximum payment is 10% of discretionary income</td>
</tr>
<tr>
<td>made to parents and must be a new borrower on or after Oct. 1, 2007, and must have</td>
<td></td>
<td>Payments change as income changes</td>
</tr>
<tr>
<td>received a disbursement of a Direct Loan on or after Oct. 1, 2011.</td>
<td></td>
<td>Individual needs to demonstrate partial financial hardship to qualify</td>
</tr>
</tbody>
</table>

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

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

Figure 15: Federal Loan Repayment Plans
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:

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

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

$$ u_{it} \sim N(0, \sigma_u^2) $$

$$ \sigma_u^2 = (\omega - v * p_i)^2 $$

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:

$$ E[(w_{i0} * e^{t*(X_i' \mu) + \sum_{1}^{t} u_{it}} - d_i)^{-\gamma}] = (w_{i0} * e^{t*(X_i' \mu) + \sum_{1}^{t} u_{it}} - d_i)^{-\gamma} $$

where $\pi_{it}$ 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:

\[
\pi_{it} = \frac{1}{2} t \sigma^2 (1 + \gamma) \frac{w_{i0} e^{t\epsilon(X_i^{\mu})}}{w_{i0} + e^{t\epsilon(X_i^{\mu})} - d_i} \]

for \( t < T+1 \)

\[
\pi_{it} = \frac{1}{2} t \sigma^2 (-\gamma) \]

for \( t \geq T+1 \)

To derive \( \pi_{it} \), note that we can write:

\[
E[u'(w_{it})] = E[u'(w_{i0} + e^{t\epsilon(X_i^{\mu})} + e^{\sum_{i} \epsilon_i} - d_i)]
\]

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

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

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

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

We first take the derivative of this expression w.r.t. \( \sigma \):

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

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. \( \sigma \), and evaluate it at \( \sigma = 0 \):

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

\[
\pi''(0) = [1 + w_{i0} u''(w_{i0} - d_i)]
\]

Under the assumption of CRRA utility, this becomes:

\[
\pi''(0) = [1 + w_{i0} u''(w_{i0} - d_i)]
\]

\[
= [1 - (1 + \gamma) \frac{w_{i0}}{w_{i0} - d_i}]
\]

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:
\[ g_i(\theta) = \sum_{t=1}^{T} \beta_t \frac{\partial d_t}{\partial \theta} (w_{i0} * e^{i\pi \mu} * e^{\pi t - d_t})^{-\gamma} - \beta^{t+1} (-d_t)(w_{i0} * e^{(T+1)\pi \mu} * e^{\pi t+1})^{-\gamma} \]

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

\[ b = \arg \min_{\theta} g_i(\theta)'g_i(\theta) \]
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 an 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 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 16: CV for Uniform 7% APR, 10 Year Fixed Plan relative to 7% with Term Flexibility and Market Prices with Term Flexibility
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:

\[
g_n(\theta) = \frac{1}{N} \sum_{i=1}^{N} \sum_{s=1}^{S} \left( \sum_{t=1}^{T} \beta_t \frac{dd_t}{d\theta_t} (w_{ist} - d(T_t))^{-\gamma} - \beta_{t+1} (-d(T_t))(w_{ist})^{-\gamma} \right)
\]

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
- \(T_{\text{max}}\) = 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 \(N \times T_{\text{max}} \times S\) 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: \((\theta = \{\gamma, \alpha_{60}, \alpha_{240}, \mu, v, \omega\})\).

3. For each individual, scale the random error by \(\omega - v \ast p_i\). 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: \(y_{ist} = X'_i\mu + y_{ist(t-1)} + u_{ist}\), where \(X'_i\mu\) is the fixed consumption trend, \(y_{ist(t-1)}\) 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 $\alpha_{60}$ or $\alpha_{240}$ 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 – $\gamma$ 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.
E Estimation of $\gamma$ using Borrowers’ Fixed/Variable Choice

The fixed variable choice impacts a shorter time horizon than the term choice, since every year the borrower is allowed to switch between the two rates; I therefore model the trade off the borrower makes between the premium paid over one year for a fixed rate loan vs. the unknown interest rate variation that could occur over the next year. Under a variable rate contract, the monthly payment is tied to the market interest rate $r_t$, which changes with uncertainty over the following year. I model this change as the process:

$$r_{t+1} = r_t + \kappa + e_i$$

where $\kappa$ is the expected mean change in interest rate over the next year, and $e_i$ represents a mean-zero, random shock with $Var(e_i) = \sigma_e^2$. To calibrate $\kappa$ and $\sigma_e^2$, I use the Federal fund futures prices, which have long been used to express the markets views on the likelihood of changes in interest rates. These futures provide a probability distribution of where the market believes interest rates will be in the next week, month, or even year. I use the probability distribution that existed when individual $i$ refinanced to calculate first $\kappa$, which is the expected mean change in interest rates over the next year, and also $\sigma_e^2$, which is market-based variance of this expectation. For example, in June 2016, the futures prices implied a probability distribution for rates in the next year given by:

<table>
<thead>
<tr>
<th>Rate Change (bp)</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-25</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>+25</td>
<td>38</td>
</tr>
<tr>
<td>+50</td>
<td>10</td>
</tr>
<tr>
<td>+75</td>
<td>1</td>
</tr>
</tbody>
</table>

This translates into $\kappa = 15$ bp, and a variance equal to 325 bp. Using these market-based expectations about interest rates requires the assumption that beliefs about $\sigma_e^2$ held by the borrowers I study are in line with those of the market and constant cross-sectionally across the sample.

An individual’s variable rate monthly payment will permanently increase or decrease when their variable interest rate changes by $\kappa + e_i$. Specifically, the monthly payment will change as follows:

$$MP_{t+1}^V = MP^V(r_{it} + \kappa) + m_i$$

where $r_{it}$ is individual $i$’s starting interest rate and $MP^V(r_{it} + \kappa)$ is the monthly payment.
calculated with the new expected interest rate. The error term \( m_i \sim N(0, \sigma^2_{MP}) \) is a nonlinear transformation of \( e_i \), which I approximate as \( \sigma^2_{MP} \approx \sigma^2_{r} \frac{dP^V}{dr} \). 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 \( MP^V \) 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.

When choosing between the two rates (conditional on term), an individual compares the known utility of the fixed rate to the expected utility of the variable rate. 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. Our assumption that \( E[r_t + \kappa + e_t] = r_t + \kappa \) and \( Var[r_t + \kappa + e_t] = \sigma^2_e \) allows me to use a certainty equivalence approach when modeling how borrowers compare expected utility from the fixed and variable rate contracts. The expected utility from choosing the variable rate (\( E[u(w - (MP^V(r_t + \kappa) + m_i))] \)) can be expressed as the utility of its certainty equivalent, \( u(w - MP^V) \) where:

\[
w - MP^V = w - MP^V(r_t + \kappa) + \frac{\sigma^2_{MP}}{2} \cdot \frac{u''(w - MP^V(r_t + \kappa))}{u'(w - MP^V(r_t + \kappa))}
\]

This value shows that while individuals pay a lower monthly payment with the variable rate, the uncertainty related to it also lowers its expected utility. The certainty equivalent decreases as an individual becomes more risk averse (\( \gamma \) increases), the expected variance of the monthly payment increases (\( \sigma^2_{MP} \)), or as the relative “stakes” of the gamble become larger (\( \frac{1}{w-\text{MP}^V(r_t+\kappa)} \)).

In contrast, the per-period utility of the fixed rate contract is known and given by \( u(w - MP^F) \). Conditional on the choice of term, the choice of fixed versus variable rate is determined by a comparison of the utility of the fixed rate contract vs. the utility of the certainty equivalent of the variable rate contract over the next year. An individual would be indifferent if:

\[
U^F_i = u(w - MP^F_i) = u(w - MP^V) = E[U^V_i] \\
\rightarrow w - MP^F_i = w - MP^V(r_t + \mu) - \frac{\gamma \sigma^2_{MP}}{2} \cdot \frac{1}{w - MP^V(r_t + \mu)}
\]

\[
\rightarrow \gamma = -\frac{(MP^V(r_t + \mu) - MP^F) \cdot 2(w - MP^V(r_t + \mu))}{\sigma^2_{MP} \cdot \tilde{i}}
\]

This equation shows how we can identify \( \gamma \) using small exogenous changes in the fixed variable spread \( \tilde{i} \). The spread \( \tilde{i} \) depends not only on an individual’s income, debt amount, and term, but also on the fixed variable spread they face. Therefore variation in interest rates that impacts the spread...
will shift $\tilde{\gamma}_i$ exogenously, changing the share of borrowers for whom the inequality holds. As the share of borrowers with $\gamma < \tilde{\gamma}_i$ increases, we should see a corresponding increasing in the percentage of borrowers choosing a variable rate. If we assume that $\gamma$ is lognormally distributed, then we get the expression that an individual will choose a variable rate if:

$$\ln(\gamma) + \epsilon_i < \ln(\tilde{\gamma}_i)$$

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 price regimes:

$$Pr(\text{Variable}_i) = \Phi(\ln(\tilde{\gamma}_i) - \ln(\gamma))$$  \hspace{1cm} (1)

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 estimate $\gamma$. The figure below plots the average value of $\tilde{\gamma}_i$ under each interest rate regime, as well as the share choosing a variable rate under each regime – recall everyone with a value of $\tilde{\gamma}_i$ above their true risk aversion $\ln(\gamma) + \epsilon_i$ parameter will choose a variable rate. The different regimes shift the distribution of indifference points substantially up and down, and the share choosing a variable rate moves in almost perfect tandem, in line with the predictions of our model.

I estimate $\gamma$ as both a constant parameter and a function of observables using maximum likelihood estimation; the results are shown in Table 10. The magnitude of $\gamma$ estimated using the fixed variable choice is 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 $\gamma$, perhaps due to the fact that they are less liquidity and credit constrained, and thus less likely to choose a variable rate for interest rate savings.

Here I plot the changes in the average value of $\tilde{\gamma}$ over price regimes against changes in the share choosing a variable rate. The two move in almost perfect tandem.
<table>
<thead>
<tr>
<th>Coefficients: $\gamma$</th>
<th>$\gamma$</th>
<th>$\gamma$ w/ observables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.85</td>
<td>-0.62</td>
</tr>
<tr>
<td>Risk</td>
<td>0.067</td>
<td></td>
</tr>
<tr>
<td>$w$</td>
<td>2.14e-05</td>
<td></td>
</tr>
<tr>
<td>Mortgage</td>
<td>-0.0523</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.136</td>
<td></td>
</tr>
<tr>
<td>Age$^2$</td>
<td>-0.00148</td>
<td></td>
</tr>
<tr>
<td>FICO</td>
<td>2.5 e-05</td>
<td></td>
</tr>
<tr>
<td>log(D)</td>
<td>-0.164</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predictions: $\gamma$ w/observables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Min</td>
</tr>
<tr>
<td>Max</td>
</tr>
<tr>
<td>N</td>
</tr>
</tbody>
</table>

Table 10: Estimates and Predicted Distribution of $\gamma$ using Fixed Variable Rate Choice
F 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:

If $w_{it} - d_i < x$

Pay $n = w_{it} - x$ in period $t$

Pay $(d_i - n)(1 + r)^{T^*+1-t} + R_{it}$ in period $T^* + 1$

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:

$$E[U(T^*)]' = E[U(T^*)] - \beta^t(u(w_{it} - d_i) - u(w_{it} - n))$$

Additional utility in period $t$ when constraint binds

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

Negative utility in period $T + 1$ when pay additional interest & penalty

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:
\[ E[U(T^*)] = E[U(T^*')] - \epsilon \]

and
\[ E[U(t')] = E[U(t)] - \epsilon \quad \forall t \neq T^* \]

This means that:

\[ E[U(T^*)] < E[U(T^*')] \]

and
\[ E[U(T^*)] > E[U(t')] \quad \forall t \neq T^* \]

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.

G Model Estimates
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Log Income Process</th>
<th>Non-Log Process</th>
<th>Constant Income</th>
<th>Simulated MM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>1.119</td>
<td>1.289</td>
<td>1.848</td>
<td>1.024</td>
</tr>
<tr>
<td></td>
<td>(1.0833, 1.1546)</td>
<td>(1.2027, 1.3735)</td>
<td>(1.5228, 2.1735)</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - Constant</td>
<td>7.49E-02</td>
<td>9399.965</td>
<td>-</td>
<td>2732.100</td>
</tr>
<tr>
<td></td>
<td>(0.055561, 0.094261)</td>
<td>(7837.9, 11101)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - Starting Income</td>
<td>7.48E-03</td>
<td>0.901</td>
<td>-0.169</td>
<td>0.910</td>
</tr>
<tr>
<td></td>
<td>(0.0060717, 0.0088882)</td>
<td>(0.79995, 1.0042)</td>
<td>(-0.4772, 0.1358)</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - log(Debt)</td>
<td>-9.35E-03</td>
<td>-611.649</td>
<td>-0.386</td>
<td>-393.600</td>
</tr>
<tr>
<td></td>
<td>(-0.010121, -0.0085806)</td>
<td>(-692.5, -532.87)</td>
<td>(-0.6329, -0.1309)</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - Home Owner</td>
<td>-8.21E-05</td>
<td>8.793</td>
<td>-0.203</td>
<td>16.577</td>
</tr>
<tr>
<td></td>
<td>(-0.0006079, 0.00069657)</td>
<td>(-96.163, 113.53)</td>
<td>(-0.41776, 0.012249)</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - # Dependents</td>
<td>5.31E-04</td>
<td>163.838</td>
<td>-0.446</td>
<td>276.350</td>
</tr>
<tr>
<td></td>
<td>(0.00011908, 0.00094268)</td>
<td>(96.153, 237.02)</td>
<td>(-0.57852, -0.3153)</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - Age</td>
<td>1.44E-04</td>
<td>108.083</td>
<td>-0.047</td>
<td>103.920</td>
</tr>
<tr>
<td></td>
<td>(0.00003136, 0.00061954)</td>
<td>(36.004, 185.86)</td>
<td>(-0.18916, 0.095549)</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - Age$^2$</td>
<td>-2.51E-06</td>
<td>-1.119</td>
<td>0.000</td>
<td>-1.790</td>
</tr>
<tr>
<td></td>
<td>(-8.739e-06, 3.7159e-06)</td>
<td>(-2.2437, -0.06739)</td>
<td>(-0.018908, 0.0019279)</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - Degree MD</td>
<td>4.05E-04</td>
<td>-391.035</td>
<td>0.449</td>
<td>-601.340</td>
</tr>
<tr>
<td></td>
<td>(-0.0026784, 0.0034888)</td>
<td>(-791.4, 12.713)</td>
<td>(-0.39542, 1.2935)</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - Degree JD</td>
<td>-4.42E-03</td>
<td>-607.457</td>
<td>0.586</td>
<td>-992.690</td>
</tr>
<tr>
<td></td>
<td>(-0.0056486, -0.003201)</td>
<td>(-796.38, -419.51)</td>
<td>(0.25414, 0.9177)</td>
<td></td>
</tr>
<tr>
<td>$\mu$ - Degree Masters</td>
<td>-2.11E-03</td>
<td>-297.765</td>
<td>0.239</td>
<td>-506.180</td>
</tr>
<tr>
<td></td>
<td>(-0.0029062, -0.0013099)</td>
<td>(-418.72, -182.43)</td>
<td>(0.025228, 0.45221)</td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>4.45E-04</td>
<td>8.018</td>
<td>-</td>
<td>440.390</td>
</tr>
<tr>
<td></td>
<td>(-0.00080498, 0.00016949)</td>
<td>(-1.782, 17.818)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$v$ - Risk Score</td>
<td>-2.22E-04</td>
<td>-33.211</td>
<td>0.571</td>
<td>-81.936</td>
</tr>
<tr>
<td></td>
<td>(-0.0038, -0.0035)</td>
<td>(-63.787, -2.635)</td>
<td>(0.47069, 0.67065)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_{40}$</td>
<td>0.89</td>
<td>0.247</td>
<td>2.038</td>
<td>-0.127</td>
</tr>
<tr>
<td></td>
<td>(-1.1462, -5.5437)</td>
<td>(0.1548, 0.3401)</td>
<td>(1.7933, 2.2825)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_{240}$</td>
<td>1.23</td>
<td>1.92</td>
<td>-12.69</td>
<td>3.33</td>
</tr>
<tr>
<td></td>
<td>(0.65173, 1.45392)</td>
<td>(1.3815, 2.4539)</td>
<td>(-13.897, -11.474)</td>
<td></td>
</tr>
</tbody>
</table>

| N                  | 11585              | 11585           | 11585           | 11585        |

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.

Table 11: Choice Model Parameter Estimates
### Additional Figures and Tables

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

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

#### Table 12: Budget Lifetime Default Rates

![Graph](image)

(a) Response of Share Choosing Variable to Exogenous Variation in the Fixed Variable Spread:

**Figure 17:** 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.
These plots show the value of $\gamma$ that would be necessary to rationalize the term choice of each individual under the assumption of constant income and $\beta = 1$.

Figure 18: Baseline Estimates of $\gamma$

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 overpredicts terms, but otherwise seems to perform well. All counterfactual exercises use these predicted term choices as a comparison point.

Figure 19: Model Fit
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 20: Cost Differential Relative to Lowest Risk Rating

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 21: Predicted Term Choices over Risk Score under Uniform 6.6% vs. Market Prices
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 22: Propensity to Refinance over CV

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 exogenous price shifts.

Figure 23: Variation in Observable Characteristics over Time
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 24: Average Size of Extra Payments Over Time Made by Borrowers

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.

Figure 25: Cross-Sectional Age Earnings Profiles, by Degree and Occupation
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 26: Levels and Changes in Other Monthly Payments Before and After Refinancing

<table>
<thead>
<tr>
<th>Initial Monthly Payment</th>
<th>Initial Auto</th>
<th>Initial Real Estate</th>
<th>Initial Non-Real Estate</th>
<th>Initial Credit Card</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>449.32</td>
<td>1927.72</td>
<td>1104.31</td>
<td>94.37</td>
</tr>
<tr>
<td>Median</td>
<td>386.00</td>
<td>1698.00</td>
<td>882.00</td>
<td>52.00</td>
</tr>
<tr>
<td>IQR (25,75)</td>
<td>246.50</td>
<td>1210.00</td>
<td>897.00</td>
<td>84.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change in Monthly Payment</th>
<th>Auto Change</th>
<th>Real Estate Change</th>
<th>Non-Real Estate Change</th>
<th>Credit Card Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>27.15</td>
<td>63.64</td>
<td>-181.06</td>
<td>-12.74</td>
</tr>
<tr>
<td>Median</td>
<td>0.00</td>
<td>0.00</td>
<td>-75.00</td>
<td>0.00</td>
</tr>
<tr>
<td>IQR (25,75)</td>
<td>0.00</td>
<td>37.00</td>
<td>490.00</td>
<td>52.00</td>
</tr>
</tbody>
</table>

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 27: Investment Balances over Lifetime

<table>
<thead>
<tr>
<th>Age</th>
<th>25th Percentile</th>
<th>75th Percentile</th>
<th>50th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

73
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.

Figure 28: Average Government Subsidy over Borrower Risk Type and Borrower Income