HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the Committee for the PhD in Business Economics have examined a dissertation entitled

Essays on the US Mortgage Market

Presented by David Hao Zhang

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature ______________________

John Campbell, Chair

Signature ______________________

Adi Sunderam

Signature ______________________

Edward Glaeser

Signature ______________________

Ariel Pakes

Signature ______________________

Robin Lee

Date: March 21, 2022

Essays on the US Mortgage Market

A DISSERTATION PRESENTED
BY
David Hao Zhang
TO
The Department of Economics

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy
IN THE SUBJECT OF
Business Economics

Harvard University
Cambridge, Massachusetts
March 2022

Dissertation Advisor:
Professor John Y. Campbell
Professor Adi Sunderam

Author:
David Hao Zhang

Essays on the US Mortgage Market

Abstract

This dissertation explores issues surrounding inequality and inefficiency in US mortgage markets. In the first chapter, I identify inefficiencies in the US mortgage market stemming from heterogeneous borrower refinancing tendencies. In the second chapter, I develop a new methodology to assess discrimination in the menus context and apply it to the US mortgage market. In the third chapter, I assess the sources of racial inequality in mortgage payments and find that it is primarily driven by racial differences in prepayment speeds.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I am grateful to my advisors John Campbell, Adi Sunderam, Edward Glaeser, Ariel Pakes and Robin Lee for their support and guidance. They introduced me to the process of economic research. This thesis would not have been possible without their patience.

In addition, I thank Paul Willen and Kris Gerardi for many conversations regarding the institutional details surrounding the mortgage market. I also benefited from the detailed feedback of Isaiah Andrews, Elie Tamer, and the late Gary Chamberlain on methodological issues. I enjoyed thinking about Adi's "research production possibilities frontier" which served as a philosophical principle underlying my research process.

Furthermore, I thank my officemates Jiacheng Feng, Gabriel Levin-Konigsberg, and Jeffrey Wang who supported me through the process. I benefitted greatly from frequent discussions with other graduate students at Harvard, including Tianwang Liu, Frank Pinter, Gregor Schubert, Audrey Tiew, Ron Yang and also too many others to list.

Finally, I thank my parents, my fiancée, and friends for supporting me through the long journey.

0

Introduction

This dissertation has three chapters. In the first chapter, I use a structural model to study the cross-subsidization in the US mortgage market due to heterogeneous borrower refinancing tendencies. Quick to refinance borrowers gain up to 3% of their loan amount relative to slow to refinance borrowers in expectation. In equilibrium, the presence of slow to refinance borrowers reduces mortgage interest rates particularly on lower upfront closing cost mortgages which have more valuable refinancing options. As a result, quick to refinance borrowers refinance excessively relative to a perfect information, no cross-subsidization benchmark, an effect that accounts for around 28% of all US refinancing and generates significant deadweight losses due to administrative resource costs. Alternative contract designs can simultaneously reduce transfers and increase total welfare.

In the second chapter, co-authored with Paul Willen, we use a new methodology to assess mortgage pricing discrimination by race. First, we show how standard regression-based approaches for assessing differences in the menus of options offered to borrowers can lead to misleading and contradictory results. Second, we introduce a methodology that can detect differences in menus based on relatively weak assumptions. More specifically, we use pairwise dominance relationships in choices that can be supplemented by restrictions on the range of plausible menus to define (1) a test statistic for equality in menus and (2) a difference in menus (DIM) metric. Third, we derive a new procedure for inference on our class of problems. Fourth, we apply our methodology to a novel data set linking 2018–2019 Home Mortgage Disclosure Act (HMDA) data to Optimal Blue rate locks. We find strong evidence for mortgage pricing differentials by race, particularly among Conforming mortgage borrowers who are relatively creditworthy.

In the third chapter, co-authored with Kris Gerardi and Paul Willen, I study the sources of interest rate differentials between borrowers. Over the period 2005 to 2015, Black borrowers paid more than 40 basis points higher mortgage interest rates than Non-Hispanic white borrowers. We show that the main reason is that Non-Hispanic white borrowers are much more likely to exploit periods of falling interest rates by refinancing their mortgages or moving. Black and Hispanic white borrowers face challenges refinancing because, on average, they have lower credit scores, equity and income.

But even holding those factors constant, Blacks and Hispanic white borrowers refinance less suggesting that other social factors are at play. Because they are more likely to exploit lower interest rates, white borrowers benefit more from monetary expansions. Policies that reduce barriers to refinancing for minority borrowers and alternative mortgage contract designs can significantly reduce racial mortgage rate inequality.

1

Closing Costs, Refinancing, and Inefficiencies in the Mortgage Market

1.1 Introduction

What are the welfare consequences of heterogeneity in borrower refinancing behavior under the predominant US mortgage contract design? In the US, many borrowers are slow to refinance their mortgages when interest rates fall, which has long been recognized as an important friction in household finance.[1] I use a structural model to quantify the distribution of refinancing inertia in the population and to study its equilibrium consequences.

In the cross-section, I find large borrower heterogeneity in refinancing tendencies. This heterogeneity in refinancing behavior has significant consequences in terms of the households' expected financial outcomes. In particular, a borrower with no refinancing inertia is able to gain up to 3% of the loan amount from an ex ante perspective relative a borrower with high refinancing inertia. As a result, the extent to which heterogeneity in borrower refinancing behavior varies in the population and by borrower demographics, which I estimate using a heterogeneous agent model, has important distributional consequences.

The presence of slow to refinance borrowers also affects mortgage interest rates in equilibrium. I show that the existence of borrowers with large refinancing inertia reduces the interest rates lenders charge, an effect that is particularly strong on lower upfront closing cost mortgages which have more valuable refinancing options. This in turn generates a distortion in the contract choices of quick to refinance borrowers and incentivizes them to refinance more than they otherwise would, which exacerbates transfers between borrowers with different refinancing speeds and generates deadweight losses in a form of administrative costs involved with excessive refinancing.

In particular, a novel result of my paper is that a significant fraction of US refinancing, around

[1] See, e.g., Schwartz and Torous (1989), Archer and Ling (1993), McConnell and Singh (1994), Stanton (1995), Green and LaCour-Little (1999a), Campbell (2006b), Agarwal, Rosen, and Yao (2016), Keys, Pope, and Pope (2016a), Johnson, Meier, and Toubia (2018a), Andersen et al. (2018), and Gerardi, Willen, and Zhang (2021).

28%, is excessive relative to a perfect information, no cross-subsidization benchmark. In other words, borrowers with low refinancing inertia refinance too much at the expense of borrowers with high refinancing inertia. This suggests that policies that reduce disparities in refinancing has the potential of simultaneously reducing inefficiency in the market while having attractive distributional properties.

By way of background, mortgage originating lenders must cover their costs. They can do so in two ways. First, they can charge the borrower upfront, through upfront closing costs known as "points."[2] Second, they can raise the interest rate on the mortgage, holding fixed its principal balance and then recovering their costs from the secondary market.[3] Borrowers can therefore choose to get a lower rate, higher upfront closing cost mortgage, or a higher rate, lower upfront closing cost mortgage. Higher rate, lower upfront closing cost mortgages by construction carry a more valuable refinancing option, and I show that they are cross-subsidized more in equilibrium. This incentivizes actively refinancing borrowers to refinance more often than they otherwise would. The extra refinancing comes from two mechanisms. First, actively refinancing borrowers become more likely to take out a mortgage with a higher rate and lower upfront closing costs, which mechanically carries a higher refinancing incentive.[4] Second, actively refinancing borrowers are able to refinance more

[2]In the industry, each mortgage point refers to 1% of the loan amount that borrowers pay upfront.

[3]Borrowers could also in principle pay their closing costs upfront and then add them to the principal balance and thus hold the interest rate fixed. In practice, this may be infeasible for many borrowers who face binding loan-to-value (LTV) and debt-to-income (DTI) constraints.

[4]This is consistent with Dave Ramsey's financial advice, which says that: "most buyers won't regain their money on mortgage points because they usually refinance, pay off, or sell their homes before they reach their break-even point." Source: https://www.ramseysolutions.com/real-

cheaply than they otherwise would by taking out a low upfront closing cost mortgage when they do refinance. Because mortgage refinancing involves administrative resources that could have been used for other economic activity, the extra refinancing that quick to refinance borrowers undertake solely to receive transfers generates deadweight losses from a social perspective.

The economic consequences of this distortion are significant. From a structural model of heterogeneity in borrower refinancing behavior, I find that approximately one quarter of all US mortgage refinancing would not have occurred without the cross-subsidization of lower upfront closing cost mortgages. Therefore, the heterogeneity in borrower refinancing behavior combined with the predominant mortgage contract design generates not only transfers between borrowers with different refinancing speeds but also deadweight losses through excessive refinancing by the quick to refinance borrowers.

I begin the analysis by documenting five stylized facts about the mortgage market in Section 1.4 to provide reduced-form evidence of this cross-subsidization and to inform the model. First, I show that almost all borrowers pay for most of their mortgage closing costs through a higher interest rate on their mortgage relative to adjusted mortgage-backed securities (MBS) yields. Second, many borrowers are slow to refinance, while others are more quick at doing so. These two facts naturally lead to a third: for the same amount of mortgage closing costs added to the rate, different borrowers will end up paying very different amounts for it, measured in terms of the ex post net present values (NPVs) of extra interest payments. Fourth, I find that this heterogeneity in ex post NPVs is

estate/what-are-mortgage-points. The financial advice to not pay more in upfront closing costs or "points" in exchange for a lower rate turns out to be correct for my benchmark optimally refinancing borrower, but only due to the cross-subsidization from slow to refinance borrowers.

predictable via ex ante demographics. In particular, for my sample of mortgages originated in 2013, Black and Hispanic borrowers paid an extra 0.43% and 0.48%, respectively, of the loan amount more in ex post NPVs of extra interest payments controlling for the amount closing costs added to the rate. This suggests a role for ex ante cross-subsidization in addition to ex post. Fifth, I find evidence of both selection of borrowers into upfront closing cost choices as well as substantial heterogeneity in borrower prepayment behavior among borrowers with the same upfront closing cost choice.

To quantify the size of the cross-subsidy and study its ex ante welfare consequences, I develop a structural equilibrium model that captures borrower heterogeneity in refinancing and moving tendencies while endogenizing borrower choices of upfront closing costs in Section 2.3. Since my novel source of cross-subsidization is about how borrowers who are slow to refinance effectively pays the closing costs of more actively refinancing borrowers, it is important for my model to capture heterogenous borrower refinancing behavior along with their equilibrium effect on mortgage interest rates. To do so, I embed the time and state dependence of borrower refinancing behavior described in Andersen et al. (2020a) into a life-cycle model that gives welfare estimates interpretable in dollar-equivalent terms. A zero-profit condition pins down the supply side. As such, my model is able to quantify the equilibrium welfare implications of heterogenous borrower refinancing behavior in dollar terms, which is an independent contribution to the literature that has primarily studied this issue in the reduced form.

Borrowers in my model are heterogeneous in terms of their (i) refinancing costs, including a time-varying ability to refinance and a hassle cost conditional on them being able to refinance, (ii) ex ante moving hazards, and (iii) discount factors. The time-varying ability to refinance and the refinancing hassle cost are separately identified from borrower delays in refinancing after their refinancing thresholds has been reached (Andersen et al., 2020a): a behavior not reconcilable with only a fixed hassle cost component and suggests a role for borrowers' time-varying ability to refinance. Conditional on refinancing costs, borrowers' ex ante moving expectations are identified from the correlation between their responses to the refinancing incentive and their subsequent moving decisions. The fact that borrowers who do not refinance despite facing a large incentive to do so are more likely to move shortly thereafter suggests that they have heterogenous ex ante moving expectations that affect their refinancing decisions. Finally, conditional on borrower moving and refinancing types, their discount factors are identified from their choices of upfront closing costs.

Three main conclusions emerge from my empirical work. First, cross-subsidization from slow-to-refinance borrowers significantly affects equilibrium prices and is larger on mortgages with lower upfront closing costs. For a calibrated borrower who is always able to refinance at zero hassle costs, a mortgage with a one percent upfront closing cost carries a 1.25% lower interest rate in the existing market equilibrium relative to a world without cross-subsidization. For mortgages with a four percent upfront closing cost, the difference is smaller at 0.26%. The intuitive reason for the larger cross-subsidization of low upfront closing cost mortgages is that, from the perspective of the lender, slow to refinance refinance borrowers overpay for their mortgage closing costs when they add it to the rate because they keep paying the higher interest rate for longer and subsequently cross-subsidize the lower upfront closing cost mortgages more.

A key advantage of my approach is that I am able to quantify the consequences of this cross-subsidization. As my second conclusion, I find that the cross-subsidization of mortgage closing costs generates large transfers between borrowers. I estimate that eliminating the cross-subsidization of mortgage closing costs by requiring all borrowers to add their closing costs to the balance of the loan[5] would reduce the average difference in borrower utility in the current world relative to the no cross-subsidization benchmark by approximately half. Thus, approximately half of the cross-subsidization in the US mortgage market due to differences in prepayment tendencies can be attributed to the equilibrium consequences of the contractual feature where borrowers finance their mortgage closing costs via the rate rather than via a standard loan. As my third conclusion, I show that the efficiency consequences of price distortions are large. In particular, I estimate that around one quarter of all US refinancing would not have occurred but for this cross-subsidization, leading to a welfare loss of around $3.5 billion per year relative to the no cross-subsidization benchmark.

Using the model, I conduct two counterfactual analyses in Section 1.7. First, I investigate borrower welfare under an alternative contract design where their closing costs have to be added to the mortgage balance. I find a reduction in cross-subsidization from $1339/borrower to $698/borrower, a decrease of 48%. Furthermore, I find an increase in average borrower utility of $556/borrower in dollar terms. Second, I study the case of automatically refinancing mortgages. This contract eliminates the cross-subsidization between borrowers with different refinancing speeds, and leads to a bigger increase in average borrower utility of $1215/borrower. My results suggests that the equity-efficiency trade-off is not binding in the US mortgage context: it is possible to reduce inequality while increasing total welfare.

---

[5] Requiring mortgage closing costs to be paid upfront or added to the balance of the loan both eliminate the cross-subsidization. I chose to model it as the latter to avoid borrower liquid savings constraints.

My paper is about how information frictions can lead to cross-subsidization and how the incentives generated by such cross-subsidization can lead to inefficiency. Without information frictions, it would be Pareto optimal to have complete markets for mortgage contracts under competitive equilibrium.[6] With information frictions, actively refinancing borrowers may seek out contracts that carry more cross-subsidization and generate deadweight losses in the process. The fact that actively refinancing borrowers may select into contracts that are more heavily cross-subsidized but are "costlier" for the lender may be considered a form of adverse selection in the sense of Einav et al. (2013). As such, my findings have close parallels in the health insurance market, where it is well-known that adverse selection can lead to sub-optimal outcomes under competitive equilibrium and that contractual mandates can improve welfare.[7] Finally, while my model is fully consistent with all borrowers being rational, if one instead views the slow to refinance borrowers who add closing costs to the rate as behavioral agents who do not understand the true cost of a higher interest rate, it can also be interpreted as an empirical model of a shrouded equilibrium as in Gabaix and Laibson (2006).

My paper is primarily related to the literature on borrower heterogeneity in mortgage refinancing behavior. Many papers document and model the large borrower heterogeneity in refinancing behavior conditional on the interest rate savings available, including Archer and Ling (1993), McConnell

[6] Arrow and Debreu (1954).

[7] See, e.g., Akerlof (1970), Rothschild and Stiglitz (1976), Wilson (1977) for early theoretical work, Einav and Finkelstein (2011) for a review, and Cutler and Reber (1998), Einav, Finkelstein, and Cullen (2010), Bundorf, Levin, and Mahoney (2012), Hackmann, Kolstad, and Kowalski (2015), Handel, Hendel, and Whinston (2015), Handel, Kolstad, and Spinnewijn (2019), Ho and Lee (2021) for examples of empirical work in this area.

and Singh (1994), Stanton (1995), Deng, Quigley, and Van Order (2000a), Agarwal, Rosen, and Yao (2016), Keys, Pope, and Pope (2016a), Johnson, Meier, and Toubia (2018a), Andersen et al. (2018), Beraja et al. (2018a), Ambokar and Samaee (2019), and Gerardi, Willen, and Zhang (2021). This literature has primarily studied borrower refinancing heterogeneity in the reduced form and is not able to quantify the degree of the cross-subsidization between borrowers with different refinancing speeds among various mortgage contracts. A contemporaneous paper, Fisher et al. (2021), studies cross-subsidization in the UK mortgage market using a structural model. My model captures richer dynamics in terms of stochastic interest rates and heterogeneous borrower moving expectations which are quantitatively important, and points out the significant impact of borrower selection of mortgage contracts in terms of upfront closing costs in determining the equilibrium transfers and deadweight losses.

My paper also contributes to the literature on life-cycle models of mortgage choice. This includes Campbell and Cocco (2003), Mayer, Piskorski, and Tchistyi (2013), Corbae and Quintin (2015), Campbell and Cocco (2015) Eichenbaum, Rebelo, and Wong (2018), Chen, Michaux, and Roussanov (2020), Campbell, Clara, and Cocco (2021) and Guren, Krishnamurthy, and McQuade (2021). Of these papers, only Eichenbaum, Rebelo, and Wong (2018) incorporate equilibrium cross-sectional heterogeneity in refinancing behavior, which they use to model the state-dependent behavior of monetary policy, but they do not endogenize the mortgage premia and subsequently do not study its implications in terms of borrower cross-subsidization and welfare.

In terms of institutions, my paper is related to a growing literature on choices of mortgage upfront closing costs, which are also called points. In this literature, Brueckner (1994) LeRoy (1996), and Stanton and Wallace (2003) present theories of mortgage points that emphasize the role of selection on borrowers' expected prepayment speeds. My empirical work takes this selection seriously and evaluates its welfare implications. Chari and Jagannathan (1989) study the role of insurance to income shocks for the institution of mortgage points, which I also incorporate in my quantitative model. Empirical work on consumer behavior with mortgage points includes Woodward and Hall (2012) who document how points may lead to sub-optimal shopping, Agarwal, Ben-David, and Yao (2017a) who show that many borrowers make the "mistake" of paying too much in points given their predicted refinancing propensities, and Benetton, Gavazza, and Surico (2020) who look at the UK context and finds that lenders may exploit heterogeneity in demand elasticities between rates and points to increase profits.[8]

The rest of this paper is structured as follows. Section 1.2 presents the background about the upfront closing cost and interest rate trade off. Section 1.3 describes the data used in the study. Section 1.4 presents motivating facts. Section 2.3 presents my model and simulation results. Section 1.6 presents estimation results. Section 1.7 describes the counterfactual analyses. Section 2.6 concludes.

1.2 Background

US borrowers face a choice between a mortgage with a higher interest and a lower upfront closing cost or a mortgage with a lower interest rate and a higher upfront closing cost. I illustrate this choice in Figure 2.9, which shows a series of options for rates and upfront closing costs from a

[8] Another strand of literature on mortgage points concerns its role in mortgage discrimination: Bhutta and Hizmo (2020), Bartlett et al. (2019a), and Zhang and Willen (2021).

lender ratesheet. The first column of the table in Figure 2.9 shows the choices of interest rates that are available to a borrower, while the 15 Day, 30 Day, and 45 Day columns show the corresponding upfront closing costs, quoted in percentages of the loan amount, that borrowers would have to pay in order to receive the rate once the loan is originated within the given lock period.[9] The quoted upfront payment to the lender are also called "points." In particular, Figure 2.9 shows how borrowers might choose a mortgage with a lower interest rate by paying more points, or a mortgage with a higher interest rate by paying fewer (or, even, negative) points.[10] Appendix Figure A.12 shows an example of how borrowers were shown a series of rate and upfront closing cost choices from a price comparison website.

---

[9] A rate is "locked" when a lender commits to originating a mortgage with the given terms within the stated lock period of, e.g., 15, 30, or 45 days.

[10] Negative points are possible to cover the other upfront closing costs borrowers may have to pay, such as transfer taxes and application fees.

Figure 1.1: Rate and upfront closing costs options in an example lender rate sheet

Rate	15 Day	30 Day	45 Day
3.500	4.043	4.213	4.303
3.625	2.910	3.080	3.180
3.750	2.104	2.274	2.364
3.875	1.649	1.829	1.919
4.000	0.917	1.097	1.187
4.125	0.238	0.408	0.508
4.250	(0.569)	(0.399)	(0.309)
4.375	(1.122)	(0.952)	(0.862)
4.500	(1.733)	(1.553)	(1.463)
4.625	(2.281)	(2.111)	(2.011)
4.750	(2.835)	(2.665)	(2.575)
4.875	(3.298)	(3.128)	(3.028)
5.000	(3.546)	(3.376)	(3.276)

Not all choices within Figure 2.9 are equally likely to be chosen. Mortgages with low or negative upfront closing costs/points are significantly more popular with borrowers than contracts with high upfront closing costs/points. I show that these low upfront closing cost mortgage contracts generate significant transfers between borrower with different refinancing types, and that the cross-subsidization of these contracts lead to pricing distortions which generate deadweight losses.

This trade-off is more important than it may first appear. As a concrete example, based on Figure 2.9, a borrower with $200k mortgage and $10,300 in closing costs (5.15% of loan amount, which is close to the average) would see their rate increase from 3.5% to 4.375%, a 87.5bps increase, if they were to add all closing costs to the rate.[11] This 87.5bps increase in rate has significant distri-

[11] Going from a 3.5% mortgage at 4.213 points to a 4.375% mortgage at -0.952 points.

butional consequences to the extent that slow to refinance borrowers pay it for a longer period of time than other borrowers. It also has efficiency consequences to the extent that slow to refinance borrowers cross-subsidize other borrowers' price of mortgage origination, leading to excessive refinancing by other borrowers.

In this paper, I characterize mortgages with low or negative upfront closing costs/points as mortgages with their price of mortgage origination added to the rate. To be more precise about the definition of the price of mortgage origination added to the rate, focusing on the setting where lenders are selling the mortgages they originate on the secondary market,[12] I decompose lenders' total origination revenue from making a loan as:

$$\underbrace{\text{lender origination revenue}}_{\text{price of mortgage origination}} = \underbrace{\text{upfront closing costs}}_{\text{paid upfront}} + \underbrace{\text{secondary marketing income}(c)}_{\text{added into rate}} \tag{1.1}$$

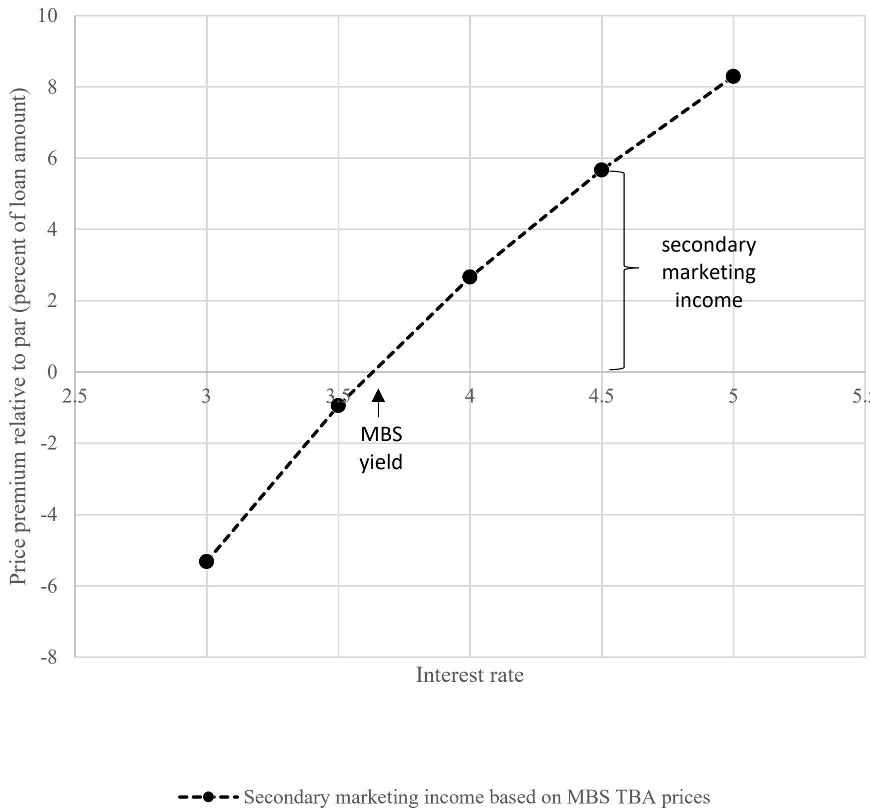
where secondary marketing income(c) refers to the net income lenders derive from selling a loan with interest rate c on the secondary market. The secondary marketing income can be alternatively described as the premium of the mortgage relative to par. To illustrate what the secondary marketing income as a function of interest rates might look like on a given day, Figure 1.2 plots the secondary market value of mortgages based on MBS TBA prices as a percentage of the loan amount at various interest rates.[13] It shows that mortgages with higher interest rates tend to be more valuable

[12] Or, equivalently, where lenders are evaluating the value of their portfolio based on their potential secondary market value.

[13] The TBA market is a highly liquid market where most MBS are traded, and is described in more detail in Vickery and Wright (2013).

on the secondary market and that originating a mortgage with a high enough interest rate generates positive secondary marketing income.

Figure 1.2: Secondary marketing income as a function of interest rates



The price of mortgage origination can be added to the interest rate of the mortgage because a mortgage with a higher interest rate carries a higher monthly payment and is more valuable on the secondary market. This adding to the rate mechanism allows the lenders to charge lower upfront closing costs while maintaining the same origination revenue from making a loan. Indeed, in Appendix A.3 I find that, based on rate sheet data, the pass-through of higher secondary marketing income to lower upfront closing costs to be nearly complete across a range of interest rate choices.

1.3 Data

For my loan-level analyses, I use a combination of three data sets. The first data set is the 2013–2020 data from Optimal Blue on rate locks. Optimal Blue is a rate-locking platform used by lenders constituting about 40% of all U.S. mortgage originations.[14] It contains information about interest rates, upfront closing costs in the form of points paid by the borrower, and time of the lock. Second, I use the 2013–2021 CRISM (Equifax Credit Risk Insight Servicing McDash Database) data, which is an anonymous credit file match from Equifax consumer credit database to Black Knight's McDash loan-level mortgage data set. It contains information on loan performance and a time-varying borrower characteristic in terms of their Equifax Risk Score. The CRISM data also allows me to classify prepayments as moves or refinances.[15] It has been frequently used to study borrower refinancing behavior.[16] Third, I use the 2013–2019 Home Mortgage Disclosure Act (HMDA) data to capture borrower demographics.

I construct two novel matches of these data sets, including a 2018–2019 Optimal Blue-HMDA

[14]Mortgage lenders use rate-locking platforms such as Optimal Blue to assist their loan originators and mortgage brokers in offering rates to their clients. This data set has been used to study issues surrounding rates and points in Bhutta and Hizmo (2020) and in mortgage pricing in Bhutta, Fuster, and Hizmo (2020).

[15]I follow the procedure of Lambie-Hanson and Reid (2018) and Gerardi, Willen, and Zhang (2021) to identify moving by classifying a prepayment as a move if the borrower's address changed within a 6-month window surrounding the prepayment date.

[16]See, e.g., Beraja et al. (2018a), Lambie-Hanson and Reid (2018), Di Maggio, Kermani, and Palmer (2020), Cunningham, Gerardi, and Shen (2021), Abel and Fuster (2021), and Gerardi, Willen, and Zhang (2021).

match and a 2013–2021 Optimal Blue-HMDA-CRISM match, for my empirical analyses. I focus on 30-year, conforming, fixed-rate mortgages for my study due to their status as the most commonly chosen form of mortgage contract in the US.[17] Details of the matching procedure as well as summary statistics can be found in the Appendix A.2.1 and A.2.

Finally, I obtain actual data on the rate and upfront closing cost menus from LoanSifter.[18] Summary statistics and more detailed descriptions of the LoanSifter data are shown in Appendix A.2.3. I show that the rate and upfront closing cost trade-off from LoanSifter on average closely matches the rate and secondary marketing income relationship as implied by MBS TBA prices from Morgan Markets in Appendix A.3.

1.4 Motivating facts

In this section, I present some stylized facts that illustrate the existence of cross-subsidization of mortgage closing costs and its sizable distributional implications. First, I show in Section 1.4.1 that almost all borrowers pay for most of their mortgage closing costs through a higher interest rate on their mortgage relative to mortgage-backed securities yields, rather than upfront. Second, I show that borrowers have heterogeneous refinancing tendencies in Section 1.4.2. Third, I show that the interaction of these two effects means different borrowers with the same closing costs added to the rate end up with very different net present values (NPVs) of their extra interest rate payments, ex

[17] Complex mortgages used to be more common before the financial crisis, but have largely vanished by the start of my sample period (Amromin et al., 2018).

[18] These two data sets have also been used in Fuster, Lo, and Willen (2017) to study the time-varying price of mortgage intermediation.

post, in Section 1.4.3. Fourth, I assess magnitude of this difference by demographic groups in Section 1.4.4. Fifth, I explore the evidence on the selection (and lack thereforeof) of borrowers with different refinancing tendencies into upfront closing cost choices in Section 1.4.5.

1.4.1 Prevalence of mortgages with closing costs added to the rate

When borrowers take out a mortgage, they have a choice between adding closing costs to the rate of the mortgage or paying them upfront. In this section I assess the extent to which mortgage closing costs are added to the rate using the 2018–2019 Optimal Blue-HMDA data. The 2018–2019 HMDA data contains information about the upfront closing costs paid by the borrower in the form of loan origination costs, and the match to Optimal Blue data enables me to obtain information on when the rate was locked which then allows me to estimate the revenue that lenders generate from the secondary market.

I estimate the extent to which mortgage closing costs are added to the rate based on Equation (1.1), which breaks down lenders' total revenue from origination as the sum of upfront closing costs and secondary marketing income. The secondary marketing component of lender revenues is estimated following the procedure of Fuster, Lo, and Willen (2017),[19] where the revenue that lenders generate

[19]The methodology of Fuster, Lo, and Willen (2017) for estimating secondary marketing income involves estimating the premium of an originated mortgage relative to par from MBS TBA prices by subtracting g-fees (the cost of GSE guarantee) from the mortgage interest rate and then using that as the coupon rate, the value of which is then derived using linear interpolation on reported MBS TBA prices between (i) coupons and (ii) trading days.

from the secondary market y as a fraction of the mortgage balance M_{it} is given as:

$$y_{it} = \frac{p_{it}^{TBA+payup}(c_{it} - gfees_t) - M_{it}}{M_{it}} \tag{1.2}$$

where $p_{it}^{TBA+payup}$ is the estimated value of the mortgage on the secondary market based on TBA prices plus "payups," for a coupon rate $c_{it} - gfees_t$ where c_{it} is the interest rate on the mortgage and $gfees_t$ is the price of the government guarantee. Payups are additional amounts that investors pay for an MBS relative to the TBA price for mortgages that have particularly favorable prepayment risk. Low-balance mortgages, for example, are less likely to be prepaid and hence tend to be more valuable in the secondary market. As a result, I add the payups based on mortgage size to the MBS TBA price.[20]

The results of my analysis are shown in Figure 1.3. The left panel in Figure 1.3a shows that lenders make on average 4.6% of the mortgage balance as revenue for each mortgage they originate. This revenue compensates the lender for their costs. First, lenders need to pay for the upfront costs of mortgage insurance, also called loan-level price adjustments (LLPAs) by Fannie Mae and Freddie Mac. Second, lenders pay for loan originator compensation, which can be 1–2% of the loan amount. Third, lenders pay for the underwriting and processing costs associated with the origination. Rela-

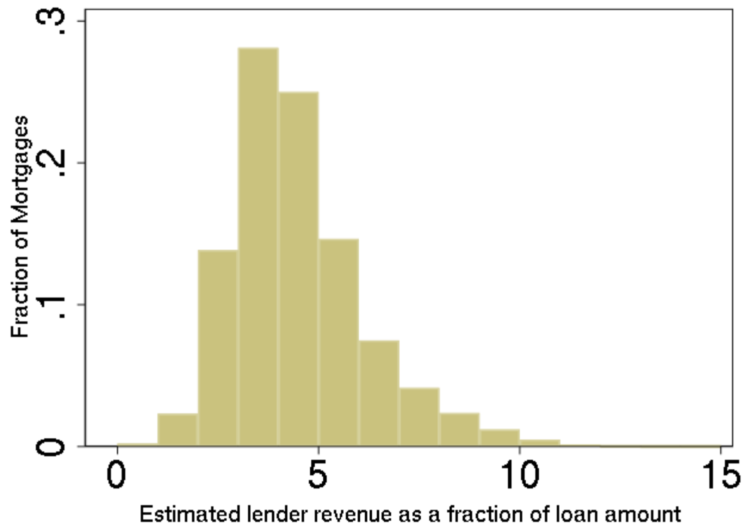
[20]A drawback of this approach of estimating secondary marketing income is that it excludes both the impact of the revenue generated from the sale of mortgage servicing rights and the fees paid to servicers from coupon payments. Fuster, Lo, and Willen (2017) argue that the two effects may approximately cancel each other out. Without explicit data on the value of mortgage servicing rights, I also compute a lower bound on the estimated lender revenues by looking at the MBS value of the net interest rate paid to investors by assuming counterfactually that mortgage servicing rights are worth zero. This lower bound is presented in Appendix Figure A.4, which still shows that the vast majority of mortgages have their closing costs paid for through the rate.

tive to these expenses, the portion that is attributable to accounting profits are low: the Mortgage Bankers' Association (MBA) reports an average production profit of 0.14% of the loan amount in 2018 and 0.31% of the loan amount in 2017.[21]

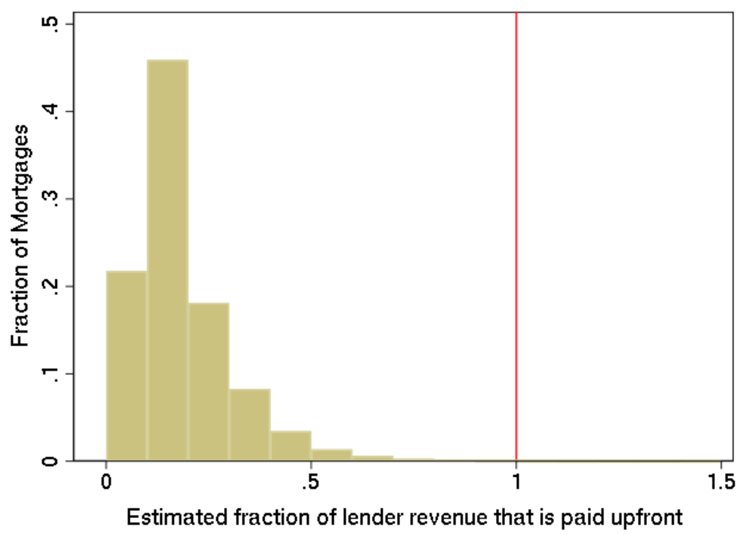
The right panel of Figure 1.3b shows that only a small fraction of lender revenue is paid as upfront closing costs, with an average of 19.5%. That is, even though most of the lender costs of origination are incurred upfront, 80.5% of the price of origination is added to the rate of the mortgage and paid over time primarily by immobile and inactively refinancing borrowers. Hence, almost all mortgages being originated in the US can be considered "low upfront closing cost" mortgages whose price of mortgage origination are prone to cross-subsidization between borrowers with different refinancing speeds.

[21] https://www.mba.org/2019-press-releases/april/independent-mortgage-bankers-production-volume-and-profits-down-in-2018. MBA also reports that average net production revenues in 2018 (excluding LLPAs) are 3.47% of the loan amount, which is consistent with my estimate of 4.6% with LLPAs.

Figure 1.3: Lender revenue and percentage paid as upfront closing costs



(a) Estimated lender revenue



(b) Fraction of lender revenue paid upfront

Conceptually, the empirical observation that lenders make most of their income from secondary marketing revenue is best characterized as closing costs being added to the rate if higher secondary marketing revenue is passed through to consumers as lower upfront closing costs. I present evidence that this is true in Section A.3.

1.4.2 Heterogenous refinancing tendencies

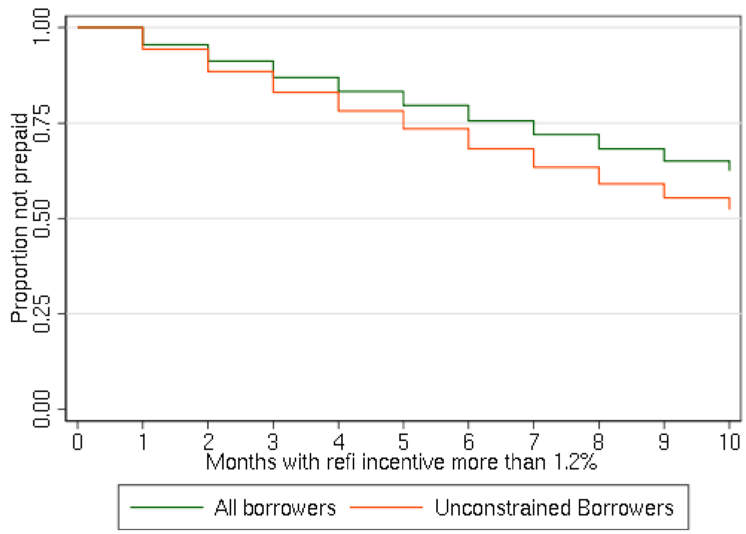
It is well-known that some borrowers are slow to refinance, while others are more quick to refinance when interest rates fall.[22] This is also true in my Optimal Blue-HMDA-CRISM sample. In particular, Figure 1.4 plots the Kaplan-Meier survival hazards of prepayment following months where the interest rate incentive for refinancing, here defined as the decrease in the 30-year Freddie Mac survey rate, is greater than 1.2%, which is larger than the optimal refinancing threshold in typical calibrations of both the Agarwal, Driscoll, and Laibson (2013b) model and my model as presented in Section 2.3. Kaplan-Meier survival hazards are also used to illustrate borrower refinancing behavior in Andersen et al. (2018).

Figure 1.4 shows the results. In particular, more than half of mortgages are not prepaid after 10 months of a relatively high refinancing incentive. While this could be due to supply-side constraints, it also shows that the same pattern holds among a group of borrowers who maintained an Equifax Risk Score of greater than or equal to 700 and an LTV of less than or equal to 80% throughout the sample and are hence unlikely to be unable to refinance due to unemployment, eligibility, or cash flow constraints. Even among this group of borrowers, I find that more half are not prepaid after 10

[22]See, e.g., Archer and Ling (1993), McConnell and Singh (1994), Stanton (1995), Agarwal, Rosen, and Yao (2016), Keys, Pope, and Pope (2016a), Johnson, Meier, and Toubia (2018a), and Andersen et al. (2018).

months of a relatively high refinancing incentive.

Figure 1.4: Kaplan-Meier survival hazards with months of interest rate incentive being greater than 1.2%

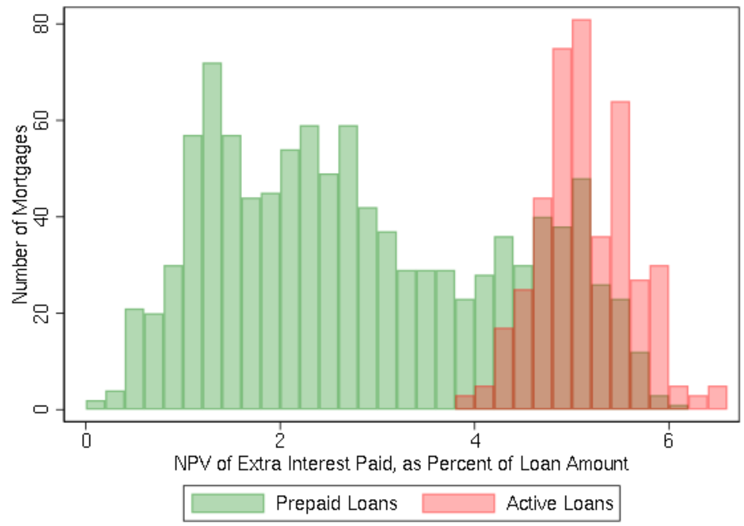


1.4.3 Cross-subsidization of closing costs added to the rate

The interaction of heterogeneity in refinancing tendencies and closing costs added to the rate implies a cross-subsidization of mortgage closing costs. To illustrate this in my data, Figure 1.5 looks at borrowers with similar amounts of closing costs added to the rate (between 4.75-5.25%) in 2013 in my Optimal Blue-HMDA-CRISM sample and compares the NPV of the extra interest rate they paid as a percentage of their loan amount.[23] Due to differences in prepayment behavior, I find large differences in how much borrowers end up paying for the 4.75-5.25% in closing costs they added to the rate, ranging from close to 0% to more than 6%.

[23]The year 2013 was chosen because it is the earliest year in my sample.

Figure 1.5: NPV of extra interest paid, 2013 mortgages with 4.75–5.25% of the loan amount in closing costs added to the rate



The reason for the variance in outcomes in Figure 1.5 is that, when the closing costs are added to the rate of the mortgage, lenders can only recover their closing costs over time through a higher interest rate payment. The principal balance of the mortgage remains unchanged. Therefore, borrowers who prepay earlier end up paying less, while borrowers who prepay later end up paying more. The transfers and deadweight losses studied in this paper come from the extent to which that borrowers who actively refinance pay less for their closing costs in expectation and receive cross-subsidization from other borrowers.

1.4.4 The predictability of cross-subsidization by demographics

Next, I examine the extent of this ex-post cross-subsidization by demographics. To do so, I run the regression on loan level data in my Optimal Blue-HMDA-CRISM sample:

$$NPV_{i,t} = \beta X_i + \gamma Z_i + \xi_{\varphi_{i,t} \times t} + \varepsilon_{i,t} \tag{1.3}$$

where $NPV_{i,t}$ is the NPV of extra interest paid for their closing costs that are added to the rate over the observed life of the mortgage; X_i is a set of demographic and credit utilization variables including race (Black, Hispanic), gender (male and female), credit card revolver status, and quartiles of education; Z_i is a set of control variables including categories of credit scores at origination, LTV, DTI, and log loan amount; $\xi_{\varphi_{i,t} \times t}$ is the amount of closing costs added to the rate by time fixed effects.

The results of this analysis are shown in Figure 1.6 and Table 1.1. I find that Black and Hispanic borrowers paid an extra 0.5% of the loan amount for their closing costs added to the rate relative to other borrowers. For a \$300,000 loan, the magnitude of this cross-subsidization is about \$1500 per loan. Furthermore, single-applicant female borrowers paid an extra 0.24% of the loan amount for their closing costs added to the rate. A limitation of this analysis is that does not take into account the potentially unexpected decline in interest rate during this period, so a model is needed to get at the welfare effects ex ante.

Figure 1.6: NPV of extra interest paid by demographic and borrower characteristics

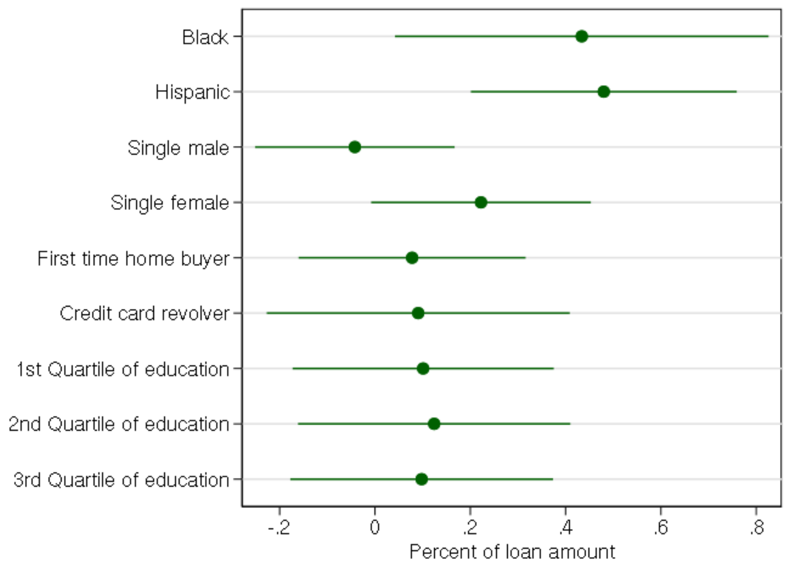


Table 1.1: Regression on NPV of extra interest paid by demographic and borrower characteristics

	(1)	
	NPV of Extra Interest Paid	
Black	0.434***	(2.17)
Hispanic	0.480***	(3.38)
Single male	-0.042	(-0.40)
Single female	0.223**	(1.89)
First-time home buyer	0.078	(0.64)
Credit card revolver	0.091	(0.56)
1st quartile of education	0.101	(0.72)
2nd quartile of education	0.124	(0.85)
3rd quartile of education	0.098	(0.70)
Log(loan amount)	-0.363***	(-2.92)
Credit Score controls	Yes	
LTV controls	Yes	
DTI control	Yes	
Constant	7.918***	(4.85)
Observations	1275	
φ by month FEs	Yes	

Robust t statistics in parentheses.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$

1.4.5 Selection in choices of upfront closing costs

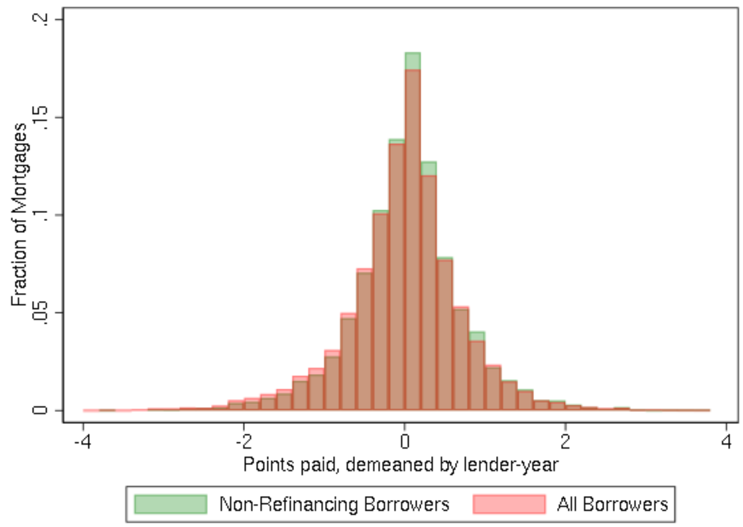
Finally, I examine borrower choices of upfront closing costs in my Optimal Blue-HMDA-CRISM data, paying particular attention to selection by borrower type. If borrowers all know their prepayment types and choose upfront closing costs solely based on their expected prepayment propensities, then there would be no cross-subsidization between borrowers despite heterogeneity in prepayment propensities. The choice of upfront closing costs would serve as a screening device that separates borrowers by type, as described in the models of Brueckner (1994), LeRoy (1996), and Stanton and

Wallace (2003). While I find some selection in the data, I also find evidence of within-choice heterogeneity in ex-post prepayment and refinancing behavior, which leaves room for cross-subsidization.

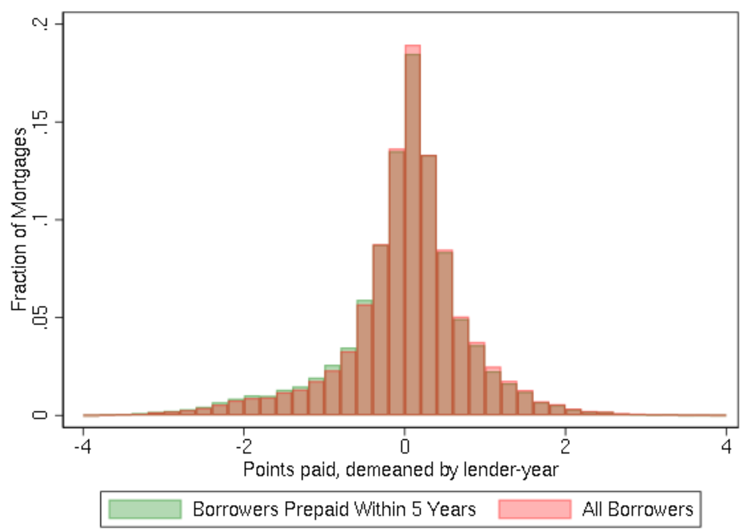
In this section, I measure borrower upfront closing costs in terms of "points," where each point is customarily one percent of the loan amount used to reduce the interest rate. Upfront closing costs consist of points plus an application fee. Negative points, also called "lender credit," that reduce the toal upfront closing costs paid are also possible. The reason I use points rather than upfront closing costs in this analysis is that, unlike the 2018–2019 Optimal Blue-HMDA data, the 2013–2021 Optimal Blue-HMDA-CRISM data contains only information on points and not any other application fees the lender may charge. To the extent that these application fees are constant within lender and loan type, my lender by county by year fixed effects within the sample of 30-year, fixed rate mortgages alleviates the effect of this measurement error.

First, I examine the extent to which borrowers with different prepayment behavior choose different levels of upfront closing costs measured in terms of points. Figure 1.7 plots the distribution of borrower choices of points by their eventual refinancing or prepayment behavior. I define a non-refinancing borrower as one who did not refinance or otherwise prepay within five years despite facing a Freddie Mac Survey Rate decrease of at least 1.4%. As the figure shows, although non-refinancing borrowers on average pay more points, and borrowers who prepay within five years on average pay fewer points, the difference is small in terms of the overall distribution.

Figure 1.7: Points paid by borrower prepayment behavior



(a) Refinancing behavior



(b) Prepayment behavior

To make sure that the result of Figure 1.7 holds even after controlling for underwriting variables, I run an OLS regression of the number of points paid with (1) an indicator function for whether the borrower is a non-refinancing borrower, and (2) an indicator function for whether the borrower prepaid within five years. Results are shown in Table 1.2. Indeed, while I find a positive correlation between non-refinancing borrowers and their payment of points, and a negative correlation between borrowers who prepay within five years and their choices of points, the magnitude of the difference is small at no more than 12 basis points. This analysis suggests that most of the heterogeneity between borrower prepayment behavior remains conditional on choices of upfront closing costs.

Table 1.2: Choices of points and refinancing/prepayment behavior

	(1) Points		(2) Points	
Non-refi borrower	0.094***	(3.91)		
5-year prepayment			-0.122***	(-3.94)
Log(loan amount)	0.121***	(3.54)	0.118***	(3.70)
Credit score controls	Yes		Yes	
LTV controls	Yes		Yes	
DTI control	Yes		Yes	
Constant	-1.419***	(-3.20)	-1.290***	(-3.25)
Observations	3935		3935	
LenderXCountyXYear FEs	Yes		Yes	

Robust *t* statistics clustered by lender and county in parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Next, I present regression estimates of how borrower choices of points correlate with their prepayment behavior with choices of points and prepayment as the dependent variable. The regressions are of the form:

$$\mathbb{1}_{i,t} = \beta X_i + \gamma Z_i + \xi_{l_{i,t} \times c_{i,t} \times t} + \varepsilon_{i,t} \tag{1.4}$$

where as before X_i is a set of demographic and credit utilization variables including race (Black and Hispanic), gender (male and female), credit card revolver status, and quartiles of education; Z_i is a set of underwriting variables including categories of credit scores at origination, LTV, DTI, and log loan amount; $\xi_{l_{i,t} \times c_{i,t} \times t}$ is the lender by county by year fixed effects. I run three regressions of this form with the indicator variable $\mathbb{1}_{i,t}$ being equal to the amount of points paid, whether the mortgage was prepaid within five years, and whether the mortgage was originated by a borrower who failed to refinance despite facing a greater than or equal to 1.4% refinancing rate incentive.

Results are shown in Table 1.3. First, in terms of points, I find that borrowers with a larger loan amount pay more points, and that the correlation is small in terms of other borrower characteristics. The correlation between point choices and predicted prepayment behavior is also weak. For example, Black and Hispanic borrowers are significantly less likely to prepay their mortgage and more likely to be a non-refinancing borrower, but their choices of points are not statistically significantly different from zero compared to the other borrowers.[24]

[24] Bhutta and Hizmo (2020) finds that minority borrowers tend to pay fewer points. The discrepancy in results can be explained by the fact that we focus on conforming mortgages rather than FHA mortgages used in Bhutta and Hizmo (2020), and is explored in more detail in Zhang and Willen (2021).

Table 1.3: Borrower choices of points and their prepayment behavior by characteristics

	(1) Points		(2) 5-year prepayment		(3) Non-Refi Borrower	
Black	0.022	(0.21)	-0.129***	(-3.60)	0.229***	(6.28)
Hispanic	0.022	(0.46)	-0.059**	(-1.89)	0.074***	(3.38)
Single male	0.020	(0.66)	-0.000	(-0.01)	0.006	(0.34)
Single female	-0.019	(-0.51)	-0.022	(-1.01)	0.016	(0.72)
First-time home buyer	0.026	(0.47)	-0.017	(-0.54)	0.026	(0.96)
Credit card revolver	-0.041	(-0.89)	0.080***	(2.37)	-0.022	(-0.58)
1st quartile of education	-0.051	(-1.18)	-0.039	(-0.99)	0.021	(1.14)
2nd quartile of education	-0.060	(-1.05)	-0.010	(-0.38)	-0.016	(-0.60)
3rd quartile of education	-0.019	(-0.41)	-0.007	(-0.19)	-0.031	(-1.10)
Log(loan amount)	0.093***	(2.83)	0.0763***	(3.96)	-0.140***	(-7.21)
Credit score controls	Yes		Yes		Yes	
LTV controls	Yes		Yes		Yes	
DTI control	Yes		Yes		Yes	
Constant	-1.069**	(-2.41)	-0.437*	(-1.95)	2.041***	(10.62)
Observations	3935		3935		3935	
LenderXCountyXYear FEs	Yes		Yes		Yes	

Robust *t* statistics clustered by lender and county in parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Another way to examine selection is to look at how borrower choices of points relate to their moving and refinancing behavior. Points do predict moving and prepayment behavior in a statistically significant manner, which is indicative of some selection being important in this market. To do so, I run the the linear probability model on an indicator variable for moving or refinancing:

$$\mathbb{1}_{i,t}(\text{move/refi}) = \sum_{j=1}^{N} \beta_j \mathbb{1}(\psi_i = j) + \gamma Z_i + \xi_{l_{i,t} \times c_{i,t} \times t} + \varepsilon_{i,t} \tag{1.5}$$

where $\mathbb{1}_{i,t}(\text{move/refi})$ is an indicator variable that is equal to either moving or refinancing; β_j are a set of coefficients on categories of points choices as represented by the indicator function $\mathbb{1}(\psi_i = j)$, and Z_i is a set of controls including the call option value of refinancing from Deng, Quigley, and Van Order (2000a), the spread of the mortgage interest rate at origination to the Freddie Mac Primary Market Survey Rate (spread at origination, or SATO) as well as its square, and the standard set of loan amount, credit score at origination (credit score), loan-to-value ratio (LTV), and debt-to-income ratio (DTI) controls. In particular, the call option value of refinancing is defined as:

$$Call\ Option_{i,k} = \frac{V_{i,m} - V_{i,r}}{V_{i,m}} \tag{1.6}$$

where

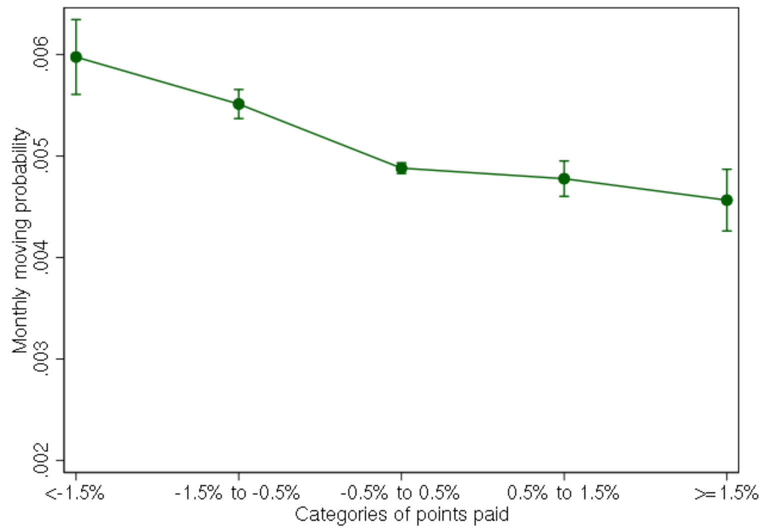
$$V_{i,m} = \sum_{s=1}^{TM_i - k_i} \frac{P_i}{(1+m_{it})^s} \tag{1.7}$$

$$V_{i,r} = \sum_{s=1}^{TM_i - k_i} \frac{P_i}{(1+c_i)^s} \tag{1.8}$$

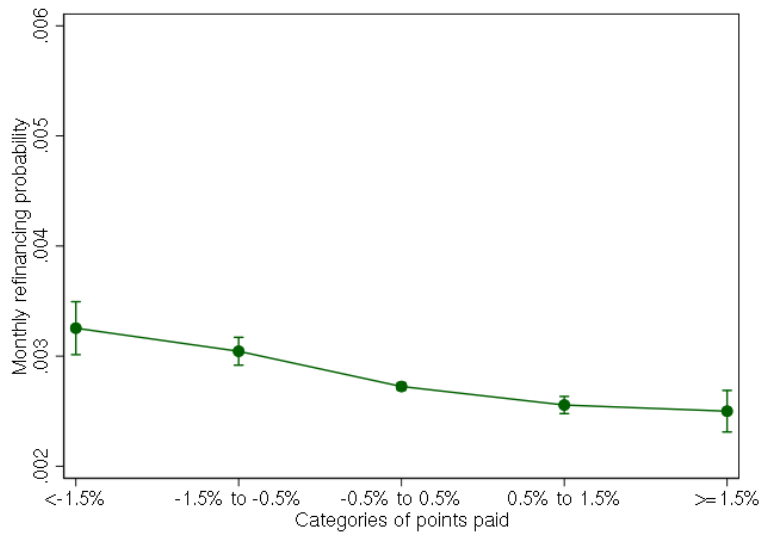
and c_i is borrower i's mortgage rate at origination, TM_i is the mortgage term, k_i is the number of months already past, m_{it} is the Freddie Mac Primary Market Survey Rate, and P_i is the size of the current mortgage payment. The *Call Option* variable represents the potential interest rate savings from refinancing, which is positively correlated with refinancing behavior. Finally, $\xi_{l_{i,t} \times c_{i,t} \times t}$ represents lender by county by year fixed effects, and $\varepsilon_{i,t}$ is the error term.

Figure 1.8 and Table 1.4 present the results. In particular, Figure 1.8a plots the predicted probabilities of moving by categories of points paid in intervals of width 1. It shows that, all else equal, the borrowers' moving hazard is decreasing in the amount of points that they pay, which is consistent with a selection story. Figure 1.8b shows the same pattern but for refinancing. In general, borrowers who pay more points are less likely to move and refinance.

Figure 1.8: Moving/refinancing probability by points paid



(a) Moving probability by points



(b) Refinancing probability by points

Table 1.4 shows the regression coefficients that underlie these results. The regression coefficients show a negative, monotone, and statistically significant relationship between the level of points paid and moving and refinancing probabilities. In terms of control variables, the *Call Option*, spread at origination *SATO*, and log of the loan amount are positively correlated with moving and refinancing.

Table 1.4: Choices of points and refinancing/prepayment behavior

	(1)		(2)	
	Moved		Refi'ed	
-1.5% to -0.5% points	-0.046***	(-2.60)	-0.021	(-1.54)
-0.5% to 0.5% points	-0.110***	(-5.33)	-0.053***	(-3.93)
0.5% to 1.5% points	-0.120***	(-4.95)	-0.070***	(-4.99)
≥1.5% points	-0.141***	(-5.11)	-0.075***	(-4.33)
Call Option	0.986***	(13.01)	1.290***	(17.81)
SATO	0.025	(0.84)	-0.135***	(-3.87)
SATO Sq	-0.131***	(-6.46)	0.063*	(-2.00)
Log(loan amount)	0.204***	(18.99)	0.095***	(9.62)
Credit score controls	Yes		Yes	
LTV controls	Yes		Yes	
DTI control	Yes		Yes	
Constant	-1.940***	(-15.73)	-0.897***	(-6.88)
Observations	8529466		8529466	
LenderXCountyXYear FEs	Yes		Yes	

Robust t statistics clustered by lender and county in parentheses.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$

These patterns imply that a model of cross-subsidization by prepayment type has to take into account both the within-choice heterogeneity in prepayment behavior as well as the selection of borrowers into point choices by their ex ante prepayment expectation. My model accomplishes both of these tasks. In particular, by estimating a distribution of ex ante moving and refinancing types and how they correlate through borrower choices of points, it simultaneously incorporates both selection and within-choice borrower heterogeneity.

1.5 Model

The motivating facts in Section 1.4 show that the institutional feature where mortgage closing costs tend to be added to the rate of the mortgage makes it susceptible to substantial cross-subsidization between borrowers with different refinancing speeds, and that this cross-subsidization has distributional implications. While revealing, this analysis was conducted in a period with declining interest rates which may overstate the magnitudes of the differences in outcomes by demographic groups. I use a structural model to answer the questions: how big are these effects ex ante? And, what are the equilibrium interactions between agents and their welfare implications?

Because borrower refinancing behavior is an important determinant of mortgage interest rates, an equilibrium model that incorporates the supply side (ie. mortgage interest rate) response to heterogeneity in refinancing behavior is needed to get at the welfare questions. As such, I build an equilibrium of mortgage choice that rationalizes the heterogeneity in borrower refinancing behavior and allows me to assess its welfare implications in dollar terms. On the demand side, following the state-of-the-art from Andersen et al. (2018), I estimate a distribution of borrower refinancing costs with two components: a fixed refinancing hassle cost and a time varying ability to refinance. In addition, borrowers differ by their moving probabilities and discount factors. These decisions are then embedded in a workhorse lifecycle model of mortgage choice from Campbell and Cocco (2015) and Chen, Michaux, and Roussanov (2020). A competitive supply side pins down mortgage interest rates at various levels of upfront closing costs and closes the model.

Calibration of the model shows evidence of large cross-subsidization of low upfront closing cost mortgages from slow to refinance borrowers. In addition, the fully estimated model allows me to measure the welfare implications of heterogeneity in borrower refinancing tendencies in equilibrium, which is an independent contribution to the literature that has largely studied this heterogeneity in the reduced form.

1.5.1 Setup

Demand side

On the demand side, households maximize non-housing consumption with time-separable utility with bequest motive for terminal wealth taking housing choice as exogenous:

$$\max \mathbb{E}_1 \sum_{t=1}^{T} \beta_i^{t-1} \frac{(C_{it})^{1-\gamma_i}}{1-\gamma_i} + \beta_i^T b_i \frac{W_{i,T+1}^{1-\gamma}}{1-\gamma}, \tag{1.9}$$

where T is the terminal age, β_i the time discount factor, C_{it} the non-durable consumption, γ_i the coefficient of relative risk aversion, and $W_{i,T+1}$ the real terminal wealth.

In terms of exogenous state transitions, I assume that the risk-free rate r_{1t} follows the model of Cox, Ingersoll, and Ross (1985), which has a natural zero lower bound. I take inflation $\pi = 1.68\%$ as a constant equal to the average in my sample.[25] Real (log)labor income L_{it}, house price H_{it}, and

[25] Inflation expectations were stable over my sample period, and a constant term for inflation al-

changes in the mortgage interest rate at an average level of upfront closing costs $\Delta\bar{c}_t$ are modelled as a vector auto-regression (VAR) with the risk-free rate r_{1t} as an exogenous covariate, the details of which are described in Appendix A.5.1. Finally, moving is treated as an exogenous mortgage refinance at an average level of upfront closing costs.

In each period, households make a decision of whether to refinance along with a consumption and savings decision. In doing so, they face financial constraints in the sense that their savings $S_{it} \geq 0$. They make a real mortgage payment P_{it}^M and earn interest r_{1t} on savings minus inflation π_t, and so in non-refinancing periods their non-durable consumption C_{it} in real terms can be written as:

$$C_{it} = \exp(L_{it}) - P_{it}^M + (r_{1,t-1} - \pi_t)S_{it-1} + \Delta S_{it} \tag{1.10}$$

Where $\Delta S_{it} = S_{it} - S_{it-1}$ is the change in the borrower's savings. In order to refinance, borrowers need to pay a cost $\tilde{\kappa}_{it}$. I model the borrowers' refinancing cost $\tilde{\kappa}_{it}$ as:

$$\tilde{\kappa}_{it} = \begin{cases} \infty, \text{ with probability } 1 - p_i^a \\ \\ \kappa_i, \text{ with probability } p_i^a \end{cases} \tag{1.11}$$

where p_i^a is the probability that a borrower is able to refinance in a particular time period. The inclu-

lows me to easily convert the nominal mortgage payment from the amortization table to real terms. In particular, the real mortgage payment under constant inflation is $P_{it}^M = \frac{1}{(1+\pi)^t} M_i \frac{c_{it}/12(1+c_{it}/12)^n}{(1+c_{it}/12)^n - 1}$. Note that because M_i is fixed, this formula does not incorporate the slight differences in the speed of principal paydown as borrowers refinance to mortgages with different interest rates. This feature is standard for the literature and the error resulting from it is likely small for minor differences in rates.

sion of time- and state-varying refinancing costs is necessary to fit the data where borrowers do not immediately refinance when facing their cut-off, as described in Andersen et al. (2018) which uses a similar setup for capturing refinancing costs.

Furthermore, I require that the refinance must leave the borrower a loan-to-value (LTV) ratio of at most 95%, which is required by Freddie Mac[26] and captures the constraints to refinancing in periods of house price decline as described in Hurst et al. (2016).

The full value function $V_{it}(c_{it}, S_{i,t-1}, \bar{c}_t, r_{1,t-1}, H_{it}, H_{i,t-1}, L_{it})$ is a function of the state variables interest rate on the mortgage c_{it}, last period savings $S_{i,t-1}$, the current market interest rate $\bar{c}_t$, last period's risk-free rate $r_{1,t-1}$, house price H_{it}, lagged house prices $H_{i,t-1}$, labor income L_{it}. Of these variables, c_{it}, S_{it} are endogencous in that they are influenced by the decision to refinance and borrower's consumption decision, while the other states are exogenous. In what follows I write the value function $\tilde{V}_{it}(c_{it}, S_{it}) = V_{it}(c_{it}, S_{it}, \bar{c}_t, r_{1,t-1}, H_{it}, H_{i,t-1}, L_{it})$ as a function of the endogenous variables only for brevity.

When first getting a mortgage, borrowers make a choice of mortgage interest rate c along with their associated upfront closing cost $\psi_{it}(c)$ to maximize their expected utility in the first period:

$$\mathbb{E}_1 U_{i1} = \max_{\Delta S_{i2}, c} \frac{(\exp(L_{i1}) - (\tilde{\kappa}_{i1} + \psi_{it}(c)M) - \Delta S_{i2})^{1-\gamma_i}}{1-\gamma_i} + \beta \mathbb{E}_1 \tilde{V}_{i2}(c, S_{i2}) \qquad (1.12)$$

[26]Freddie Mac's requirements for refinancing are described in https://sf.freddiemac.com/general/maximum-ltv-tltv-htltv-ratio-requirements-for-conforming-and-super-conforming-mortgages. Fannie Mae has a slightly looser LTV requirement of at most 97%: https://singlefamily.fanniemae.com/media/20786/display.

In the following periods, borrowers make a mortgage payment $P^M(c_{it})$. And in periods where the borrower is able to refinance, their utility can be written as the maximum of what can be obtained by refinancing and not refinancing:

$$\mathbb{E}_t U_{it}^a = \max \begin{cases} \max_{\Delta S_{it}} \frac{(\exp(L_{it}) - P^M(c_{it}) + (r_{1,t-1} - \pi_t) S_{it-1} - \Delta S_{it})^{1-\gamma_i}}{1-\gamma_i} + \beta \mathbb{E}_t \tilde{V}_{i,t+1}(c_{it}, S_{it}) \\ \max_{\Delta S_{it}, c} \frac{(\exp(L_{it}) - P^M(c_{it}) - (\tilde{\kappa}_{it} + \psi_{it}(c) M) + (r_{1,t-1} - \pi_t) S_{it-1} - \Delta S_{it})^{1-\gamma_i}}{1-\gamma_i} + \beta \mathbb{E}_t \tilde{V}_{i,t+1}(c, S_{it}) \end{cases} \tag{1.13}$$

where the first line of Equation (1.13) corresponds to the borrower's utility from not refinancing and continuing to get the interest rate c_{it}, while the second line corresponds to the borrower's utility from refinancing to the rate c which affects the upfront closing cost they pay $\psi_{it}(c)$.

Similarly, the borrower's utility given that they are not able to refinance is:

$$\mathbb{E}_t U_{it}^{na} = \max_{\Delta S_{it}} \frac{(\exp(L_{it}) - P_{it}^M - (r_{1t} - \pi_t) S_{i,t-1} - \Delta S_{it})^{1-\gamma_i}}{1-\gamma_i} + \beta \mathbb{E}_t \tilde{V}_{i,t+1}(c_{it}, S_{it}). \tag{1.14}$$

Finally, I model moving as an exogenous costless refinance to the new mortgage with an interest rate $\bar{c}_t$ that is associated with an average level of closing costs, which occurs with probability p_i^m for borrower i. Therefore, the borrower's utility upon moving is:

$$\mathbb{E}_t U_{it}^m = \max_{\Delta S_{it}} \frac{(\exp(L_{it}) - P_{it}^M - (r_{1t} - \pi_t) S_{i,t-1} - \Delta S_{it})^{1-\gamma_i}}{1-\gamma_i} + \beta \mathbb{E}_t \tilde{V}_{i,t+1}(\bar{c}_t, S_{it}). \tag{1.15}$$

Combined, the value function of the borrower can be written as:

$$\mathbb{E}_t V_{it} = (1 - p_i^m)(p_i^a \mathbb{E}_t U_{it}^a + (1 - p_i^a)\mathbb{E}_t U_{it}^{na}) + p_i^m \mathbb{E}_t U_{it}^m. \tag{1.16}$$

Supply side

A supply side to the model is needed compute mortgage premia with counterfactual mortgage contract designs. I assume that the supply side is perfectly competitive and that lenders set the rate and upfront closing cost/points trade-off based on the MBS value of mortgages. That is, in equilibrium the relationship between the upfront closing costs paid as a fraction of the loan amount ψ_{it} for borrower i at time t and the mortgage interest rate c is pinned down by a zero profit condition:

$$\pi_{it}^l = \psi_{it} M + \varphi_t(c) M - \bar{m}_t^l - m_i^l(M) = 0 \tag{1.17}$$

where π_{it}^l is lender profit from a originating loan to borrower i at time t, $\varphi_t(c)$ is the MBS premium of the mortgage as a percent of the loan amount at the time of origination, and $\bar{m}_t^l$ is average marginal cost incurred by the lender for originating the loan, and $m_i^l(M)$ is the borrower and loan amount specific marginal cost incurred by the lender for originating the loan. Assuming that the marginal cost of loan origination $\bar{m}_t^l + m_i^l(M)$ does not vary by the borrower's choice of points, we have by re-arranging:

$$\psi_{it}(c) = \frac{\bar{m}_t^l + m_i^l(M)}{M} - \varphi_t(c). \tag{1.18}$$

So that, all else equal, a mortgage with a higher interest rate c and MBS value $\varphi_t(c)$ would require fewer upfront points ψ_{it}. In particular, my model implies that the MBS value of mortgages with a higher interest rate will be passed-through to borrowers in terms of lower upfront closing costs. This is approximately true in reality, as I show in Figure A.2.

To close the model, I estimate the MBS value of mortgages $\varphi_t(c)$ based on an expected NPV method where the cashflows from MBS are assumed to be discounted using the risk-free rate r_{1t} plus an option-adjusted spread (*OAS*) term that compensates for the the liquidity and prepayment risk. The *OAS* has been used and evaluated as a proxy for expected MBS returns in Gabaix, Krishnamurthy, and Vigneron (2007), Song and Zhu (2018), and Boyarchenko, Fuster, and Lucca (2019), and Diep, Eisfeldt, and Richardson (2021).[27] More specifically, I compute investor cashflows given an empirical prepayment hazard function $\hat{p}_{t'}$ and a cumulative remaining balance $\hat{q}_{t'} = \prod_{j=t}^{t'-1}(1-\hat{p}_j)$. When a borrower prepays, the lender gets remaining principal $B^M_{t'}$. Otherwise, they get a payment $P^M(c)$.[28] The MBS premium of the mortgage, in dollar value terms, is then:

$$\varphi_t(c)M = E_t \sum_{t'=t}^{t+T} \delta_{t'}\hat{q}_{t'}[(1-\hat{p}_{t'})P^M(c) + \hat{p}_{t'}B^M_{t'}] - M \tag{1.19}$$

where the discount factor is based on the cumulative risk-free rate in period j, r_{jf}, plus an estimated

[27] Another method of valuing MBS is via multivariate density estimation, as in Boudoukh et al. (1997), but that does not allow me to get counterfactual prices under alternative prepayment behavior or with alternative mortgage contract designs.

[28] $P^M(c)$ is the nominal version of the real mortgage payment used in the demand model.

OAS term that compensates for liquidity and prepayment risk:

$$\delta_{t'} = \frac{1}{\prod_{j=t}^{t'}(1 + r_{jf} + OAS)}. \tag{1.20}$$

Such that, conditional on empirical prepayment hazards $\hat{p}_t(M)$, the *OAS* is the only free parameter in the supply side model. In particular, with the model of $\varphi_t(c)$, the mortgage interest rate at an average level of upfront closing costs $\bar{c}_t$ and the level of the risk-free rate r_{1t} pins down $\bar{m}_t^l + m_t^l(M)$ exactly and allows me to recover the mortgage interest rate and upfront closing cost trade-off for all levels of upfront closing costs and with counterfactual cash flows (ie. the no cross-subsidization case, or with alternative mortgage contract designs) while holding $\bar{m}_t^l + m_t^l(M)$ as fixed. Details of the OAS estimation is shown in Appendix A.5.2.

Combined with the demand side, my model can be viewed as an equilibrium model of mortgage premia, in line with Campbell and Cocco (2015) and Campbell, Clara, and Cocco (2021), but with the addition of heterogenous borrower refinancing costs and endogenous upfront closing costs. A key assumption in this model is the perfectly competitive supply side. If the supply side were not perfectly competitive as is assumed here, my counterfactual results would still hold if lenders charge a constant markup across loans but should be interpreted in terms of consumer welfare instead of social welfare.

1.5.2 Computation

I solve the household life-cycle model using a value function approximation approach that is novel to this literature. In particular, instead of discretizing a state space S, I approximate:

$$V(S) \approx \hat{V}(S) \tag{1.21}$$

where $\hat{V}(S)$ is estimated using a Ridge regression on a fully interacted third order polynomial expansion of all the state variables, where the ridge parameter is chosen via cross-validation. The use of a sieve approximation to conduct value function iteration is proven to be consistent in Arcidiacono et al. (2013) and has been applied in other empirical contexts.[29] The advantage of this approach is that it allows value functions on relatively large state spaces to be computed tractably.

1.5.3 Cross-subsidization by upfront closing cost choice: a calibration

Using the model, I illustrate the cross-subsidization of low upfront closing cost mortgages from the perspective of a quick to refinance borrower through a calibration. First, I show in Section 1.5.3 that, in the model environment under the prevailing rate and upfront closing cost trade-off, quick to refinance borrowers benefit from getting mortgages with lower upfront closing cost, whereas the same is not true in a counterfactual environment where lenders are able to price based on borrower refinancing types. Section 1.5.3 shows that this is because low upfront closing cost mortgages are

[29] See, e.g., Keane and Wolpin (1997) and Barwick and Pathak (2015).

cross-subsidized for quick to refinancing borrowers from a pricing perspective. Section 1.5.3 shows that these borrowers are incentivized to refinance more when they get lower upfront closing cost mortgages. Finally, Section 1.5.3 summarizes the economic intuition of the cross-subsidization from a transfer and efficiency perspective.

All of the analysis in this section is conducted for a calibrated borrower with parameters described in Table 1.5, where β, M, p^m are the median of the estimates from Section 1.6, OAS is as estimated in Appendix A.5.2, and $p_i^a = 1, \kappa_i = 0$ are chosen to represent the behavior of an optimally refinancing borrower.

Table 1.5: Parameters for an illustrative calibration of cross-subsidization from the perspective of a quick to refinance borrower

Parameter	Value
β_i	0.92
γ_i	2
M_i	\$223,784
p_i^m	0.12
κ_i	0
p_i^a	1
Initial liquid assets	\$50,000
Initial risk-free rate	0.80%
Initial mortgage rate	3%
Initial income	\$72,000
Initial house price	\$300,000
OAS	0.22%

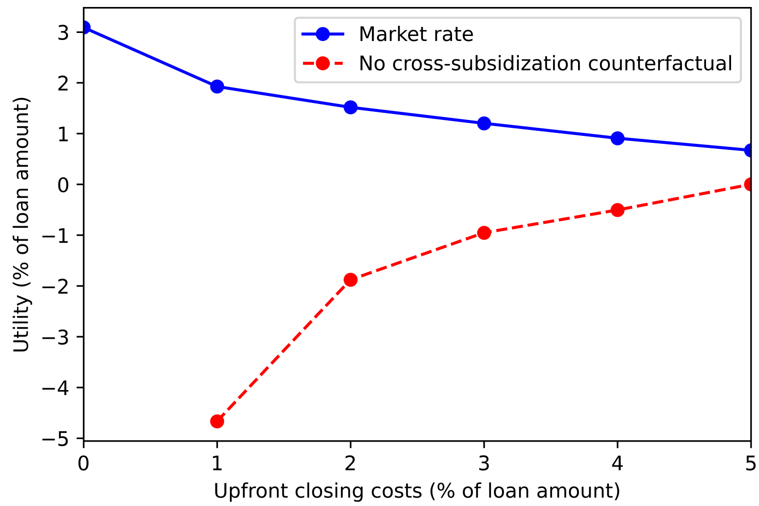
While the set of parameters in Table 1.5 bakes in a lot of assumptions, I note that my calibration results are not overly sensitive to the choice of parameters or even the model in general: in Appendix A.7 I show that essentially the same results can be obtained by using a different model of Agarwal, Driscoll, and Laibson (2013b) with Brownian motion interest rates. The results that are indicative of a significant cross-subsidization of low upfront closing cost mortgage for quick to refinance borrowers emerge in both specifications.

Borrower utility by upfront closing cost choice

I present the results on borrower utility by upfront closing cost choice in Figure 1.9. In particular, for the set of borrowers with parameters described in Table 1.5, I plot the calibrated quick to refinance borrower's utility when they choose a given level of upfront closing cost both during the initial period and whenever they refinance their mortgage. The solid line represents this utility under the market exchange rate between the mortgage interest rate and upfront closing costs. The fact that this line is downward sloping suggests that the calibrated quick to refinance borrower gains from choosing a lower upfront closing cost mortgage.

To compare with a case where there is no cross-subsidization, I iterate the borrower and lender problems jointly using backward induction, where the supply side condition implied by Equation (1.18) from the last period with the borrower's model-implied prepayment probabilities are used as the rate and upfront closing cost trade-off that borrowers face in the current period. This allows me to get the borrower value and equilibrium interest rates in the no cross-subsidization world. I plot the borrower utility under the no cross-subsidization rate in the dashed line of Figure 1.9.

Figure 1.9: Calibrated borrower utility by upfront closing cost choice, $p_i^a = 1, \kappa_i = 0$



As shown in Figure 1.9, the calibrated quick to refinance borrower benefits from choosing a low upfront closing cost mortgage under the market exchange rate between interest rate and upfront closing cost mortgages.[30] However, this is not the case in the counterfactual environment where lenders are able to price based on the quick to refinance borrower's specific type, when the borrower actually benefits from paying more in terms of upfront closing costs. The reason for the difference between the solid and dashed lines is due to the pricing effect of the cross-subsidization. In particular, quick to refinance borrowers face an interest rate on lower upfront closing cost mortgages that is significantly lower than otherwise. The following section analyzes this effect in more detail.

[30]This is consistent with Dave Ramsey's financial advice, which says that: "most buyers won't regain their money on mortgage points because they usually refinance, pay off, or sell their homes before they reach their break-even point."

The pricing effect of cross-subsidization

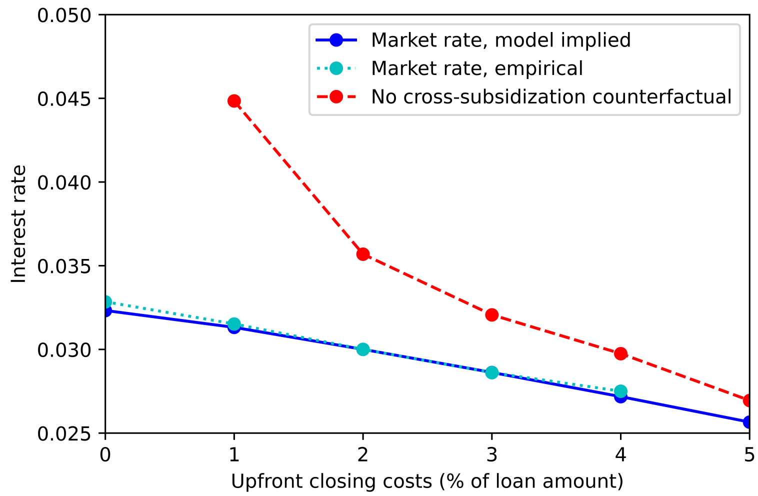
Figure 1.10 illustrates this pricing impact of cross-subsidization by plotting the implied interest rates from the joint iteration of borrower and lender values in the dashed line. The market rate and upfront closing cost rate-off as implied by the model is shown in the solid line, and the empirical rate and upfront closing cost trade-off is presented in the dotted line.

As Figure 1.10 shows, the interest rate trade-off is higher, and steeper, for the calibrated quick to refinance borrower in the no cross-subsidization counterfactual. This suggests that the market interest rate for low upfront closing cost mortgages is especially lower than the no cross-subsidization case due to the cross-subsidization from slow to refinance borrowers. In terms of numbers, I find that a mortgage with a one percent upfront closing cost would carry a 1.25% higher interest rate in the no cross-subsidization case relative to the existing market equilibrium, whereas the difference is only 0.26% for a mortgage with a four percent upfront closing cost.[31]

---

[31] The effect of the presence of inactively refinancing borrowers in reducing the interest rate on low upfront closing cost mortgages has some parallel in Handel (2013), where high switching cost consumers help stabilize the market for health insurance.

Figure 1.10: Market interest rate vs no cross-subsidization counterfactual interest rate

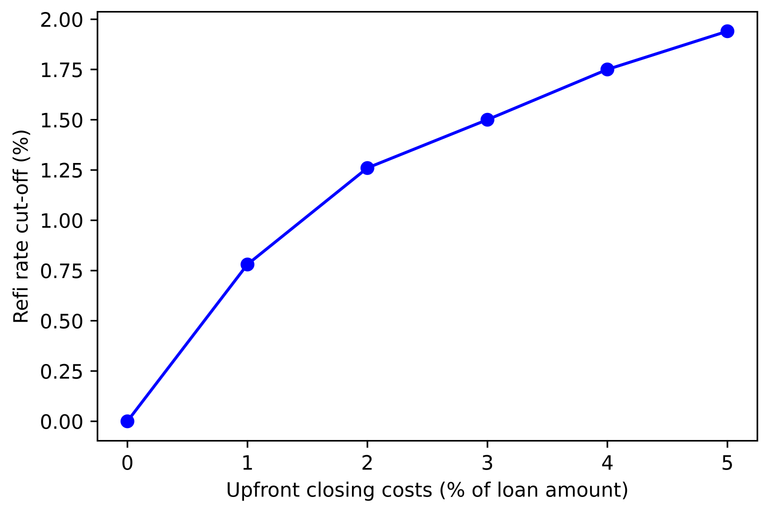


Optimal refinancing strategy by upfront closing costs

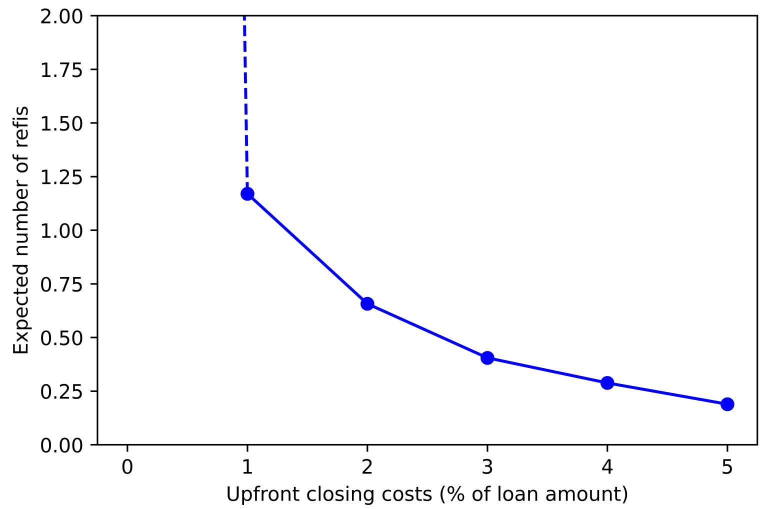
Next, I examine the source of the quick to refinance borrower's change in behavior after getting a low upfront closing cost mortgage. In general, a lower upfront closing cost implies that the quick to refinance borrower is incentivized to increase their the number of refinances: by getting a lower upfront closing cost mortgage and refinancing whenever interest rates fall, quick to refinance borrowers are able to gain a utility advantage. This is illustrated in Figure 1.11.

Figure 1.11: Optimal refinancing strategy by upfront closing costs

(a) Refi cutoff



(b) Expected number of refinances



In particular, Figure 1.11a shows that, for an actively refinancing borrower, the optimally refinancing rate cut-off is much lower when the borrower gets a low upfront closing cost mortgage.

This is because the expected interest rate savings outweigh the upfront closing costs paid. In turn, this leads to more refinancing than otherwise, as shown in Figure 1.11b. As the earlier Figure 1.9 shows, this extra refinancing only profitable for the quick to refinance borrower to undertake due to the cross-subsidization of low upfront fee mortgages.

Economic intuition on transfers and inefficiencies

I showcase the economic intuition behind these results in Figure 1.12, where I plot illustrative demand curves for mortgage originations for a non-refinancing borrower and an actively refinancing borrower. The demand curve for a non-refinancing borrower is vertical, representing that their quantity of upfront closing is fixed and due to exogenous factors (e.g., moving). The demand curve for an actively refinancing borrower is downward sloping in the price, representing the fact that an actively refinancing borrower would refinance more if the price of originations is lower, as the interest rate savings from refinancing become higher than the price of a new origination.

The social marginal cost of mortgage origination is represented as a solid horizontal line. For non-refinancing borrowers, the price they face is this cost shifted upwards as the cost of origination gets added to the rate and they end up paying more for each origination, which is illustrated in Figure 1.12a. For actively refinancing borrowers, their effective price of mortgage origination is shifted downwards from the social cost, as illustrated in Figure 1.12b. An important distinction between the two panels is in the change of borrower behavior. To the extent that actively refinancing borrowers originate more mortgages than they otherwise would due to this cross-subsidization, they introduce a social deadweight loss represented by the triangle indicated by the arrow in Figure 1.12b.

Figure 1.12: Deadweight loss from cross-subsidization of the price of mortgage refinancing

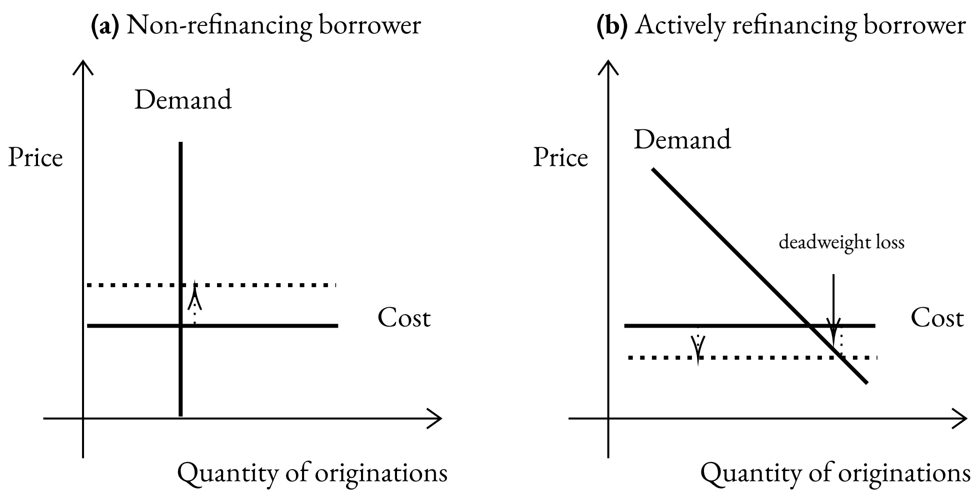


Figure 1.12 may also be interpreted in terms of price elasticities. To the extent that non-refinancing borrowers' quantity of mortgage origination are less price elastic, the effect of their cross-subsidization involves less of a change in behavior. Therefore, the economic distortions in the model mostly attributed to the changes in the incentives faced by the actively refinancing borrowers who refinance excessively.

1.6 Estimation

To estimate the model, I allow $p_i^a, \kappa_i, \beta_i, p_i^m, M_i$ to vary by individual, where p_i^a is the probability that an individual is available to refinance in a particular time period, κ_i is the individual's refinancing hassle cost when they do refinance, β_i is the discount factor, p_i^m is the individual's moving probability, and M_i is the individual's mortgage size. I fix the coefficient of risk aversion $\gamma = 2$, liquid savings at origination to \$50k, and a bequest motive of $b = 200$ in accordance with Campbell and Cocco (2015). To maintain comparability to the TBA market, I further restrict my analysis to 30 year purchase mortgages with a balance above \$150k, FICO above 680, and LTV below 85% following Fusari et al. (2020).

I first present the identification argument in Section 1.6.1, then estimation procedure in Section 1.6.2, then results in Section 1.6.3, some calibration based on my estimates in Section 1.5.3, and finally the implications of my estimates for transfers and welfare in Section 1.6.4.

1.6.1 Identification

Of the unknown parameters, the distribution of M_i is observed. I discuss the identification for the distribution of p_i^a, κ_i, β_i, p_i^m as follows. First, the time-varying ability to refinance p_i^a and hassle costs κ_i are separately identified from borrower responses to the time series movement of the interest rate incentive. More specifically, if the only heterogeneity in borrower refinancing behavior were due to hassle costs, borrowers would refinance immediately when their refinancing cutoff is reached. This is rejected in the data as many borrowers wait long after the interest rate has fallen to their eventual refinancing rate, suggesting that a time-varying refinancing cost is at play. This line of reasoning is also used in Andersen et al. (2020a).

Of the other parameters, ex ante moving probabilities p_i^m are identified from the interaction between the interest rate incentive and borrower refinancing behavior. In particular, borrowers who do not refinance when faced with a large interest rate incentive are more likely to subsequently move. This suggests that moving is not just an ex post shock and that there is heterogeneity in moving expectations ex ante. Finally, conditional on refinancing and moving probabilities, discount factors β_i are identified from borrower choices of upfront closing costs. In general, because upfront closing costs involve an initial outlay, they are more attractive to borrowers with a higher discount factor. The choices of borrowers who choose low upfront closing cost mortgages despite being unlikely to refinance or move are rationalized with a lower discount factor.

1.6.2 Parametrization

I estimate the distribution of the borrower types using mortgage performance data. More specifically, I use a Logit-Normal distribution[32] to model p_i^a, β_i, p_i^m, a Log-Normal distribution to model κ_i, and allow p_i^a, β_i to be correlated via a coefficient ρ. The precise parametrization is as follows:

$$\begin{bmatrix} p_i^a \\ \beta_i \end{bmatrix} \sim Logit\left(MultivariateNormal\left(\begin{bmatrix} \mu_{p^a}(b,h) \\ \mu_\beta \end{bmatrix}, \begin{bmatrix} \sigma_{p^a}^2 & \rho\sigma_{p^a}\sigma_\beta \\ \rho\sigma_{p^a}\sigma_\beta & \sigma_\beta^2 \end{bmatrix} \right) \right) \tag{1.22}$$

$$p_i^m \sim Logit(Normal(\mu_{p^m}(b,h), \sigma_{p^m})) \tag{1.23}$$

$$\kappa_i \sim LogNormal(\mu_\kappa(b,h), \sigma_\kappa) \tag{1.24}$$

where $\mu_{p^a}(b,h), \mu_{p^m}(b,h), \mu_\kappa(b,h)$ can depend on a Black and Hispanic dummy represented by b

[32] The Logit-Normal distribution is the distribution generated by $Y = \frac{\exp(X)}{1+\exp(X)}$ with a normally distributed X. This formulation allows me to model observations that are between zero and one, as well as correlations between them, in closed form.

and h, respectively. This gives me 15 parameters:

$$\theta = (\mu_{p^a}(0,0), \mu_{p^a}(1,0), \mu_{p^a}(0,1), \sigma_{p^a}, \mu_{p^m}(0,0), \mu_{p^m}(1,0),$$
$$\mu_{p^m}(0,1), \sigma_{p^m}, \mu_{\kappa}(0,0), \mu_{\kappa}(1,0), \mu_{\kappa}(0,1), \sigma_{\kappa}, \mu^{\beta}, \sigma_{\beta}, \rho)$$

to estimate. I focus on the correlation ρ between a borrower's probability of being able to refinance and their discount factor because variation in the distribution of κ is small. Intuitively, this is because when borrowers do refinance, they tend to do so for relatively low interest rate savings (ie. in the range of 1%), which would not be reconcilable with a high refinancing hassle cost κ. Therefore, time-varying ability to refinance appears more important in the data, and I also estimate its correlation with the borrowers' discount factors.

In the data, I observe borrowers' prepayment decisions which combines moving and refinancing.[33] I construct the likelihood based on prepayment decisions, which implicitly treats all non-model implied refinancing as a move. Therefore, the moving probability p_i^m in my model captures all exogenous prepayment. The likelihood function for a prepayment decision y_{jt} for loan j at time t given a set of parameters $x_i = \{p_i^a, \kappa_i, \beta_i, p_i^m, M_i\}$ is then:

$$l_{jt}(x_i) = (1 - y_{jt})^{1-p_{jt}(x_i)} y_{jt}^{p_{jt}(x_i)}. \tag{1.25}$$

Furthermore, at time $t = 0$, the likelihood of observing the borrower with 's choice of upfront

[33] I also separately observe moving and refinancing decisions for a subset of prepayments.

closing costs ψ_{i0} that is equal to the optimal choice implied by the model of $\psi_0^*(x_i)$:[34]

$$l_j'(x_i) = \mathbb{1}(\psi_{i0} = \psi_0^*(x_i)). \tag{1.26}$$

To estimate the model, I simulate individuals with a grid for $x_i = \{p_i^a, \kappa_i, \beta_i, p_i^m, M_i\}$ based on a set of parameters θ, with $x_i \sim \mathcal{F}(\theta)$ where $\mathcal{F}(\theta)$ is the distribution of types from Equations (1.22) to (1.24). I then get their model implied optimal point choices ($\psi^*(x_i)$, in whole numbers from -2 to 2) and time-varying prepayment (i.e., refinancing and moving) decisions for each loan-time observation $p_{jt}(x_i)$, and search for the set of parameters that maximizes the likelihood of the data following the standard maximum likelihood formulation:

$$\mathcal{L} \propto \sum_j \log \left(\sum_{i=1}^{nsim} l_j'(x_i) \prod_{t=1}^{T_j} l_{j,t}(x_i) \right), x_i \sim \mathcal{F}(\theta), \tag{1.27}$$

where $nsim = 2000$ is the number of simulations used to compute the likelihood function.

1.6.3 Results

In this section I present my estimates for the distribution of borrower types in the population. The hyper-parameters and their standard errors are shown in Appendix Table A.7, and I plot their distributions in the rest of this section.

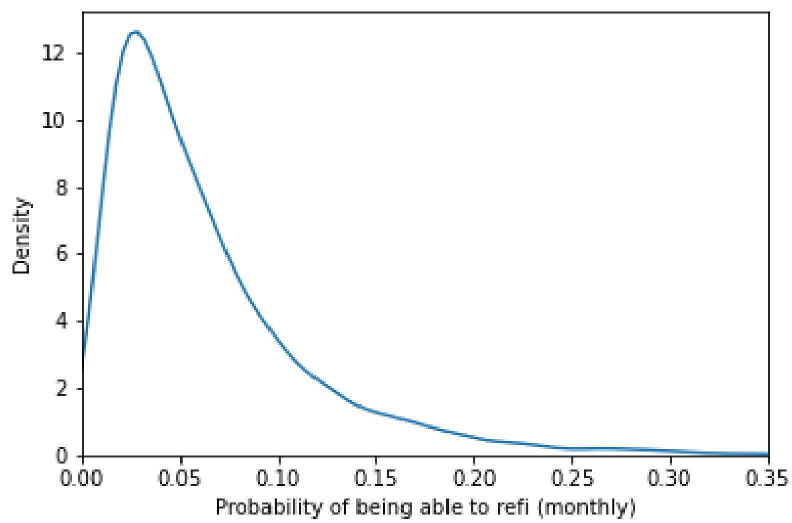
Figure 1.13 presents the estimates on the distribution of refinancing types in the population. In

[34]Since I only observe points and not application fees prior to 2018, I assume a real application fee of $2000 following Agarwal, Driscoll, and Laibson (2013b).

the left panel in Figure 1.13a, results show that most borrowers have a low probability of being able to refinance in a particular month, with some variance. Mean able-to-refinance probability is 6.0% monthly, or 52% annualized. This is consistent with my stylized fact in Section 1.4.2 showing that around half of all borrowers fail to refinance following ten months of a relatively high refinancing incentive. In the right panel in Figure 1.13b, the results show that the implied hassle cost of refinancing for most borrowers is low. Taken together, the results suggest that most of the inaction in refinancing is due to a Calvo-style time-varying ability to refinance rather than hassle costs. The identification in the data is that borrowers who eventually refinance tend to do so at relatively low interest rate savings (for example, at around 1%), which implies a low hassle cost for refinancing for most borrowers despite a time-varying inability to do so.

Figure 1.13: Distribution of borrower refinancing types

(a) Probability of being able to refi



(b) Hassle cost for refinancing

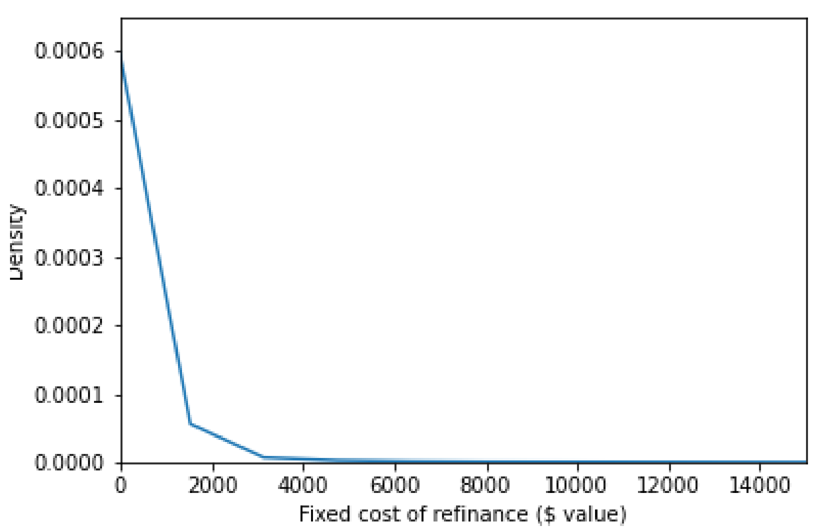
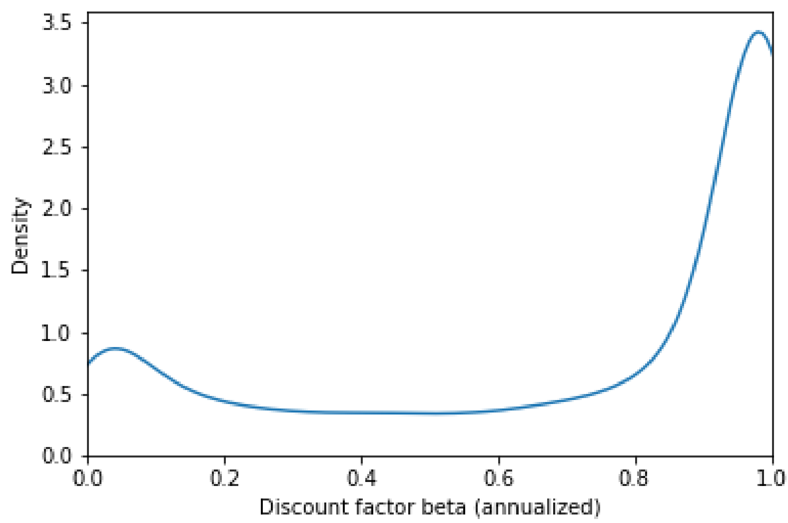


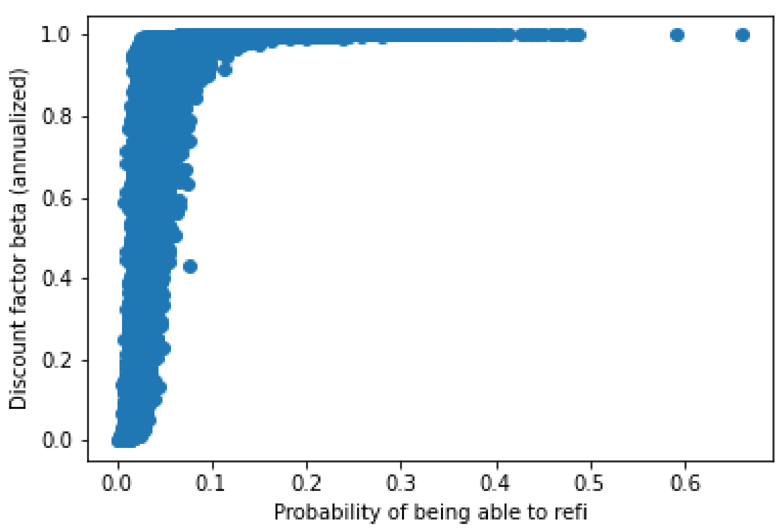
Figure 1.14 presents my estimates for borrower discount factors and their correlation with their time-varying ability to refinance. Figure 1.14a plots the distribution of discount factors, which is above 0.9 for most borrowers, but there is a small group of borrowers with discount factors closer to 0.0. The discount factors are identified from borrower choices of upfront closing costs, and the existence of many borrowers with low refinancing/moving probabilities but nevertheless get higher interest rate, lower closing cost mortgages is rationalized in the model via borrower myopia. Figure 1.14b shows a strong correlation between the likelihood of being able to refinance and the discount factor. It is a scatterplot drawn from the multivariate Logit-Normal distribution of Equation (1.22). It shows that many borrowers with a probability of being able to refinance in a particular month of less than 5% also have a discount factor significantly lower than 0.9. On the other hand, borrowers with a probability of being able to refinance in a particular month of greater than or equal to 5% tend to have a discount factor above 0.95.

Figure 1.14: Discount factor and its correlation with refinancing ability

(a) Discount factor

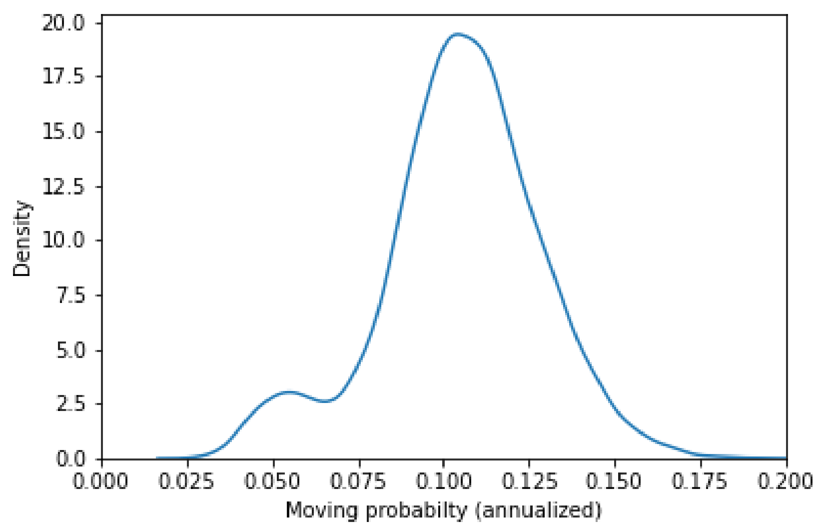


(b) Scatter plot of the probability of being able to refi and discount factor



Finally, Figure 1.15 presents my estimates of the distribution of moving probabilities by borrower. Ex ante expectations of probabilities are identified from the joint interaction of refinancing hazards and the interest rate incentive to refinance. As Figure 1.15 shows, annualized moving probabilities are centered around 11% per year, with some groups of borrowers having a lower moving probability. Appendix Figure A.6 plots these distributions by the racial group of the borrower.

Figure 1.15: Moving probability



1.6.4 Implications for transfers and welfare

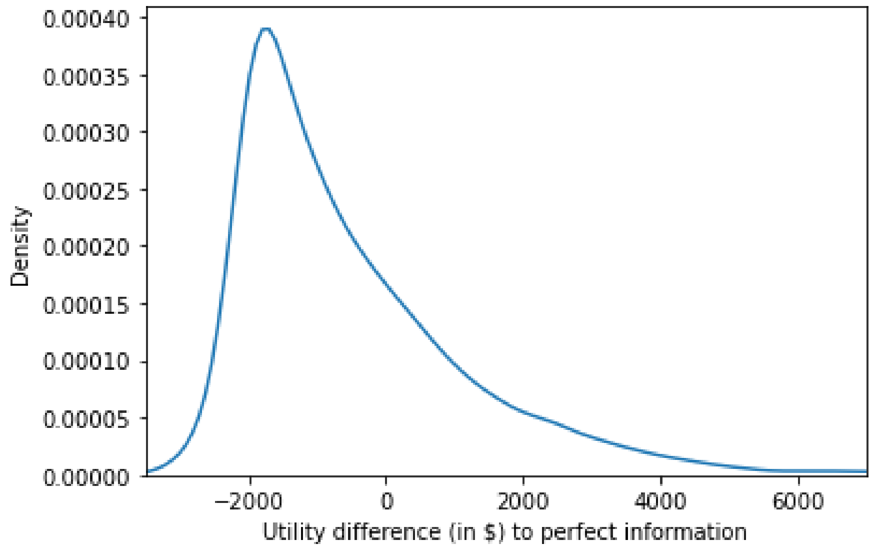
In this section I use my empirical estimates to examine the deviation of borrower behavior from the perfect information benchmark. Doing so allows me to reveal the transfers and efficiency consequences of heterogeneity in borrower refinancing behavior when interacted with the financial contract design of adding closing costs to the rate of the mortgage.

Figure 1.16 plots the differences in utility in the actual world versus the no perfect information, no cross-subsidization benchmark. I find an average welfare loss of $445 per mortgage, most of it borne by borrowers with a probability of being able to refinance of less than 5%. Given that there are around 8 million new mortgages being originated per year, the welfare loss from the closing cost channel of cross subsidization is around $3.6 billion per year. In addition, the average utility difference to the perfect information benchmark, in absolute dollar value terms, is $1339/borrower, suggesting an average difference in utility of 1% of the loan amount from slow to refinance borrowers to quick to refinance borrowers.[35] By comparison, the difference in utility from comparing the benchmark quick to refinance borrower as calibrated in Section 1.5.3 to a non-refinancing borrower that has $p_i^a = 0$ but are otherwise similar for a mortgage with zero upfront closing costs is 3.5% of the loan amount.

[35] This difference in utility is approximated by doubling the average utility difference to the perfect information benchmark, or $2678/borrower, and dividing by the average loan size of $252,000.

Figure 1.16: Differences in utility in the actual world versus the perfect information benchmark

(a) Raw Density



(b) By borrower refinancing ability

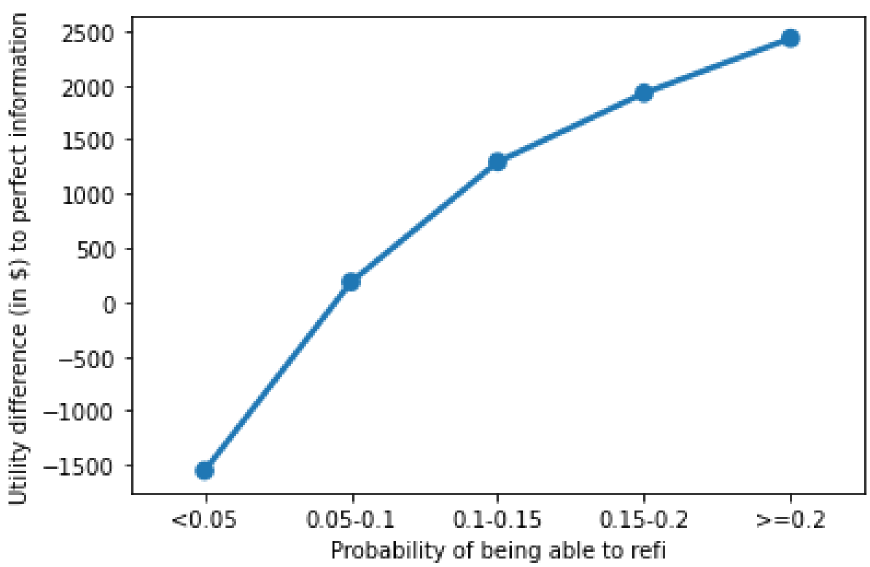
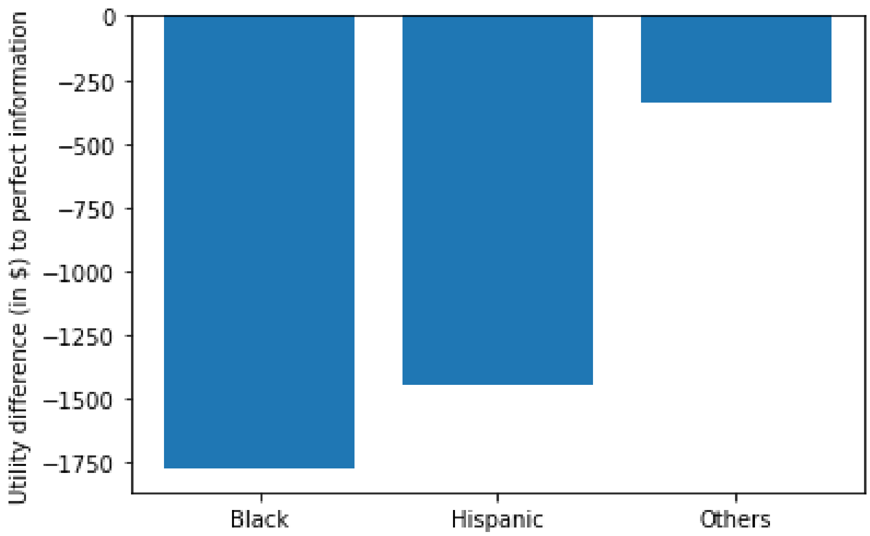


Figure 1.17 plots the welfare effects of the cross-subsidization by racial group. The welfare effects

are -\$1776 per households for Black borrowers, -\$1448 per households for Hispanic borrowers, and -\$366/borrower for other households. The welfare impact is negative for all racial groups in part due to the deadweight loss generated by the cross-subsidization of mortgage closing costs, but it is particularly strong for minorities.

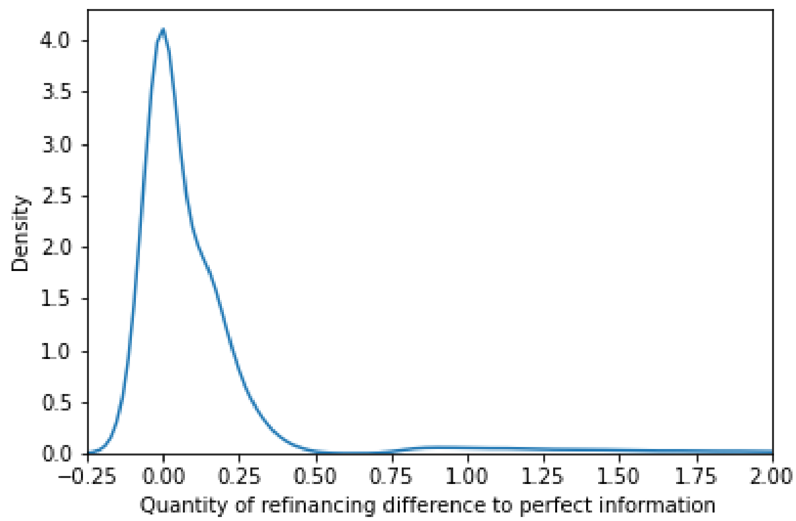
Figure 1.17: Welfare effects of cross-subsidization by racial group



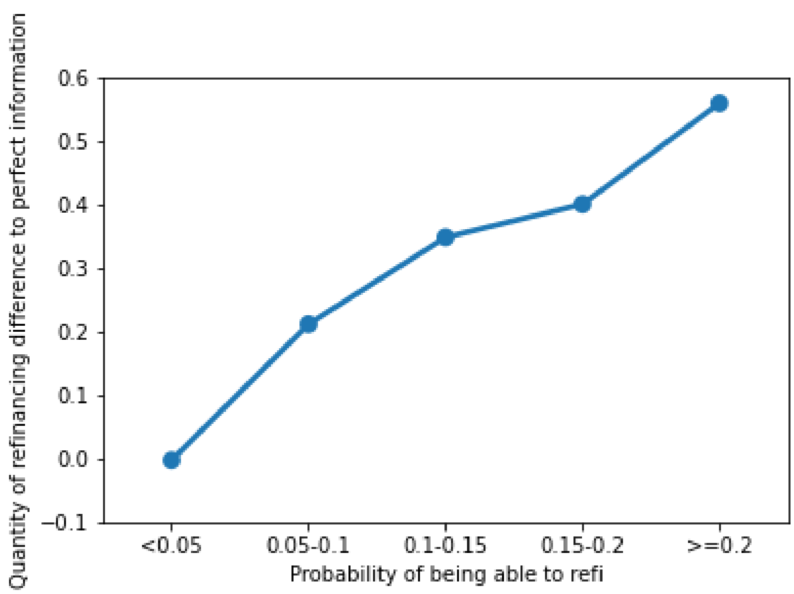
To get at the excessive refinancing incentives generated by the cross-subsidization of mortgage closing costs, Figure 1.18 plots the differences in the expected number of refinances per new origination in the actual world versus the perfect information, no cross-subsidization benchmark. I find an average increase of 0.13 refinances per new purchase origination. My model implies an average number of refinances per new purchase origination of 0.47. Therefore, it implies that 27.5% of the total US mortgage refinancing volume may be considered excessive relative to the perfect information benchmark.

Figure 1.18: Differences in the expected number of refinances in the actual world versus the perfect information benchmark

(a) Raw Density



(b) By borrower refinancing ability



1.7 Counterfactuals

I conduct two counterfactual analyses. First, I consider an alternative mortgage contract design where closing costs have to be added to the balance of the loan. The advantage of this design is that it eliminates the cross-subsidization of mortgage closing costs: all borrowers have to pay for their own price of mortgage origination. Second, I consider the case of automatically refinancing mortgages, which is a mortgage whose interest rate resets downwards automatically to a lower rate when the market rates falls by more than 1%. This contract has been discussed in Campbell (2006b). In both of these cases, I compute the updated borrower and lender value functions, and I re-estimate the equilibrium using the same zero profit condition on the supply side. To avoid complications with multiple equilibria, I restrict myself to counterfactuals where upfront closing cost choices are fixed.

1.7.1 Adding closing cost to balance

First, I consider the utility changes of borrowers when they add their closing cost to the balance of the loan. That is, their new mortgage balance becomes $M' = M(1 + c^l(M))$, and their mortgage payment becomes:

$$P_{it}(M') = M' \frac{c_{it}/12(1 + c_{it}/12)^n}{(1 + c_{it}/12)^n - 1}. \tag{1.28}$$

In periods where borrowers are able to refinance, their utility can still be written as the maximum of what can be obtained by refinancing and not refinancing, except that refinancing increases the balance of the loan from M to M'. Hence, M becomes an endogenous state variable that we add to the model which affects the size of the mortgage payment $P_{it}(M')$. The expected utility in periods where borrowers are able to refinance, $\tilde{U}^a_{i,t}$, is then:

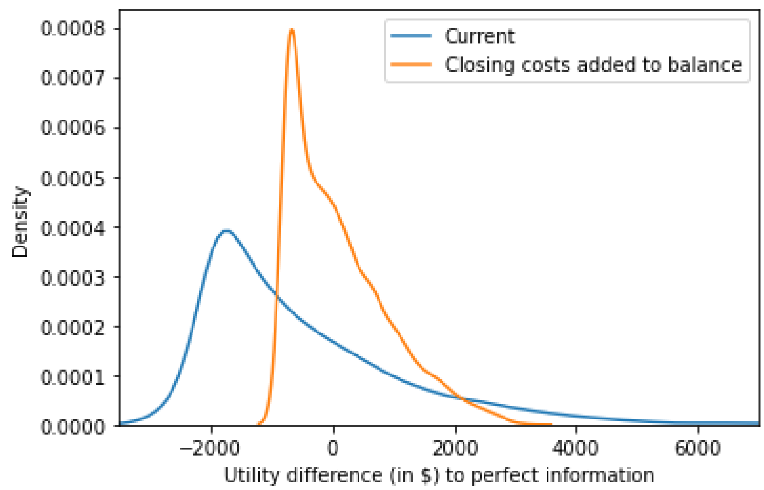
$$\mathbb{E}_t \tilde{U}^a_{i,t} = \max \begin{cases} \max_{\Delta S_{it}} \frac{(\exp(L_{it}) - P_{it}(M) - (r_{1,t-1} - \pi_t) S_{i,t-1} - \Delta S_{it})^{1-\gamma_i}}{1-\gamma_i} + \beta \mathbb{E}_t \tilde{V}_{i,t+1}(c_{it}, S_{it}, M), \text{ if no move/refi} \\ \max_{\Delta S_{it}} \frac{(\exp(L_{it}) - P_{it}(M') - \tilde{\kappa}_{it} - (r_{1,t-1} - \pi_t) S_{i,t-1} - \Delta S_{it})^{1-\gamma_i}}{1-\gamma_i} + \beta \mathbb{E}_t \tilde{V}_{i,t+1}(c_{it}, S_{it}, M'), \text{ if refi} \end{cases} . \tag{1.29}$$

I simulate borrower utility and prepayment behavior under this counterfactual with borrower utility when they are able to refinance being described by Equation (1.29) instead of Equation (1.13). I then obtain the implied aggregate borrower behavior and lender values based on my estimated distribution of borrower types in Table A.7, conditional on the paths of interest rates as estimated in Section A.5.1. Finally, I decrease the initial mortgage interest rate for all borrowers, holding fixed their prepayment behavior, until the zero profit condition in Equation (1.18) is satisfied on average, which is an equilibrium effect of this contract design that increases in borrower utility.

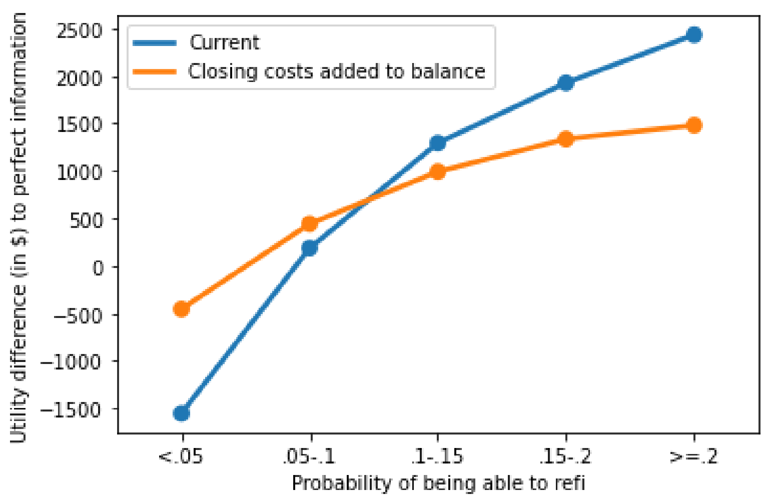
Results are shown in Figure 1.19a, with a significantly narrower range of utility differences relative to the perfect information benchmark. When closing costs are added to the balance of the mortgage, there are still gains from actively refinancing relative to not refinancing, albeit less than in the current world. This reduces the cross-subsidization between borrower types. In particular, the average utility difference to the perfect information benchmark, in absolute dollar value terms, falls by around half from $1339/borrower in the current world to $698/borrower in this counterfactual world. The same reduction in cross-subsidization can be inferred from Figure 1.19b, which plots the mean utility difference to the perfect information case, in dollar terms, by buckets of borrower refinancing ability.

Figure 1.19: Counterfactual utility from adding cost to balance

(a) Distribution



(b) By borrower refinancing ability



In terms of total welfare, I find that on average consumer welfare relative to the perfect information benchmark rises from -$446/borrower to $110/borrower. Not only is the negative welfare impact of excessive refinancing eliminated in this contract design, but there is also a welfare gain due to the relaxation of financial constraints as closing costs can be added to the balance. In the current world, actively refinancing borrowers can only pre-commit to not undertaking costly refinancing activity by paying more in upfront closing costs, which is itself costly due to financial constraints. Otherwise, they would have to take a higher initial interest rate and refinance more which carries administrative resource costs. The addition of mortgage closing costs to the balance both eliminates the cross-subsidization of mortgage closing costs and resolves this commitment problem. As a result, it is able to simulatenously reduce transfers by borrowers with different refinancing tendencies and also increase total welfare.

Appendix Figure A.7 plots the counterfactual change in utility by racial group under the alternative contract design of adding all closing costs to the balance of the loan. All racial groups gain from this counterfactual, with Black borrowers gaining on average $1566, Hispanic borrowers gaining $1325, and other borrowers gaining $472. The average welfare gain under this counterfactual is $556.

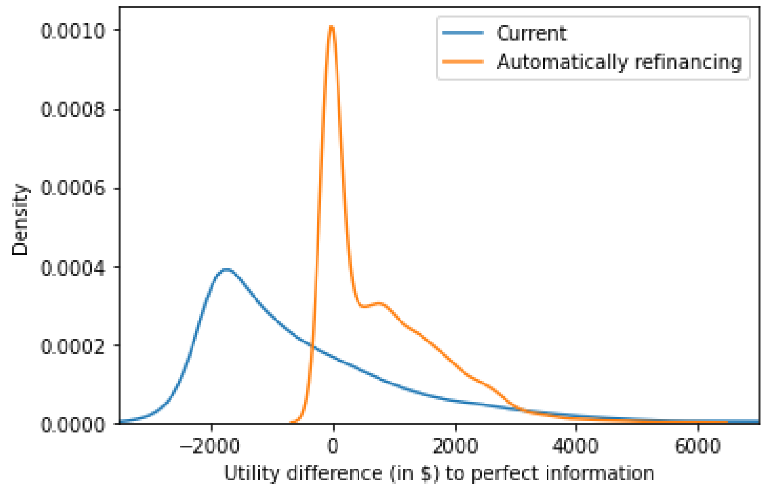
1.7.2 Making mortgages automatically refinancing

Second, I consider a counterfactual where mortgages are automatically refinancing and are originated with zero upfront closing costs. In this case, I keep the same demand as in Section 2.3 but automatically change the mortgage interest rate from c to r whenever $c - r > 1$ conditional on the paths of interest rates as estimated in Section A.5.1. Furthermore, I eliminate the possibility of refinancing as that is no longer relevant. Finally, I increase the initial mortgage premia over the risk-free rate for all borrowers until the zero profit condition in Equation (1.18) is satisfied in the counterfactual, which is an equilibrium effect of this contract design that decreases borrower utility.

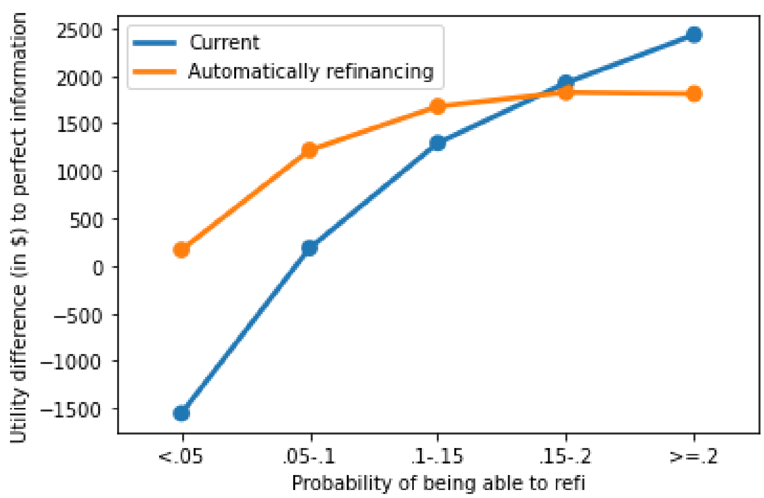
Results are shown in Figure 1.20. I terms of distribution, automatically refinancing mortgages also feature lower average utility difference to the perfect information benchmark compared to the current world. In particular, I find that this statistic falls from $1339/mortgage to $773/mortgage. Furthermore, the automatically refinancing mortgages counterfactual feature a greater welfare improvement relative to the current world compared to adding closing costs to the balance of the loan, at $1216/mortgage. This significant improvement is due to the resource cost savings of refinancing, and is concentrated among the actively refinancing borrowers as shown in Figure 1.20b. Appendix Figure A.8 plots the counterfactual change in utility by racial group under the alternative contract design of automatically refinancing mortgages, showing that all racial groups would on average increase their utility in this counterfactual.

Figure 1.20: Counterfactual utility from automatically refinancing

(a) Distribution



(b) By borrower refinancing ability



Conceptually, there are two main channels through which automatically refinancing mortgages can increase total welfare. First, they can eliminate the excessive refinancing incentives from the cross-subsidization of mortgage closing costs. Second, they also generate resource savings by eliminating the administrative and hassle costs of refinancing. To the extent that automatically refinancing mortgages present real resource savings to the economy and enable a more efficient pass-through of monetary policy not modelled here, it may be an attractive contract design for policymakers to consider.

1.8 Conclusion

The broad lesson of my paper is that in markets for consumer financial products, seemingly small contractual details can have significant equity and efficiency implications. I illustrate this lesson quantitatively in the US mortgage market where borrowers typically choose to finance their closing costs through the rate. I show that this contractual feature exacerbates transfers between borrower refinancing types while also generating deadweight losses through incentivizing excessive origination. In terms of policy, my results suggest that two alternative mortgage contract designs—(1) adding closing costs to the balance of the loan and (2) having automatically refinancing mortgages—can simultaneously reduce inequality in the market and improve total consumer welfare.

2

Testing for Discrimination in Menus

[1]

---

[1] co-authored with Paul Willen

2.1 Introduction

Whether mortgage lenders discriminate against minority borrowers is an important question both in terms of academic research and in regard to its policy relevance.[2] However, the task of assessing whether lenders discriminate by offering minority borrowers worse prices is complicated by the fact there are two dimensions to mortgage pricing in the US: the interest rate and the upfront fees charged by the lender. In particular, borrowers can choose to pay higher upfront fees (in the industry referred to as paying discount points) in return for lower interest rates. Or, conversely, they can get the lender to pay some of their closing costs in exchange for a higher interest rate. We show that the availability of this choice between a higher upfront fee and a higher interest rate creates a "menu problem" that makes the detection of lender discrimination nontrivial, with the methods implemented in the literature susceptible to false and contradictory results. The menu problem is also broadly applicable to empirical analyses of differences in opportunity under multi-dimensional choice. We propose a novel identification argument and a new procedure for inference to

[2] Since the financial crisis, the Department of Justice has reached settlements of well over $500 million with lenders that overcharged Black and Hispanic borrowers in violation of the Fair Lending Act, as explained in Bhutta and Hizmo (2020). These settlements include $335 million with Bank of America (on behalf of Countrywide), $175 million with Wells Fargo, and $55 million with JP Morgan Chase. On June 12, 2019, Sen. Elizabeth Warren wrote on Twitter, "For generations, lenders have given African American & Latino families fewer loans at worse terms than similar white borrowers. Tech alone won't fix the problem. A new analysis found that discrimination is hardwired into lending algorithms. I want answers." https://twitter.com/senwarren/status/1138909674781237253.

deal with this menu problem, and apply it to a new dataset on mortgage pricing by race.

Many studies find that minority consumers pay higher interest rates compared with observationally similar white consumers in the mortgage market.[3] While this can be interpreted as evidence that lenders systematically discriminated against minority borrowers by offering them worse pricing on their mortgages, the existence of a rate-point trade-off leads to an alternative explanation: minority consumers were offered the same menu of choices but ended up with higher rate/lower point combinations because, for example, they could not afford to pay points.[4] Explaining rate gaps with offsetting point gaps does not rule out structural disparities between racial groups, but it has policy implications that are very different from those of one in which the lenders themselves are systematically offering minority consumers inferior menus of rates and discount point options. Given data on the borrowers' chosen mortgage rates and points (but not the menus borrowers faced,

[3]See, for example, Black and Schweitzer (1985), Boehm, Thistle, and Schlottmann (2006), Bocian, Ernst, and Li (2008), Kau, Keenan, and Munneke (2012), Ghent, Hernández-Murillo, and Owyang (2014), Cheng, Lin, and Liu (2015), Bartlett et al. (2021). Relatedly, Munnell et al. (1996) and Tootell (1996) find that minority borrowers are more likely to be rejected for mortgages; Black, Boehm, and DeGennaro (2003) find that minority borrowers pay higher yield spreads when refinancing their mortgage; and Ambrose, Conklin, and Lopez (2020) find that minority borrowers pay more in broker fees particularly when faced with a white broker.

[4]The logic the different rates could reflect different options on the same menu was well understood in the industry. Skanderson, Darius, and Butler (2014), for example, write that, "[v]arious interest rates are available to any given borrower, at any given point in time, depending upon his or her needs and preferences [regarding points and closing costs]. The fact that similar borrowers selected different rates on this basis does not mean that they were treated differently in a sense that should be relevant for fair lending compliance. Nevertheless, such differences in borrower choices can create the appearance of pricing disparities if they happen to be correlated with a prohibited basis and are not captured in the data and accounted for in a fair lending analysis."

which are not typically observable), our objective is to examine whether lenders discriminated against Black borrowers in the sense of offering them a distribution of menus that was worse than the one offered to observationally similar white consumers, a practice we call *discrimination in menus*.

Our first contribution is to point out that there exists a nontrivial econometric problem involved with assessing differences in the distribution of menus offered to minority and non-Hispanic white borrowers. Researchers might think we can address heterogeneity in preferences over rates and discount points by racial group by simply controlling for the discount points in their regressions. However, we show that the approach of controlling for rates and discount points can lead to contradictory results. For example, it is sometimes observed that controlling for rate, minority borrowers pay the same closing costs as white borrowers (so it looks like there is no mortgage pricing differential by race), but controlling for closing costs, minority borrowers pay a higher rate (so it looks like there is a mortgage pricing differential by race). This is related to the "reverse regression" problem of Goldberger (1984) which appears naturally in the menu setting. The menu problem also goes beyond this "reverse regression" problem. Even when forward and reverse regressions are consistent, we show that false positives in the sense of detecting discrimination when none exists and false negatives in the sense of failing to detect discrimination when it does exist can still appear. Finally, we show that even a seemingly foolproof comparisons of means—

that is, checking if minority consumers on average pay both a higher interest rate and more in lender fees—can still lead to false positives and false negatives if interpreted as evidence of discrimination in menus.[5] These issues emerge when minority and non-Hispanic white borrowers are likely to make different decisions (that is, have different unobserved preferences) over menu items, which is particularly relevant because the possibility of heterogeneity in preferences across racial groups is usually the reason researchers would seek to control for the choices of discount points in the first place.

The menu problem is important for the mortgage pricing discrimination literature. There are two main methods by which a large literature assesses discrimination in mortgage pricing given the rate and discount point trade-off. First, Courchane and Nickerson (1997), Bhutta and Hizmo (2020), and Bartlett et al. (2021) look at whether Black borrowers paid more in points conditional on rate.[6] Second, Woodward (2008) and Bartlett et al. (2019a) compare the interest rates of minority and white borrowers after adjusting for points using a known range of rate-point trade-offs, and find that minorities paid higher

[5]It follows that adjusting by a known range of rate and lender fee trade-offs and then comparing means can be similarly problematic.

[6]In samples of FHA mortgages, Courchane and Nickerson (1997) find a differential in points paid by race conditional on rate, while Bhutta and Hizmo (2020) do not. One possible reason for this difference is that Bhutta and Hizmo (2020) improved on the earlier literature by using a much larger sample and constructing a more uniform sample of loans, which are improvements we largely adopt. Bhutta and Hizmo (2020)'s study also includes an analysis of the mean levels of rates and points paid conditional on covariates, which is similar in spirit to the second method we discuss. Finally, beyond comparing prices, they find that lenders received more revenue from loans that were made to minorities once points and secondary marketing revenue are added together.

point-adjusted-rates for mortgages. As we discussed, both of the existing methods used in the literature, (1) controlling for rates and (2) comparing mean levels of rates and discount points after adjusting by a known range of slopes, can lead to false positives and false negatives and can contradict one another as an assessment of mortgage pricing discrimination. Indeed, we show that in our FHA sample the choice of either (1) or (2) does lead to apparently contradictory results on whether mortgage pricing discrimination exists. Our robust solution to the menu problem would therefore allow researchers and regulators to assess discrimination in mortgage markets in a more internally consistent and theoretically sound way.

The menu problem also extends well beyond the mortgage setting. Generally speaking, the problem is relevant whenever a researcher wishes to assess disparities in opportunity given data on choices while allowing for heterogeneous preferences across groups. For example, when workers make decisions that trade off wages and hours worked, researchers may wish to assess the extent to which the gender gap in pay may be explained by the choice of hours, as in the model of Goldin (2014). The menu problem implies that popular measures of gender inequality, such as the gender pay gap conditional on hours or even the gender pay gap after adjusting for all relevant average compensating differentials, are not necessarily informative about whether the data on wages can be explained by heterogeneous preferences over hours worked. Our robust metrics can also be useful for this type of problems depending on the institutional details the researcher has access to. Therefore, the problems we point out and the solution we propose may be of broad interest.[7]

As a solution to the menu problem, we propose (1) a new metric for detecting whether there exists a difference in the distribution of menus offered to two groups and (2) a new lower bound measure for assessing differences in menus (DIM) for the extent to which one group of consumers would like to switch to another group's menus. Both metrics are based on pairwise dominance relationships in the data (that is, a mortgage with a lower rate and paying fewer points dominates a mortgage with a higher rate and more points) that can be supplemented by industry knowledge. Based on these pairwise relationships, we ask the question of whether the data *can* be rationalized by a model of equality in menus but heterogeneity in preferences, and if not, we compute an average difference in menus perceived by one group of consumers when switching to another group's menus. Unlike the existing methodology used in the literature, our metrics are robust to any form of unobserved differences in preferences across borrower groups.

The sample counterparts to both of our metrics can be computed as solutions to optimal transport problems, which are computationally well understood and can be efficiently computed through linear programming. As a technical contribution, we also derive a new

[7]The labor literature has also analyzed differences in worker productivity and wage inequality within the context of structurally specified models of labor supply (e.g. Hwang, Reed, and Hubbard (1992) and Bell (2019)). Our illustrations highlight the role of seemingly innocuous assumptions, such as the restriction of productivity differences to a single component of unobserved heterogeneity, that is typically used in these models.

approach to uniformly valid inference for the value of optimal transport problems, which we implement for our metrics to distinguish between statistical noise and actual differences in menus. Conventional approaches to inference, such as bootstrapping, fail for optimal transport problems, because the objective function can be non-differentiable (Fang and Santos, 2018). We prove that optimal transport problems are directionally differentiable in the sense of Shapiro (1991) and Fang and Santos (2018). We then apply the asymptotic results of Fang and Santos (2018), which we combine with a Bonferroni correction following Romano, Shaikh, and Wolf (2014) and McCloskey (2017) to address sampling error in the directional derivatives. We show that this approach leads to asymptotically uniformly valid size control for hypothesis testing in the value of optimal transport problems, and test it in a Monte Carlo simulation. Our new approach to inference in optimal transport may be useful for other researchers who wish to conduct inference on the value of optimal transport problems, many of which are described in Galichon (2016).

Empirically, we use our metrics to assess mortgage pricing discrimination in the 2018–2019 Home Mortgage Disclosure Act (HMDA) data matched to Optimal Blue rate locks. We show that we can detect a difference in menus offered by the same lender in the same county and within narrow covariate groups for conforming mortgages for both Black and Hispanic borrowers relative to non-Hispanic white borrowers. Furthermore, we show that on average Black borrowers getting conforming mortgages would be willing to increase their interest rate by at least 1.7 basis points in order to switch to the menus of non-Hispanic white borrowers. Similarly, Hispanic borrowers are on average willing to pay 1.5 basis points more in interest rate in order to switch menus with non-Hispanic white borrowers. Our finding that racial differences in lender pricing remains relevant for conforming mortgages is consistent with Bartlett et al. (2021), although the amount of interest rate discrimination we detect is smaller in magnitude. On the other hand, we do not detect interest rate discrimination in FHA mortgages, which is consistent with Bhutta and Hizmo (2020). Within conforming mortgages, the differences in menus we detect are particularly concentrated among borrowers with lower loan-to-value (LTV) ratios and are not explained by multi-product discounts or loan originator compensation.[8] The fact that the mortgage pricing discrimination we detect is concentrated among the more creditworthy conforming mortgage borrowers is consistent with the less risky non-Hispanic white borrowers being more likely to be offered discounts during the search and negotiation process.[9]

[8]Relatedly, Ambrose, Conklin, and Lopez (2020) find in a pre-2008 sample period that the more creditworthy minority borrowers pay higher broker fees than observationally similar non-Hispanic white borrowers, even though their default risks are similar. Regulation Z of 2011 now forbids mortgage brokers from varying the fees they charge across borrowers except as a function of the loan amount, shutting down this particular channel of disparity, though similar forces may be play in the disparities we uncover.

[9]Negotiation in mortgage markets is common. Studies that look at the search and negotiation process in mortgage markets include Allen, Clark, and Houde (2014), Allen, Clark, and Houde (2019), and Bhutta, Fuster, and Hizmo (2019). A reason why minority borrowers may behave as if they have higher search costs is in Agarwal et al. (2020c), where borrowers internalize a higher probability of rejection. Price discrimination by such effective search costs would violate the US fair-lending law, to the extent it results in disparate impact by race within lender, according to Bartlett et al. (2021).

In our empirical application, we focus on the important and basic question of whether we can detect differences in the rate and discount point menus faced by observably similar non-Hispanic white and minority mortgage borrowers. The answer to this question is important because one needs to ascertain whether there *exists* disparate treatment before considering its causes. Our use of the word "discrimination" to refer to this disparate treatment conditional on observables is in line with the recent literature and is justified by the unusual institutional details and the regulatory framework developed around this market. First, as explained in Bartlett et al. (2021), the US fair-lending law imposes a requirement of no pricing differentials by race conditional on observables in our setting, so our results are naturally interesting from a regulatory perspective. Second, for the types of mortgages we focus on, lenders are insured from the risk of default by either the GSEs or the FHA, who compensate investors for any losses of principal but allows them to benefit from the more favorable prepayment risk of minorities. Therefore, in contrast to some stereotypes, mortgages from Black and Hispanic borrowers are likely significantly more valuable than those of observably similar white borrowers due to their lower prepayment risk, as shown in the simulations of Kau, Fang, and Munneke (2019a) and in analyses of the actual mortgage-backed securities prices in Gerardi, Willen, and Zhang (2020). As a result, even if lenders did use unobservables that are correlated with race to price for expected loan performance in a possibly illegal manner, it would be unlikely to justify the unfavorable pricing to minorities we find.

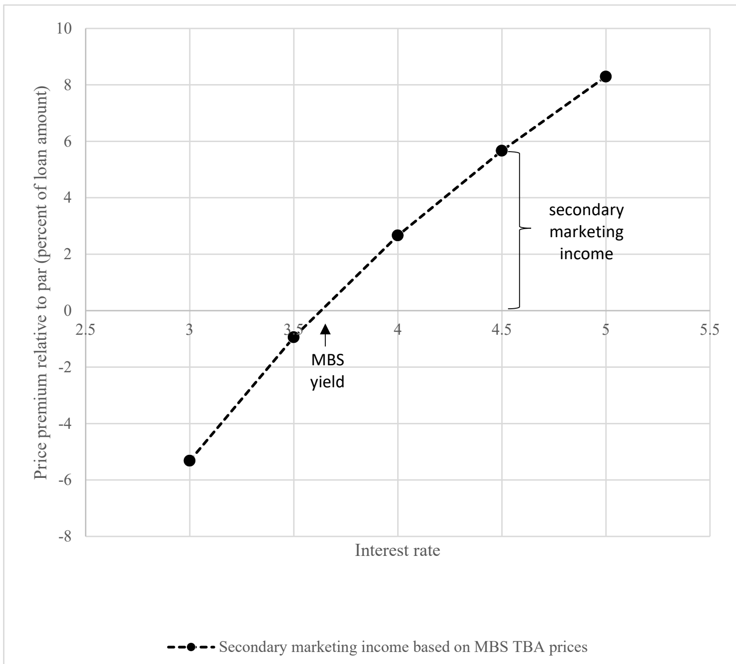
The rest of this paper is structured as follows. Section 2.2 explains the motivation of our paper by exploring why heuristic approaches to analyzing discrimination in menus may be misleading. It also provides intuition for our approach. Section 2.3 formally defines our metrics for assessing discrimination in menus. Section 2.4 describes a methodology for conducting inference on our metrics. Section 2.5 shows our data and empirical results. Section 2.6 concludes.

2.2 The menu problem

2.2.1 Problem description

In this section we introduce the menu problem and discuss why intuitively appealing approaches for assessing discrimination in menus may be misleading. By way of background, there are two dimensions of pricing for mortgages in the United States: an upfront fee/discount points and the interest rate, where each point is customarily worth 1 percent of the loan amount. Consumers can have the option of picking a particular rate and discount point combination that best suits their preferences and financial constraints. We plot those choices from an example rate sheet in Figure 2.1. We also present a screenshot illustrating this trade-off from an online mortgage price comparison service in Appendix Figure A.12. In particular, borrowers can pay discount points to reduce their interest rate or receive money from the lender to help cover their closing costs by getting lender credit (paying negative points). The sense in which we think about lender discrimination in menus, then, is for minority borrowers to receive a worse rate-point schedule than white borrowers.

Figure 2.1: An example set of menu items from a lender rate sheet.



In practice, the researcher often observes the distribution of borrower choices x but not the underlying menus $\mathbf{m}$ where $x \in \mathbf{m}$ is chosen from. The menu problem then emerges as the problem of inference on the extent to which a matched group of borrowers who are by construction observationally similar in terms of covariates faced the same distributions of menus. More specifically, testing for equality of menus can be written as testing the null hypothesis $\mathcal{H}_0$ that the distributions of menus being offered to both groups are equal. That is, suppose $\mathbf{m}_1 \sim \mathbf{M}_1$ for borrowers in group 1 and $\mathbf{m}_2 \sim \mathbf{M}_2$ for borrowers in group 2, the menu problem is the hypothesis testing problem where:

$$\mathcal{H}_0 : \mathbf{M}_1 = \mathbf{M}_2, \tag{2.1}$$

$$\mathcal{H}_1 : \neg\mathcal{H}_0. \tag{2.2}$$

An important stated objective of the applied literature on mortgage pricing discrimination, and indeed the literature on differences in opportunity more broadly, is assessing equality in menus as specified in Equation (2.1). Nevertheless, existing methodology used in this literature tends to fall short of being able to credibly assess inequality in menus. We point out a series of difficulties associated with using existing methodology to test $\mathcal{H}_0 : \mathbf{M}_1 = \mathbf{M}_2$, which we call the "menu problem."

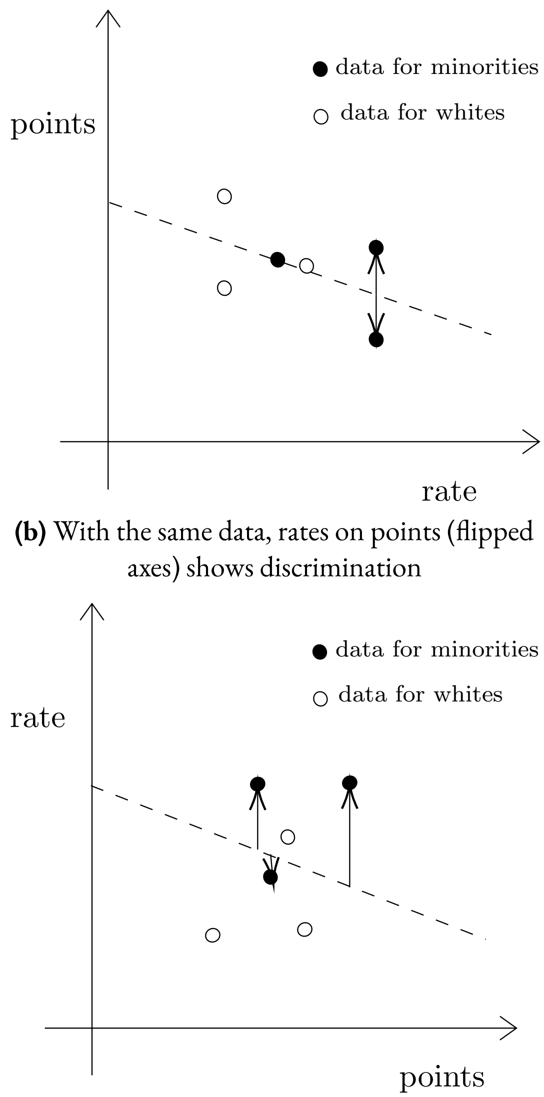
One natural approach to assessing whether lenders offered minority and non-Hispanic white borrowers different menus is to control for one dimension of the menu, that is, conditional on the distribution of covariates, to estimate whether minority borrowers who received the same interest rate as white borrowers paid more in discount points. This approach was used in Courchane and Nickerson (1997), Bhutta and Hizmo (2020), and Bartlett et al. (2021). A problem with this approach, however, is that it can lead to contradictory estimates depending on whether the researcher chooses to control for rates or points.[10] The situation in Figure 2.2 shows that it is possible for a regression of points on rate to show no discrimination against minorities while a regression of rate on points shows discrimination with the same example data. In this figure, we represent example data from minority and white borrowers using black and white dots, respectively; regression line by the dashed line; and the difference to the regression line by the solid arrows. Figure 2.2a shows that a regression of points on rate and borrower race would show a zero coefficient for minorities, with the two arrows balancing each other out. On the other hand, Figure 2.2b shows that, using the same data, a regression of rate on points would instead give a positive coefficient for minorities. This sort of contradiction is not particular to the linear regression case, and as we show in Appendix Figure A.13, it can appear with general conditional expectations.

[10]Bhutta and Hizmo (2020) also runs lender-specific regressions, and Bartlett et al. (2021) also controls for lender-specific slopes in a robustness check. However, the problems identified here exists as long as there are within-lender hetergeneity in menus, which we find is prevalent in Section 2.2.2.

Figure 2.2: How the choice of which menu dimension to control for can lead to contradictory findings of discrimination

(a) Points on rates shows no discrimination

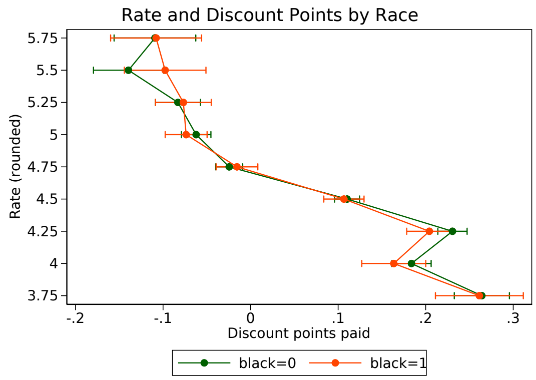
(b) With the same data, rates on points (flipped axes) shows discrimination



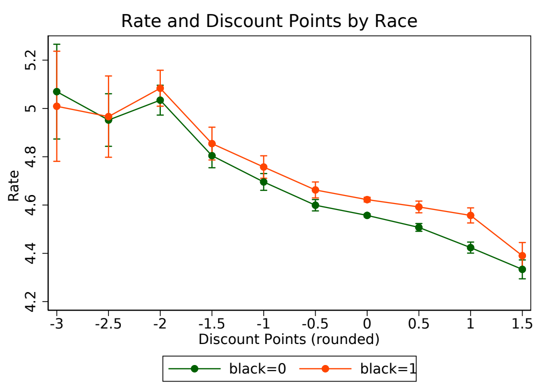
This possibility is not merely a theoretical curiosity. Figure 2.3a plots the predicted values of discount points paid in a regression with race dummies interacted with levels of interest rate spread rounded to the nearest quarter of a percent for the FHA mortgages in our sample. It shows that, conditional conditional on the rate paid, Black and white borrowers pay similar numbers of points. On the other hand, using the same data, Figure 2.3b shows the predicted values of interest rate paid in a regression with race dummies interacted with levels of discount points rounded to the nearest half a point. It suggests that, conditional conditional on the number of points paid, Black borrower pays higher rates than white borrowers. Thus, the regression approach can give strikingly clear, though ultimately misleading, results.

Figure 2.3: An empirical example for the choice of which menu dimension to control for can lead to contradictory findings of discrimination

(a) Points on rates shows no discrimination



(b) With the same data, rates on points (flipped axes) shows discrimination

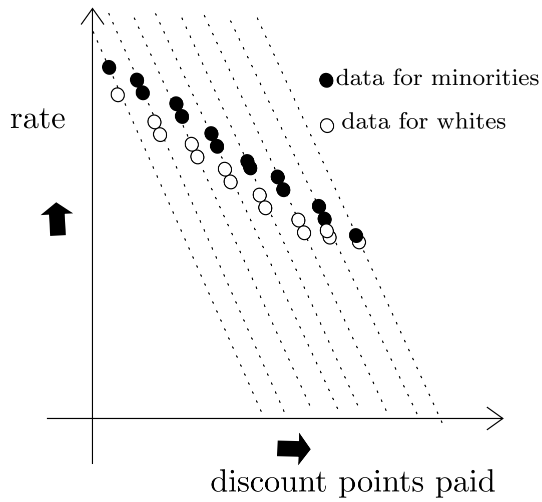


This situation in Figures 2.2 and 2.3 can be viewed as a version of the reverse regression problem of Goldberger (1984). This problem emerge naturally in the menus setting, since when borrowers are choosing from a menu of options it is in theory not clear which direction the regression should be run. The problem with testing equality in menus goes beyond the reverse regression problem, as there can be false positives and false negatives even when the forward and reverse regressions are consistent (and even when a simple comparison of means is consistent) which we will discuss in the rest of this section.

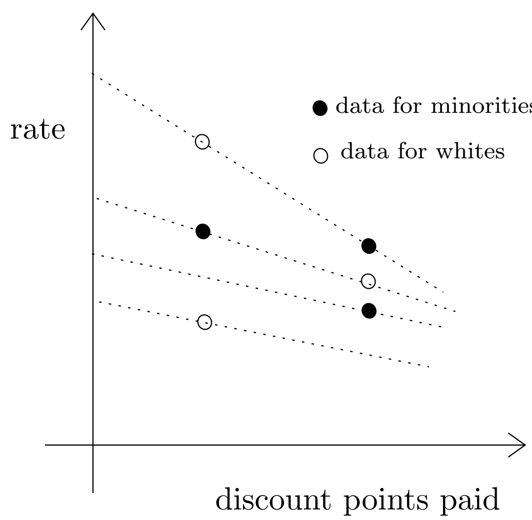
Figure 2.4 shows how even when forward and reverse regressions consistently detect discrimination or no discrimination, the heuristic of controlling for one dimension of the menu can lead to false positives and false negatives if interpreted as a test for equality in menus as in Equation (2.1). In this figure, we represent example data from minority and white borrowers using black and white dots, respectively, and menus by dotted lines where each dotted line is one potential menu from which a borrower may draw. The left panel of Figure 2.4a shows a false positive situation in which minority borrowers paid more in rate, controlling for points, and more points, controlling for rate, even though lenders offered both minority and white consumers the same distribution of menus. The only difference in borrower behavior by group is that minority consumers chose to pay fewer points on every menu. That is, even though $\mathbf{M}_1 = \mathbf{M}_2$ in reality, as represented by a common set of dotted lines facing both groups, it appears as if minorities are worse off controlling for either dimension of the menu. In the right panel of Figure 2.4b, we illustrate a false negative situation in which minority consumers paid the same rate conditional on points but faced a worse distribution of menus, since the bottom menu (the most advantageous menu) was offered only to white borrowers while the second-to-bottom menu was offered only to minority borrowers. Furthermore, Figure 2.4b can be constructed in a way such that the variance of the rate and discount points paid are equal, such that forward and reverse regressions both give the same false negative but there does exist discrimination in menus. That is, in the situation of Figure 2.4b, $\mathbf{M}_1 \neq \mathbf{M}_2$ in reality, but controlling for either dimension of the menu the researcher would find no difference between the two groups.

Figure 2.4: False positives and false negatives from controlling for one direction of the menu

(a) False positive, minority borrowers paid more in rate (points), controlling for points (rate), but the menus were the same



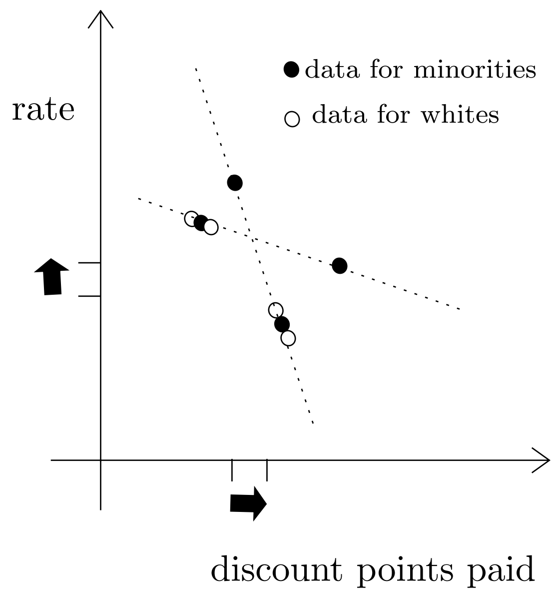
(b) False negative, minority borrowers paid the same average rate, controlling for points, as white borrowers, but their menus were worse



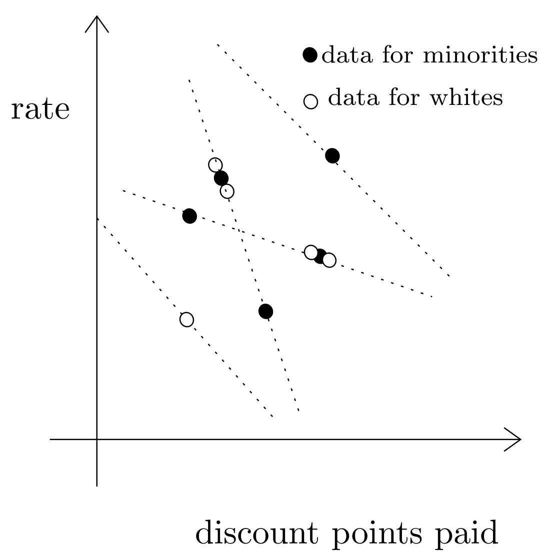
A second intuitively appealing approach for assessing discrimination in menus is to compare means, thus avoiding the problem of having to estimate menu slopes from the data. In other words, the researcher may wish to check if minority consumers paid more on average in both rates and discount points, such that they are disadvantaged in both dimensions compared with observationally similar white borrowers. A variation of this approach is to take a pre-defined range of rate-point trade-offs as the slope estimated from external sources, which is done in Woodward (2008) and Bartlett et al. (2019a). While this avoids the problem that regressions may incorrectly estimate menu slopes, it can still lead to false positives and false negatives when slopes are not constant across menus, thus breaking the assumption (1) that all menus share the same shape. This is a realistic problem, because we know from rate sheet data that an unobserved heterogeneity in slopes does exist in the mortgage setting, as the rate-point trade-offs do vary substantially across lenders and over time (Figure 2.9). Figure 2.5a illustrates how, when slopes differ across menus, a false positive in which minority consumers pay more on average in terms of both rates and points but faced the same distribution of menus as white borrowers can occur ($\mathbf{M}_1 = \mathbf{M}_2$). And Figure 2.5b illustrates a false negative possibility in which minority borrowers paid the same average rates and points as white borrowers but did face worse menus ($\mathbf{M}_1 \neq \mathbf{M}_2$).

Figure 2.5: False positives and false negatives from checking if minority borrowers paid more on average in both rates and points

(a) False positive, minority consumers paid more on average in both rates and points, but their menus are the same



(b) False negative, minority consumers pay the same average rate and points as white borrowers, but their menus were worse



The mechanism for when a comparison of means would lead to false positives in the case of Figure 2.5a is that in the example, minority borrowers respond less to differences in the slopes of the rate-point menus compared with white borrowers. This possibility is empirically relevant, because there are two directions in which constraints can drive borrower choices of rates and closing costs. First, if borrowers are cash constrained when getting the loan, they may need to get a lower closing cost mortgage (pay fewer discount points) regardless of what rate-point trade-offs lenders offer. Second, if borrowers are debt-to-income (DTI) constrained, they may need to pay more points to buy down the rate to increase their borrowing limit.[11] Therefore, the finding that minority borrowers on average pay more than white borrowers may simply reflect the fact that minority borrowers are more constrained in their choices, which is a form of "disadvantage" that is not necessarily due to lenders discriminating against them by offering them different menus.

While simple in hindsight, the situation of Figure 2.5a also illustrates how a seemingly innocuous assumption that differences in menus faced by different agents can be summarized by a single monotonically additive term is actually a strong assumption when there are differences in both menu intercepts and slopes in reality. Perhaps due to tractability, such an assumption is popular in models of labor productivity (e.g. Hwang, Reed, and Hubbard (1992), Bell (2019)), where agents are assumed to face structurally defined menus of trade-

[11] See, for example: https://www.thetruthaboutmortgage.com/dti-debt-to-income-ratio/.

offs indexed by a single productivity term, and estimation proceeds by devising a method to consistently estimate the distribution of that term. Our illustration implies that by ignoring the multidimensional way unobserved heterogeneity in trade-offs can enter into the model (for example, in terms of both levels and slopes), a researcher can be led to conclude that, for example, one group of workers is more productive than another group when one group is simply more flexible than another group in terms of their preferences over wage-amenities trade-offs. Therefore, whether a single index assumption is reasonable or not depends on how much of a concern the situation illustrated in our Figure 2.5a would be in the specific empirical context in which they are applied.[12]

Fundamentally, the shortcomings of existing empirical methodologies, when applied to the menu problem, can be summarized as a combination of omitted variables bias and misspecificiation bias. Roughly speaking, the assumptions underlying the heuristic approaches of assessing discrimination in menus are that (1) all menus share the same shape, with unobserved heterogeneity in menus being due to an additive error term, and (2) that this shape can be correctly estimated/inferred from other data. The omitted variables problem stems from the fact that the preferences of consumers, since they are unobserved, can lead to bias when the slopes of menus are estimated from data by controlling for them, thus breaking

[12]Another setting where a single (log-)additive productivity assumption is popular is in the production function literature (e.g. Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg, Caves, and Frazer (2015)). Therefore, while methods from the production estimation literature are useful for estimating average returns to capital and labor, they are also problematic if applied as a test of equality in menus as in our mortgage context.

implicit assumption (2). This is the main problem with simply controlling for some dimensions of the menu, as illustrated in Figures 2.2 and 2.4. While it is less recognized in the mortgage pricing discrimination literature, much work in labor productivity estimation is indeed aimed at estimating the slopes of trade-offs and addressing the omitted variables problem in (2). A more subtle problem is that correctly estimating the average slopes of trade-offs is not sufficient for assessing equality in menus. Models that assume away heterogeneity in slopes and interpret residuals to average trade-offs as differences in menus can generate a mis-specification error when heterogeneity in slopes does exist in reality, which can lead to misleading inference by breaking assumption (1). This is the problem illustrated in Figure 2.5.

2.2.2 The relevancy of the "menu problem" in the mortgage setting

As we illustrated in Figure 2.3, the possibility for contradiction in regression estimates of differences in menus is an empirically relevant problem in the mortgage discrimination setting in the sense that it is present in our data. More fundamentally, the problems illustrated in Section 2.2.1 depend on (1) the existence of unobserved heterogeneity in menus and (2) in the differences in preferences between borrower racial groups. We argue that these issues are first order in the mortgage setting.

Empirically, we find that there exists large unobserved heterogeneity in menus across borrowers within narrow covariate groups and from the same lender, suggesting that existing methods that condition on these observables cannot eliminate the problems we point out. In particular, we use the data described in Section 2.5 to sample pairs of borrowers who got a mortgage with the same lender, on the same day, with the same mortgage type (Conforming or FHA) and the same loan-level price adjustment (LLPA) and loan amount group, and compute the implied slopes given by their rate and point choices, which we plot in Figure A.14. Within this sample and among borrowers whose rate and point choices were not identical, 51% of Conforming pairs and 52% of FHA pairs had choices that dominated one another. Furthermore, Figure A.14 also shows that the implied exchange rate between them have large dispersion. Finally, R^2 of our regressions of points on rate in Tables A.11, as well as those of the literature, are well below 1. These observations all imply that there exists large unobserved heterogeneity in the menus, even conditional on narrow covariate and lender groups.

The reason that researchers have emphasized the role of discount points in studies of mortgage pricing discrimination is due to a presumption that Black borrowers has different preferences over menus items compared to White borrowers, perhaps due to differences in liquidity or borrowing constraints. Indeed, we observe Black and White borrowers making different choices in the data. In particular, Table 2.3 of our empirical application shows that Black borrowers pays more points than similar White borrowers for Conforming mortgages and fewer points than similar White borrowers for FHA mortgages. However, as we point out, there is a difficulty in attributing this difference in choices to a difference in the menus presented to Black and White borrowers or the heterogeneity in preferences between the racial groups. The advantage of our methodology is that it allows us to say that, under weak assumptions about preferences and menus, some of the racial difference in choices can *only* be rationalized by a difference in the distribution of menus facing Black and White borrowers among Conforming mortgage borrowers.

2.3 Robust Metrics for Assessing Discrimination in Menus

In this section, we define our robust metrics for assessing discrimination in menus. We consider equality in menus to mean equality in the distribution of menus offered to Black and non-Hispanic white borrowers, conditional on their observables. First, we present intuition for our approach in Section 2.3.1. We then specify our model of borrower choice more formally in Section 2.3.2. We keep our model fairly simple; menus are treated simply as a collection of items. Then, we define a direct test metric for equality in menus in Section 2.3.3 and a more welfare-relevant differences in menus metric for whether one group of consumers would like to switch to another distribution of menus in Section 2.3.4. We discuss the power of our identification results in Section 2.3.5. We leave inference on these metrics to Section 2.4. To increase clarity, we focus our exposition on the mortgage setting with two choice dimensions. Nevertheless, our general approach can be applied to contexts with more than two dimensions and may be useful in other settings where similar restrictions can be made.

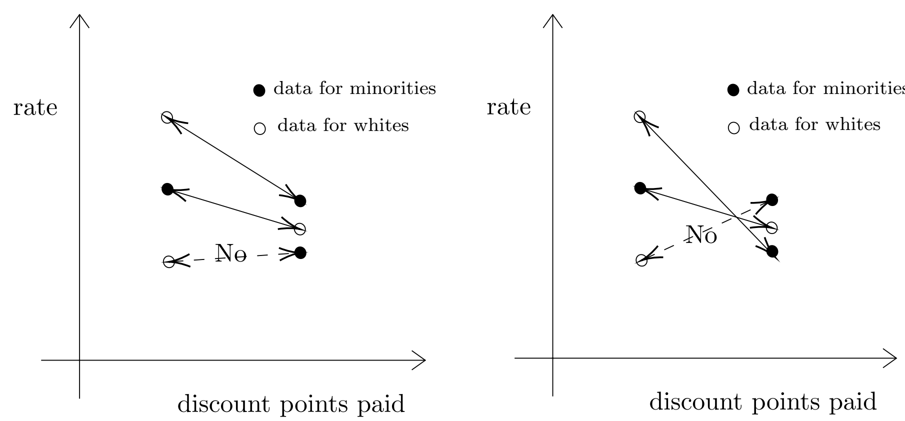
2.3.1 Intuition

Our metrics for assessing discrimination in menus are based on whether the data *can* be rationalized by a model in which all groups of borrowers faced the same distribution of menus. Ignoring for now sampling error to build intuition, the common thread in the false positive situations of Figures 2.4a and 2.5a is that there does exists a possible common distribution of menus that rationalizes the choice distributions of both minority and white borrowers, in the sense of there being a possible one-to-one match between minority and white borrowers where within each match both of the borrowers' choices could have come from the same menu. This is the criteria we use for assessing equality in menus.

By construction, our metrics are robust to false positives, since we will detect discrimination in menus only when there is no way to rationalize the data under the assumption of a common distribution of menus, regardless of the nature of any preference heterogeneity between the two groups of borrowers. Our methodology can also detect discrimination in menus where the existing heuristic approaches to the menu problem fail to do so. In particular, we illustrate in Figure 2.6 how the situation of Figures 2.4b, in which regressions controlling for either rates or points, would show a false negative, but the data fail our one-to-one matching condition under the assumption that borrowers with pairwise strictly dominated choices (that is, paying more in terms of both rates and points) could not have shared menus with one another. In other words, there is no way to construct a common distribution of menus for minority and white borrowers that explains the data. Analogously, the false negative example from comparing means in Figure 2.5b also fails our criterion.

Figure 2.6: How data from Figure 2.4b fail a "perfect matching" condition

(a) Possibility #1 for matching fails **(b)** Possibility #2 for matching also fails



To summarize, we define new metrics based on whether a set of preferences can rationalize the data under equality in menus, which are robust to the false positives. Furthermore, in some situations, such as those in Figures 2.4b and 2.5b, our metrics can detect discrimination when existing, heuristic approaches fail to do so. Nevertheless, a drawback of our approach is that it still leaves some possibility for false negatives, because the mere existence of a set of preferences that explains the data under equality in menus does not mean that it is the true set of preferences. This is a weakness compared with experimental approaches that may allow the researcher to directly observe menus, but the advantage of our approach is that it requires only data on outcomes and few assumptions on the data-generating process. Furthermore, our metrics are *sharp* in the partial identification sense that false negatives can only occur when there exists a common distribution of menus rationalizing the data that cannot be ruled out under the assumptions given.

2.3.2 Model

We now define our model more precisely. A menu item has values over k dimensions of attributes,[13] which we encode by $x \in \mathbb{X} \subset \mathbb{R}^k$. A menu $\mathbf{m} \subseteq \mathbb{X}$ is a set of such menu items that are presented to the borrowers. When borrower i is presented with a menu $\mathbf{m}$, we observe them making a choice that maximizes their utility over menu items $u_i(x)$. That is, we observe choices x_i where:

$$x_i \in \arg\max_x \{u_i(x) : x \in \mathbf{m}\}. \tag{2.3}$$

To keep the distribution of menus Lebesgue measurable and to implement the inference procedure of Section 2.4, we make the simplifying assumption that the set of items available to choose from is finite:

[13] In our context, the two dimensions of mortgage pricing are interest rates and discount points.

Assumption 1. *(Finiteness) The set of possible menu items, $\mathbb{X}$, is finite.*

Under Assumption 1, we can consider a probability distribution over possible menus $\mathbf{m} \sim \mathbf{M}$. This setup is fairly general and follows from consumers having standard (that is, complete and transitive) preferences over menu items.[14]

2.3.3 A robust test for inequality in menus

Suppose borrowers with a similar distribution of covariates in groups 1 and 2 face menus $\mathbf{m}_1 \sim \mathbf{M}_1$ for borrowers in group 1 and $\mathbf{m}_2 \sim \mathbf{M}_2$ for borrowers in group 2. The researcher wishes to compare the distribution menus across two groups of borrowers. More specifically, testing for equality of menus can be written as testing the null hypothesis that the distributions of menus being offered to both groups are equal. That is:

$$\mathcal{H}_0 : \mathbf{M}_1 = \mathbf{M}_2, \tag{2.4}$$

$$\mathcal{H}_1 : \neg\mathcal{H}_0. \tag{2.5}$$

To go from data on choices to statements about menus, we place restrictions on the

[14]Our Equation (2.3) that consumers maximize utility over menu items does rule out more behavioral representations of preferences over menus, such as in Gul and Pesendorfer (2001) and Ellis and Masatlioglu (2021) where the existence of some menu items may "tempt" consumers to change their rankings of other menu items. In that case, our statistical test of inequality in menus would still be valid, but the interpretation of our more welfare-relevant differences in menus (DIM) metric would be nuanced.

choices that could have been plausibly made from the same menus in terms of borrower preferences. For mortgages, it is plausible to assume that paying more in both interest rates and discount points is a dominated choice (and indeed would not be offered as a choice by the loan originator), which is the intuition we use in Figure 2.6 to reject equality in menus in that situation. We formalize this as Assumption 2:

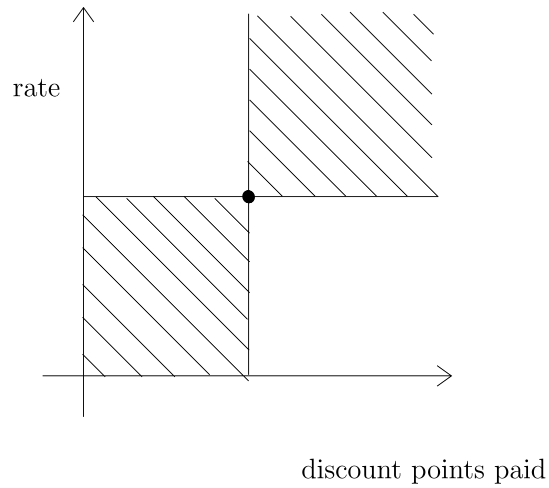
Assumption 2. *(Dominance) Paying more in rates and points is dominated. More formally, let* $x_1 = [r_1, y_1], x_2 = [r_2, y_2]$*, where* r_1, r_2 *represents rates and* y_1, y_2 *represents points. Then if* $r_1 > r_2, y_1 \geq y_2$*, or* $r_1 \geq r_2, y_1 > y_2$*,* $u_i(x_1) < u_i(x_2), \forall i$*.*

We illustrate in Figure 2.7 the restrictions on the observed choices that may come from the same menu under Assumption 2. The observed choice of the borrower, shown as the black dot, implies that they did not have the lower-left dashed quadrant available on their menu, since otherwise they would have chosen it. Similarly, any choice in the upper-right quadrant could not have been from the same menu as the choice indicated, as that agent would have an incentive to switch to that choice.

Note that while Assumption 2 is defined in the form of preferences of borrowers, it could have been alternatively formulated in the sense of dominated choices being unlikely to be *offered* to borrowers, which would have led to the same restrictions. Importantly, this implies that our test of equality in menus would hold even if borrowers are behavioral such that Assumption 2 fails, as long as lenders' menus fall within the range we specify in

Assumption 3. The main purpose of Assumption 2 is to allow us to compute a welfare-relevant metric for assessing differences in menus (DIM) later on.

Figure 2.7: Restriction on what cannot lie on the same menus from dominance.



While Assumption 2 is sufficient for rejecting equality in menus in the example situation of Figure 2.6, in our empirical application it is too weak of a restriction to be informative by itself. For our empirical analyses, we further adopt the industry rule of thumb of Bartlett et al. (2019a) that each point paid reduces the interest rate on a mortgage by one-eighth to one-fourth for conforming mortgages, with an expanded range for FHA mortgages. This is an assumption about menus rather than about preferences, which we formalize as Assumption 3:

Assumption 3. *(Restriction on Menus) In menus, each point paid reduces the rate by between* $[a, b]$*. More formally,* $x_1 = [r_1, y_1]$*, and* $x_2 = [r_2, y_2]$ *can lie on the same menu*

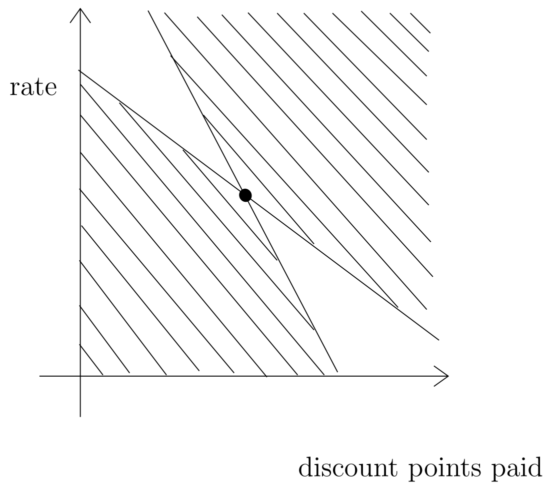
$\{x_1, x_2\} \subseteq \mathbf{m}$ *only if:*

$$a \leq \frac{r_2 - r_1}{y_2 - y_1} \leq b \text{ or } x_1 = x_2, \tag{2.6}$$

where $0 \leq a \leq b \leq \infty$, *and* $x_1 = [r_1, y_1], x_2 = [r_2, y_2]$, *with* r_1, r_2 *representing rates and* y_1, y_2 *representing points.*

We illustrate in Figure 2.8 the effect of defining a menu set based on Assumption 3. As Figure 2.8 indicates, the range of possible choices that could have come from the same menu as that of the consumer with the choice illustrated by the black dot is more restricted under this assumption, compared with using only dominance relationships in terms of preferences as in Assumption 2. Thus, this improves our ability to detect discrimination in menus. Since Assumption 3 is weakly stricter than Assumption 2 for our test statistic for equality in menus, the only place where Assumption 2 is used in our empirical application is to make welfare comparisons of menu distributions.

Figure 2.8: Restriction on what cannot lie on the same menus, from industry knowledge.

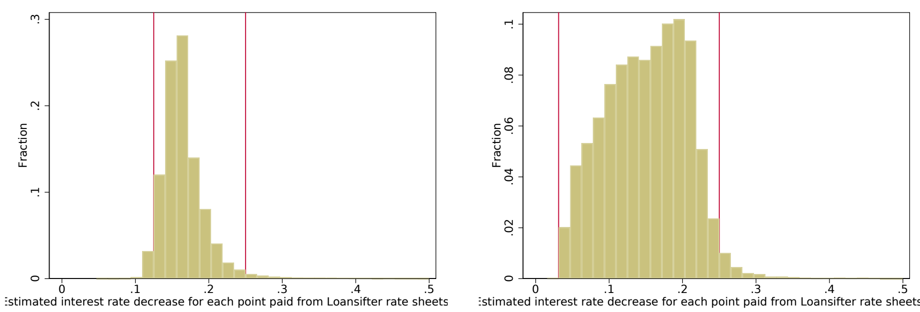


Empirically, we find that the industry rule of thumb from Bartlett et al. (2019a), which motivated Assumption 3, covers a vast majority of menus based on data from a sample of lender rate sheets that enumerates the rate and discount point menus. Using the 2014 LoanSifter data from Fuster, Lo, and Willen (2019), we estimate slopes of menus within their sample of 30-year purchase mortgages across seven different MSAs (Chicago, Houston, Los Angeles, Miami, New York City, Seattle, and San Francisco) and a range of loan amounts, FICO scores, and LTVs. The sample construction is discussed in more detail in Fuster, Lo, and Willen (2019). We estimate the slopes of the rate-point trade-off by taking the difference in the interpolated rate from 0 points to 2 points and dividing by 2. In this sample, the rule of thumb that each point paid is worth 1/8 to 1/4 of a point covers 94.4 percent of all rate sheet observations for conforming mortgages. For FHA mortgages, we use an expanded rule that each point is worth 1/32 to 1/4 in rate, which covers 97.6 percent of all observations. We illustrate this in Figure 2.9. Thus, we believe that our Assumption 3 is reasonable.

Figure 2.9: Rate sheet evidence for our menu slopes Assumption 3.

(a) Conforming mortgages, Bartlett et al. (2019a) restriction in red

(b) FHA mortgages, our expanded restriction in red



In addition, we see substantial heterogeneity between the menu slopes across lender-weeks in Figure 2.9, perhaps reflecting market power or lender- and time-specific costs. As we explained earlier, the existence of this heterogeneity interacted with possible differences in preferences between the two groups makes an approach like ours necessary for computing robust metrics of discrimination in menus.

Under Assumption 3, let $x_1 = [r_1, y_1], x_2 = [r_2, y_2]$. We define an indicator function for whether choices x_1, x_2 could have come from the same menu:

$$\varphi(x_1, x_2) = \begin{cases} 1 \text{ , if } a \leq \frac{r_2 - r_1}{y_2 - y_1} \leq b \text{ or } x_1 = x_2 \\ \\ 0 \text{ , otherwise.} \end{cases} \tag{2.7}$$

After defining this function, we have the following identification result for when vectors of choice probabilities $\mathbf{p}_1 = [p_1(x_1), p_1(x_2), \ldots], \mathbf{p}_2 = [p_2(x_1), p_2(x_2), \ldots]$ from observationally similar groups of borrowers can be rationalized under the null hypothesis of equality in the distribution of menus $\mathcal{H}_0 : M_1 = M_2$:

Theorem 1. *Under Assumptions 1 and 3, choice probabilities* $\mathbf{p}_1, \mathbf{p}_2$ *can be generated from the same underlying distribution of menus* $\mathbf{M}_1 = \mathbf{M}_2$ *if and only if there exists a coupling with probability mass function* $\pi(x_1, x_2) : \mathbb{X} \times \mathbb{X} \to [0, 1]$ *with implied marginal densities* $\sum_{x_2} \pi = \mathbf{p}_1, \sum_{x_1} \pi = \mathbf{p}_2$ *such that:*

$$T \equiv 1 - E_\pi \varphi(x_1, x_2) = 0. \tag{2.8}$$

Proof. See Appendix A.8.1. □

Theorem 1 formalizes the intuition from Section 2.2 that equality in menus should imply a way to "match" observations to one another such that each pair can come from the same set of menus. In particular, π serves as a coupling or "matching function," where each of its entries $\pi(x_1, x_2)$ represents the extent to which choice probabilities $p_1(x_1)$ and $p_2(x_2)$ came from the same menu, and the requirement of Equation (2.8) is that this coupling should lie entirely in the area where $\varphi(x_1, x_2) = 1$. Based on Theorem 1, we are ready to define a test statistic that looks at the extent to which this matching is deficient:

Definition 1. *Our test statistic for equality in menus, $\hat{T}$, given sample choice probabilities $\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2$, is:*

$$\hat{T} = \min_{\pi(x_1,x_2)} 1 - E_\pi \varphi(x_1, x_2), s.t. \sum_{x_2} \pi = \hat{\mathbf{p}}_1, \sum_{x_1} \pi = \hat{\mathbf{p}}_2, \pi \geq 0, \tag{2.9}$$

where the extent to which $\hat{T} > 0$ indicates a failure of the perfect matching condition in Theorem 1 in sample, which is evidence against equality in menus. It measures the fraction of each sample that cannot be matched to one another. The statistic $\hat{T}$ in Equation (2.9) is a finite dimensional optimal transport objective that can be efficiently computed using a linear program. To deal with sampling error inherent in $\hat{T}$, we discuss in Section 2.4 how hypothesis testing for $\hat{T}$ from Equation (2.9), as an optimal transport objective, can be consistently simulated.

2.3.4 Metric for assessing differences in menus

The direct test of inequality in the distribution of menus in Section 2.3.3, while indicative of discrimination in menus, has the drawback that it may not be the object of interest for researchers. The fact that the distribution of menus presented to one group is different in some aspect from the distribution of menus presented to another group may not be of welfare consequence, a problem that Abadie (2020) discusses in more detail. Rather, for the purposes of comparing menus across two distributions, we want a metric for assessing whether the distribution of menus from one group is meaningfully "better" than that of another group. For this purpose, we ask the question: *If Black consumers were instead assigned white menus, how much better off would they be?*

Conceptually, we consider the object of interest to be the change in welfare when Black consumers were instead assigned white menus, under an assignment rule $\pi(i, j)$ that maps each consumer $i \in \mathcal{I}_1$ from group 1 to the menu of consumer $j \in \mathcal{I}_2$ from group 2. Giving all consumers the same welfare weight, this objective can be represented by Equation (2.10):

$$\Delta W_{\mathcal{I}_1 \to \mathcal{I}_2, \pi} = \sum_{i \in \mathcal{I}_1, j \in \mathcal{I}_2} \pi(i, j)(u_i(\mathbf{m}_j) - u_i(\mathbf{m}_i)). \tag{2.10}$$

To get at $\Delta W_{\mathcal{I}_1 \to \mathcal{I}_2, \pi}$, we proxy for the utility difference $u_i(\mathbf{m}_j) - u_i(\mathbf{m}_i)$ through a metric $d_{i \to j}(\mathbf{m}_i, \mathbf{m}_j)$ that measures the extent to which consumer i would be willing to increase the interest rates on their loan in order to switch from menu $\mathbf{m}_i$ to menu $\mathbf{m}_j$:

$$u_i(\mathbf{m}_j) - u_i(\mathbf{m}_i) = d_{i\to j}(\mathbf{m}_i, \mathbf{m}_j) \equiv \sup\{\delta \in \mathbb{R} : u_i(\{x + \delta e^r, x \in \mathbf{m}_j\}) \geq u_i(\mathbf{m}_i)\}, \tag{2.11}$$

where e^r represents a basis vector that is equal to 1 at the location indexing interest rates. If consumers have constant marginal utility over interest rates such that utility can be represented as $u_i(x = [r, y]) = r + f(y)$, then $d_{i\to j}(\mathbf{m}_i, \mathbf{m}_j)$ is directly proportional to the utility change for consumer i after switching to menu j. Even if consumers do not have constant marginal utility over interest rates, it is still meaningful as a "willingness to pay" metric, since $d_{i\to j}(\mathbf{m}_i, \mathbf{m}_j)$ measures the extent to which consumer i would be willing to increase the interest rates to switch from $\mathbf{m}_i$ to $\mathbf{m}_j$. In the rest of this section we will show how we can compute an informative lower bound for this metric given the data, $\underline{d}_{i\to j}(x_i, x_j) \leq d_{i\to j}(\mathbf{m}_i, \mathbf{m}_j), x_i \in \mathbf{m}_i, x_j \in \mathbf{m}_j$, which then leads our differences in menus measure.[15]

To define our lower bound, we make an additional assumption that menus are complete in points, such that all choices of points are available to borrowers, which we formalize as Assumption 4. This is an approximation, since lenders may limit the choices of points

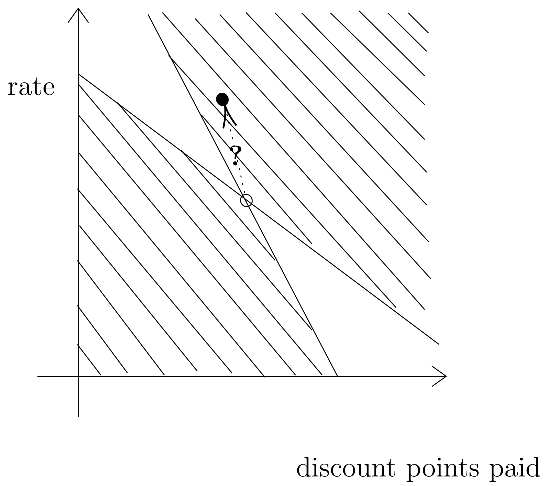
[15] We note that while we define our $\underline{DIM}_{1\to 2}$ over "willingness to pay" in terms of interest rates, we could have also defined it using points. However, due to the possible existence of cash on hand constraints that are likely binding for many consumers, constant marginal utility is very unlikely to hold for points, which makes welfare aggregation less credible.

to certain decimals (for example, 0.134, 0.266, ...) rather than literally the full range, but the practical implications of such small gaps in menus are likely small. Furthermore, there may be information constraints on the part of borrowers such that they do not "see" their full choice set (that is, some borrowers may not know that they can pay/receive points). Our metric is robust to this information constraints problem since as long as the subsets of choices borrowers "see" are held constant, our lower bound metric would remain valid.

Assumption 4. *(Completeness) The menus are complete in discount points. More specifically,* $\forall \mathbf{m}, \forall y', \exists x = [r, y'] \in \mathbf{m}$.

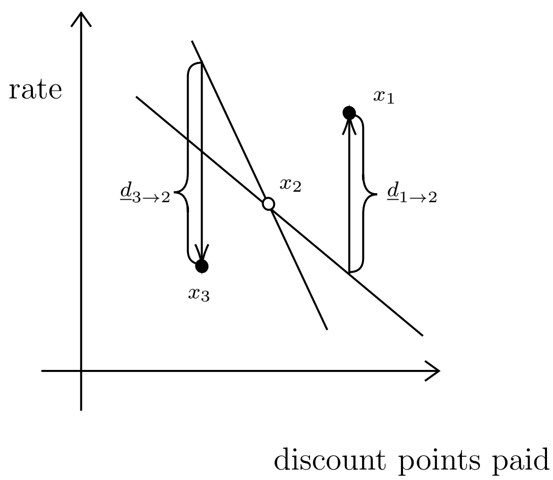
The effect of Assumption 4 is illustrated in Figure 2.10. Under the assumption that the mortgage menus are complete in discount points, we can meaningfully say that the minority borrower whose choice is represented by the black dot would have preferred the menu of the white borrower whose choice is represented by the white dot, because there exists a level of discount points such that all possible choices in the white borrower's menu dominate the minority borrower's choice. Otherwise, the minority borrower might not have preferred the white borrower's menu because the white borrower's menu could have been a singleton that the minority borrower dislikes. Therefore, adding the assumption of menu completeness in points sharpens the comparison of menus.

Figure 2.10: Impact of assuming that menus are complete in either rates or points



We illustrate in Figure 2.11 how we can construct a lower bound for the willingness to pay in terms of interest rates $\underline{d}_{i\to j}(x_i, x_j)$ under Assumption 4. There, borrower 1, who made a choice x_1 from an unobserved menu $\mathbf{m}_1$, has made a choice that is dominated (in terms of paying a higher rate at the same level of points) by any possible menu of borrower 2, who made a choice x_2 from a menu $\mathbf{m}_2$. This implies that, by revealed preference of borrower 1, the menu that borrower 2 faced is better than borrower 1's menu, or $\mathbf{m}_2 \succ_1 \mathbf{m}_1$. For borrower 1 to possibly become indifferent between $\mathbf{m}_1$ and $\mathbf{m}_2$, $\mathbf{m}_2$ needs to be shifted up by at least the amount indicated in the figure in the dimension of interest rates. In other words, a lower (sharp) lower bound for $d_{1\to 2}(\mathbf{m}_1, \mathbf{m}_2)$ is $\underline{d}_{1\to 2}(x_1, x_2)$, in the sense that borrower 1 is willing to pay at least $\underline{d}_{1\to 2}(x_1, x_2)$ more in interest rate in order to get borrower 2's menu. Similarly, the menu faced by borrower 2 would need to be shifted downward by at most the negative $\underline{d}_{3\to2}$ before it dominates x_3's choice. Therefore, borrower 3, with choice x_3, would need to pay at least $-\underline{d}_{3\to2}$, or receive at most $\underline{d}_{3\to2}$, before being willing to switch to borrower 2's menu.

Figure 2.11: Lower bound for borrower 1's willingness to pay a higher interest rate in order to get borrower 2's menu.



Formalizing the intuition from Figure 2.11, we define our lower bound for the willingness of borrower 1 to switch to borrower 2's menu under Assumptions 1 through 4 as follows:

$$\underline{d}_{1\to2}(x_1, x_2) \equiv r_1 - r_2 + a\max(y_1 - y_2, 0) + b\min(y_1 - y_2, 0) \leq d_{1\to2}(x_1, x_2), \quad (2.12)$$

where in the first line j indexes points, and in the second line we take it to the mortgage setting and let $x_1 = [r_1, y_1], x_2 = [r_2, y_2]$, where r_1, r_2 are rates and y_1, y_2 are points. The lower bound for how much borrower 1 would be willing to pay to switch to the menus of borrower 2, $\underline{d}_{1\to 2}(x_1, x_2)$, then allows us to define our differences in menus (DIM) metric:

Theorem 2. *Under Assumptions 1 through 4, choice probabilities* $\mathbf{p}_1, \mathbf{p}_2$ *implies a DIM measure:*

$$\underline{DIM}_{1\to 2} = \min_{\pi(x_1,x_2)} E_\pi \underline{d}_{1\to 2}(x_1, x_2), s.t. \sum_{x_2} \pi = \mathbf{p}_1, \sum_{x_1} \pi = \mathbf{p}_2, \pi \geq 0 \tag{2.13}$$

where $\underline{DIM}_{1\to 2}$ *serves as a lower bound for the average willingness to pay in terms of interest rates for borrowers in group 1 to switch menus with borrowers in group 2. If borrowers have constant marginal utility in interest rate, then:*

$$\underline{DIM}_{1\to 2} \leq \Delta W_{\mathcal{I}_1 \to \mathcal{I}_2, \pi} \tag{2.14}$$

where utility is measured in terms of interest rate paid.

Proof. See Appendix A.8.2. □

Theorem 2 shows that, when all consumers have the same constant marginal utility over interest rates (normalized to 1), our DIM metric is as a lower bound for the change in welfare for when consumers in group 1 are instead assigned menus from group 2 in an arbitrary way $\Delta W_{\mathcal{I}_1, 1\to 2, \pi_{1\to 2}}$. If instead consumers do not have constant marginal utility over interest rates, then the $\underline{DIM}_{1\to 2}$ metric could still be interpreted as the average increase in interest rates consumers in group 1 would be willing to pay in order to switch to menus from group 2. Furthermore, by Theorem 1, equality in menus would imply that $\underline{DIM}_{1\to 2} \leq 0$, so a finding that $\underline{DIM}_{1\to 2} > 0$ is also rejection of equality in menus in a welfare relevant way.

The sample analogue of the DIM metric follows immediately from Definition 2.

Definition 2. *Our empirical differences in menus metric,* $\underline{D\hat{I}M}_{1\to 2}$*, given choice probabilities* $\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2$*, is:*

$$\underline{D\hat{I}M}_{1\to 2} = \min_{\pi(x_1,x_2)} E_\pi \underline{d}_{1\to 2}(x_1,x_2), s.t. \sum_{x_2} \pi = \hat{\mathbf{p}}_1, \sum_{x_1} \pi = \hat{\mathbf{p}}_2, \pi \geq 0 \tag{2.15}$$

In terms of inference, the sample DIM metric in Definition 2 is also the value of a finite dimensional optimal transport problem. We discuss hypothesis testing in this class of problems in Section 2.4.

Extension: a "simple DIM"

An simple extension of our DIM metric is:[16]

$$\underline{DIM}^s_{1\to2} = E_{x_1\sim f_1(x_1),x_2\sim f_2(x_2)}\underline{d}_{1\to2}(x_1,x_2) \tag{2.16}$$

Where $f_1(x_1)$ is the distribution of borrower choices in group 1, and $f_2(x_2)$ is the distribution of borrower choices in group 2.

The $\underline{DIM}^s_{1\to2}$ metric as defined in Equation (2.16) has a straightforward interpretation: it is a lower bound for the expected willingness to pay of borrowers in group 1 after switching to *random* group 2 borrowers' menus. This is compared to the worst-possible assignment as in our Definition 2. The drawback of this metric is it requires an additional assumption to be a valid test of differences in menus, such as random assignment of menus conditional on the covariates groups within which it is computed. Such an assumption may be violated. For example, if, for borrowers in both groups, the heterogeneity in menus is due to unobservables that may be correlated with preferences, it is possible that the $\underline{DIM}^s_{1\to2}$ metric would produce a false positive. Our DIM metric in Definition 2, on the other hand,

[16] This section owes its existence to Daniel Ringo. The authors have omitted this possibility from the initial version of the paper in favor of the more robust alternative as presented in Definition 2. Dr. Ringo has independently arrived at this metric and used it to replicate our main results during his discussion of our paper at the 2021 Federal Reserve Day-Ahead Conference on Financial Markets and Institutions. In doing so, he illustrated the computational advantages of this simpler metric which may be useful in some contexts.

is robust to such a violation.

Nevertheless, the "simple DIM" metric has the advantage of being easy to compute and conduct inference on. Its sample counter-part is:

$$\underline{D\hat{I}M}^{s}_{1\to 2} = E_{x_1\sim\hat{f}_1(x_1),x_2\sim\hat{f}_2(x_2)}\underline{d}_{1\to 2}(x_1, x_2) \tag{2.17}$$

which may be computed by simply sampling random pairs of choices x_1, x_2, conditional on covariates, and then averaging their $\underline{d}_{1\to 2}(x_1, x_2)$ metric. Standard techniques such as bootstrapping may then be used for inference.

In our empirical application, the "simple DIM" gives similar qualitative results to our DIM metric. Thus, researchers may find such a metric useful for data exploration before moving on to our more robust metric. Furthermore, if the researcher believes that random assignment of menus conditional on the covariates groups is a valid assumption to make, the "simple DIM" is also a valid metric for assessing difference in menus.

2.3.5 When does our method have power?

Our metrics for differences in menus are robust in the sense that they are immune to the false positives problem from which the existing methods suffer, but they may still generate false negatives in that there exist scenarios where the data are rationalizable under equality in the distribution of menus, but in fact the distribution is different. Generally speaking, our method has power only to the extent that the borrowers' choices cannot be rationalized by an equal distribution of menus under the restrictions on preferences and menus that we have made. In this subsection, we give some examples of scenarios where this would or would not occur. In our later empirical analysis, we show that our methodology does have enough power to be useful in detecting discrimination in mortgage markets.

Figure 2.12 illustrates a scenario in which white and minority borrowers face the same default menu, but some white borrowers are offered a discretionary discount in terms of points. In the figure, the menu represented by the dashed line is shifted leftward for a white borrower, but all minority borrowers face the original menu. This shift makes the bottom-left choice by the white borrower not matchable to any of the minority borrowers' choices, so our one-to-one matching condition for equality in menus in Theorem 1 is broken. Our Theorem 1 would therefore have power to detect a difference in the distribution of menus offered to white and minority borrowers in this case. Furthermore, when the discretionary discount being given to white borrowers is large enough relative to the range of permissible menus, our DIM metric defined in Theorem 2 would also show that minority borrowers would be willing to pay to switch to the white borrowers' menus.

Figure 2.12: Power to detect discrimination when discretionary discounts offered to only some white borrowers make the data not rationalizable under equality in menus

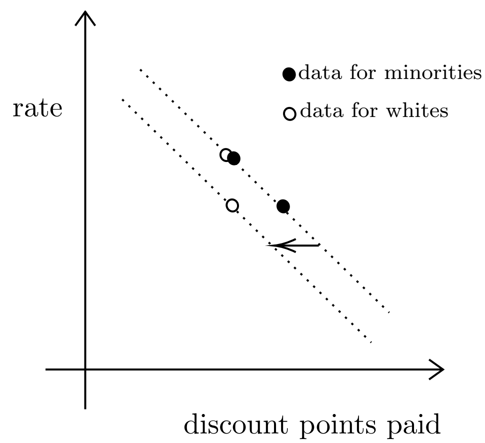
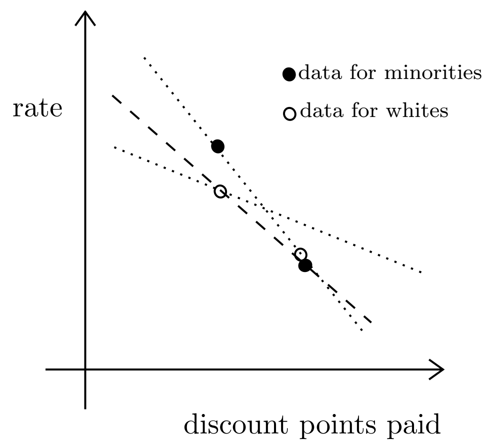


Figure 2.13 gives an example situation in which minority and white borrowers were presented with different menus as illustrated in the dotted lines, with white borrowers on average paying less in both rates and points, but the data are rationalizable under equality in menus. More specifically, while the data were generated by menus represented by the dotted lines which differ in distribution between white and minority borrowers, the data can be rationalized with one minority and one white borrower both choosing from the hypothetical menu represented by the dashed line and the other two borrowers choosing from the remaining dotted line. In this scenario, we cannot rule out that the true distribution of menus is represented by the dashed line plus a dotted line rather than the two dotted lines, and therefore our metrics from both Definition 1 and Definition 2 would fail to reject equality in menus.

Figure 2.13: No power when the data can be rationalized by (incorrect but plausible) distribution of menus that are equal across the racial groups



While our approach cannot detect differences in menus in the situation of Figure 2.13, there is a good reason for this: It is possible that the borrowers' choices were truly generated by the dashed line and the dotted line such that the racial groups did in fact face the same distribution of menus. It is indeed impossible to rule out that possibility given the data presented. This is the sense in which our metrics are sharp. Empirically in Section 2.5, we show that we are able to reject equality in menus for conforming mortgages, which shows that our method does have enough power to be useful in our mortgage setting.

2.4 Inference for Optimal Transport

In this section, we devise a new procedure for conducting hypothesis testing on the values of optimal transport problems that includes our metrics derived in Section 2.3. This is needed because the objective values of our model can be non-differentiable with respect to $\mathbf{p}_1, \mathbf{p}_2$ which implies that simple bootstrap methods may not be consistent (Fang and Santos, 2018). While some existing methods in the literature can be applied in more general contexts, they are either too conservative or converge too slowly based on our simulations, which we discuss at the end of this section. Our new procedure may be of independent interest for other researchers who wish to apply optimal transport methods to economics; some of those applications are listed in Galichon (2016).

We consider hypothesis testing on the value ϕ of a finite dimensional optimal transport problem with cost function $\varphi(x_1, x_2), x_1, x_2 \in \mathbb{X}$ and marginal distributions $\mathbf{p}_1, \mathbf{p}_2$:

$$\hat{\phi}(\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2) = \min_{\pi(x_1,x_2)} E_\pi \varphi \text{ s.t. } \sum_{x_2} \pi = \hat{\mathbf{p}}_1, \sum_{x_1} \pi = \hat{\mathbf{p}}_2, \pi \geq 0, \tag{2.18}$$

where the hypothesis is in the form of the value of the optimal transport $\phi(\mathbf{p}_1, \mathbf{p}_2)$ as a function of the true marginal distributions $\mathbf{p}_1, \mathbf{p}_2$ being less than or equal to some value ϕ_0:

$$\mathcal{H}_0 : \phi(\mathbf{p}_1, \mathbf{p}_2) \leq \phi_0, \tag{2.19}$$

$$\mathcal{H}_a : \phi(\mathbf{p}_1, \mathbf{p}_2) > \phi_0. \tag{2.20}$$

The form of the null hypothesis in Equation (2.19) is especially relevant to us because both our test of equality in menus (that is, whether $T \leq 0$) and our lower bound DIM metric (that is, whether $DIM \leq DIM_0$) can be expressed in terms of it. We provide a methodology to conduct this hypothesis test by looking at the asymptotic distribution of $\hat{\phi}$ and finding a critical value to compare the observed $\hat{\phi}$ to under $\mathcal{H}_0$. This can then be inverted into a confidence interval for the true value of ϕ.

As an overview, we combine the directional derivatives approach of Fang and Santos (2018) and Shapiro (1990) with a size correction of Romano, Shaikh, and Wolf (2014) and McCloskey (2017), which allows us to conduct hypothesis testing for optimal transport with uniform size control. To do so, we prove the directional differentiability for optimal transport problems on finite domains, and show how the general approach can be implemented as a linear program with complementarity constraints (LPCC).

We use the definition of Hadamard directional differentiation from Fang and Santos (2018), with some notational differences tailored to the optimal transport setting. Here, the value of an optimal transport represents a map $\phi : \mathbb{D}_\phi \to \mathbb{R}$, where $\mathbb{D}_\phi = \mathcal{P}_\mathbb{X} \times \mathcal{P}_\mathbb{X}$, $\mathcal{P}_\mathbb{X}$ is the set of probability measures on $\mathbb{X}$. Let $\mathbb{D}_0 = \{P_1 - P_2 : P_1 \in \mathcal{P}_\mathbb{X}, P_2 \in \mathcal{P}_\mathbb{X}\}$ be the set of possible differences in probability measures, and $\theta = \{p_1, p_2\}$ be the marginal distributions, then:

Definition 3. *(Fang and Santos, 2018) A map $\phi : \mathbb{D}_\phi \to \mathbb{R}$ is said to be Hadamard directionally differentiable at $\theta \in \mathbb{D}_\phi$ tangentially to the set $\mathbb{D}_0$, if there is a continuous* *linear map* $\phi'_\theta : \mathbb{D}_0 \to \mathbb{R}$ *such that:*

$$\lim_{n\to\infty} \left\| \frac{\phi(\theta + t_n \mathbf{h}_n) - \phi(\theta)}{t_n} - \phi'_\theta(\mathbf{h}) \right\| = 0 \tag{2.21}$$

for all sequences $\{\mathbf{h}_n\} \subset \mathbb{D}_0$ *and* $\{t_n\} \subset \mathbb{R}^+$*, such that* $t_n \to^+ 0, \mathbf{h}_n \to \mathbf{h} \in \mathbb{D}_0$ *as* $n \to \infty$ *and* $\theta + t_n \mathbf{h}_n \in \mathbb{D}_\phi$ *for all n.*

The main difference between the Hadamard directional differentiability and the typical notion of differentiability is that t_n approaches zero from the positive direction in Definition 3. Loosely speaking, the directional derivatives represent the change in the value of the function for a small change in its inputs "in the direction $\mathbf{h}$" for each $\mathbf{h}$.

We show in Theorem 3 that the value of all Monge-Kantorovich optimal transport problems with bounded cost functions on finite spaces is Hadamard directionally differentiable in the sense of Definition 3. In particular, Theorem 3 is a generalization of Sommerfeld and Munk (2018), which shows that the Wasserstein metric (the value of an optimal transport problem with the cost function restricted to distance metrics) on finite spaces is directionally differentiable.[17]

Theorem 3. *The value* ϕ *of an optimal transport problem with cost function* $\varphi(x_1, x_2), x_1, x_2 \in \mathbb{X}$*, where* $M = \sup|\varphi| < \infty$ *and* $\dim(\mathbb{X}) < \infty$*, is Hadamard directionally differentiable,*

[17] It is also related to Tameling, Sommerfeld, and Munk (2019), who prove that the Wasserstein distance on countable metric spaces is directionally differentiable. Our Theorem 3 can be similarly extended to countable metric spaces under the assumption that the cost function φ is continuous.

with derivative equal to:

$$\phi'_{\mathbf{p}_1,\mathbf{p}_2}(\mathbf{h}_1,\mathbf{h}_2) = \max_{\mathbf{u},\mathbf{v}\in\Psi^*(\mathbf{p}_1,\mathbf{p}_2)} \mathbf{h}_1^T\mathbf{u} + \mathbf{h}_2^T\mathbf{v} \tag{2.22}$$

where $\Psi^*(\mathbf{p}_1,\mathbf{p}_2) = \{\mathbf{u},\mathbf{v} : \mathbf{p}_1^T\mathbf{u} + \mathbf{p}_2^T\mathbf{v} = \phi(\mathbf{p}_1,\mathbf{p}_2), u(x_1) + v(x_2) \leq \varphi(x_1,x_2)\forall x_1,x_2\}$ *is the set of dual solutions to the linear programming problem, for all* $\{\mathbf{p}_1,\mathbf{p}_2\} \in \mathbb{D}_\phi$*, tangentially to the set* $\mathbb{D}_0$.

Proof. See Appendix A.8.3. □

Under i.i.d. sampling, we know that $\hat{\mathbf{p}}_1-\mathbf{p}_1$ and $\hat{\mathbf{p}}_2-\mathbf{p}_2$ approach a multivariate Normal distribution:

$$\hat{\mathbf{p}}_1 - \mathbf{p}_1 \to^d \mathrm{N}\left(0, \begin{bmatrix} p_{1,1}(1-p_{1,1}) & -p_{1,1}p_{1,2} & \dots \\ -p_{1,2}p_{1,1} & p_{1,2}(1-p_{1,2}) & \dots \\ \dots & \dots & \dots \end{bmatrix}\right), \tag{2.23}$$

and likewise for $\hat{\mathbf{p}}_2 - \mathbf{p}_2$, such that by construction, Assumptions 2.1 and 2.2 of Fang and Santos (2018) are satisfied. Then, Theorem 2.1 of Fang and Santos (2018) immediately implies that:

$$r_n[\phi(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2) - \phi(\mathbf{p}_1,\mathbf{p}_2)] = \phi'_{\mathbf{p}_1,\mathbf{p}_2}(r_n[(\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2) - (\mathbf{p}_1,\mathbf{p}_2)]) + o_p(1), \tag{2.24}$$

such that the asymptotic distribution of $\phi(\{\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2\})$ can be obtained via the directional Delta method. The remaining challenge for inference is that the true $\mathbf{p}_1, \mathbf{p}_2$ used in $\phi'_{\{\mathbf{p}_1,\mathbf{p}_2\}}$ in Equation 2.24 are not known, and thus must be estimated. While we could have used the plug-in analogue $\phi'_{\{\hat{\mathbf{p}}_1,\hat{\mathbf{p}}_2\}}$, that would converge only pointwise and not uniformly, which is known in the moment inequalities literature to be a poor approximation to the finite sample properties of estimators for which there are discontinuities in the pointwise asymptotic distribution (Andrews and Soares, 2010). We approach this problem following the logic of Romano, Shaikh, and Wolf (2014) and McCloskey (2017). More specifically, we suppose there exists confidence bands for $[\mathbf{p}_1, \mathbf{p}_2]$ at level β such that:

$$\limsup_{n\to\infty} \Pr([\mathbf{p}_1, \mathbf{p}_2] \in \hat{\mathcal{P}}_{n,\beta}) \geq 1 - \beta. \tag{2.25}$$

Many uniform confidence bands satisfying Equation (2.25) are available, for example from Montiel Olea and Plagborg-Møller (2019). Then, we take our estimate of the directional derivative as the maximum directional derivative within this confidence band:

$$\hat{\phi}'_{\beta}(\mathbf{h}_1, \mathbf{h}_2) = \max_{\mathbf{u},\mathbf{v}\in\Psi(\mathbf{p}_1,\mathbf{p}_2):[\mathbf{p}_1,\mathbf{p}_2]\in\hat{\mathcal{P}}_{\beta}} \mathbf{h}_1^T\mathbf{u} + \mathbf{h}_2^T\mathbf{v}, \tag{2.26}$$

where $\Psi = \{\mathbf{u}, \mathbf{v} : \mathbf{p}_1^T\mathbf{u} + \mathbf{p}_2^T\mathbf{v} \leq \phi_0, u(x_1) + v(x_2) \leq \varphi(x_1, x_2) \forall x_1, x_2\}$ are the set of dual solutions under the null hypothesis $\mathcal{H}_0 : \phi \leq \phi_0$.

Next, we define the critical value for ϕ by using our estimated maximum directional derivative from Equation (2.26) at level $1-\alpha+\beta$:

$$\hat{c}_{1-\alpha+\beta} = \inf\{c \in \mathbb{R} : \Pr(\hat{\phi}'_{\beta}(\mathbf{h}_1, \mathbf{h}_2) \leq c) \geq 1-\alpha+\beta\}, \tag{2.27}$$

where the distribution of $\mathbf{h}_1, \mathbf{h}_2$ is the asymptotic distribution of $\hat{\mathbf{p}}_1 - \mathbf{p}_1, \hat{\mathbf{p}}_2 - \mathbf{p}_2$. In the following Corollary 1, we will prove uniform coverage for when the observed value $\hat{\phi}$ is less than the critical value $\hat{c}_{1-\alpha+\beta}$, and suggest a computationally tractable version of it as a linear program with complementarity constraints (LPCC).[18]

Corollary 1. *Suppose we have uniform confidence bands for* $[\mathbf{p}_1, \mathbf{p}_2] \in \hat{\mathcal{P}}_\beta$ *that provide uniform coverage as in Equation (2.25), then under* $\mathcal{H}_0 : \phi \leq \phi_0$*:*

$$\limsup_{n\to\infty} \Pr(\hat{\phi} - \phi_0 \geq \hat{c}_{1-\alpha+\beta}) \leq \alpha, \tag{2.28}$$

where $\hat{c}_{n,1-\alpha+\beta}$ *is computed as in Equation (2.27).*

Proof. See Appendix A.8.4. □

Corollary 1 implies that uniformly valid hypothesis testing for the value of ϕ can be conducted by first computing a set of uniform confidence bands $[\mathbf{p}_1, \mathbf{p}_2] \in \hat{\mathcal{P}}_\beta$, and then

[18] As explained in Hsieh, Shi, and Shum (2021), LPCCs are well understood computationally and are implemented in software such as Knitro: https://www.artelys.com/docs/knitro/2_userGuide/complementarity.html.

maximizing over all directional derivatives within these bands as in Equation (2.26). Directly maximizing over the directional derivative defined in Equation (2.26) is computationally intensive, because it involves optimizing over a nonlinear dual value constraint $\mathbf{p}_1^T\mathbf{u} + \mathbf{p}_2^T\mathbf{v} \leq \phi_0$. To deal with this, we replace it with an equivalent complementary slackness formulation $\pi^T\mathbf{s} = 0, \pi \geq 0, \mathbf{s} \geq 0$ where $u(x_1) + v(x_2) + s(x_1, x_2) = \varphi(x_1, x_2)$, which implies that the elements of π and $\mathbf{s}$ cannot be positive simultaneously. Following the operations research shorthand, we represent this constraint by $\pi \leq 0 \perp \boldsymbol{s} \geq 0$. The derivation of this equivalence can be found in standard texts on optimal transport/linear programming. In particular, Hsieh, Shi, and Shum (2021) use a similar set of conditions for their projection method. Based on this equivalency, the problem of finding critical values for the null hypothesis $\mathcal{H}_0 : \phi \leq \phi_0$ versus the alternative $\mathcal{H}_a : \phi > \phi_0$ can be computed through the following LPCC which we implement:

$$\hat{\phi}'_{\beta}(\mathbf{h}_1, \mathbf{h}_2) = \max_{\mathbf{u},\mathbf{v},\mathbf{p}_1,\mathbf{p}_2,\mathbf{s},\pi} \mathbf{h}_1^T\mathbf{u} + \mathbf{h}_2^T\mathbf{v}, \tag{2.29}$$

$$\sum_{x_2} \pi = \mathbf{p}_1, \tag{2.30}$$

$$\sum_{x_1} \pi = \mathbf{p}_2, \tag{2.31}$$

$$E_{\pi}\varphi \leq \phi_0, \tag{2.32}$$

$$u(x_1) + v(x_2) + s(x_1, x_2) = \varphi(x_1, x_2), \tag{2.33}$$

$$[\mathbf{p}_1, \mathbf{p}_2] \in \hat{\mathcal{P}}_{\beta}, \tag{2.34}$$

$$\pi, \boldsymbol{s} \geq 0, \tag{2.35}$$

$$\pi \leq 0 \perp \boldsymbol{s} \geq 0. \tag{2.36}$$

To test our econometric approach, we conduct a Monte Carlo simulation with two possibilities for points $\{0, 1\}$ and five possibilities for rate $\{3, 3.25, 3.5, 3.75, 4\}$. Furthermore, Black and white borrowers choose each of the rate-point options with probability $\frac{1}{10}$ such that the null discrimination of no discrimination in menus is satisfied. We compute $\hat{\mathcal{P}}_{\beta}$ using the plug-in sup-t band of Montiel Olea and Plagborg-Møller (2019). Let $\beta = \frac{1}{10}\alpha$ following Romano, Shaikh, and Wolf (2014), and show the simulated probability that we reject equality in menus $\mathcal{H}_0 : \phi_0 = 0$ at the $\alpha = 0.01, 0.025, 0.05, 0.1$ levels in Table 2.1.

As Table 2.1 shows, our approach has the approximately correct size across a wide range of sample sizes and significance levels.

Table 2.1: Control of our size-corrected directional derivative approaches to inference

	Significance level			
Sample size	1%	2.5%	5%	10%
$n_1 = n_2 = 500$	0.6	2.3	4.7	9.3
$n_1 = n_2 = 1000$	0.8	2.9	5.3	9.3
$n_1 = n_2 = 5000$	1.1	2.3	4.7	9.1
$n_1 = n_2 = 10000$	0.7	2.0	3.8	7.6
$n_1 = n_2 = 50000$	0.8	1.9	4.1	8.6

Compared with existing methodology that can be applied to the optimal transport context, the advantage of our procedure is that it achieves uniform coverage without being overly conservative. In particular, Hsieh, Shi, and Shum (2021) have a novel projection method for parameter inference in mathematical programming problems, which is a broader set of problems than optimal transport, but their approach is conservative. In our empirical context, this conservativeness tends to make the confidence intervals uninformative. Another approach that is theoretically valid in this setting is the general m-out-of-n subsampling method of Politis and Romano (1994), but in addition to requiring the researcher to choose a subsample size m, it can require very large samples for convergence.

We look at the control of these competing approaches at the 5 percent significance level under our simulation setting in Appendix Table A.9. In that table, HSS (2021) refers to the projection method of Hsieh, Shi, and Shum (2021), the m-out-of-n subsampling approach refers to the method of Politis and Romano (1994), and in the final column, the size-corrected directional derivatives approach is replicated from Table 2.1 for comparison. Table A.9 shows that the HSS (2021) approach rejects the null with probability close to 0 percent, consistent with its conservativeness. On the other hand, the m-out-of-n subsampling approach tends to reject at rates greater than 5 percent for all values of m we tried, and appears to require more than 50,000 observations in order to converge to the correct rejection rate, which is significantly larger than our available sample size.[19]

In Section 2.5, we also report one-sided confidence intervals from the inversion of our hypothesis test. Nothing in our econometric theory precludes us from also testing the other direction and reporting two-sided confidence intervals instead.[20] However, since our economic theory is focused on getting a *lower bound* for the existence and welfare effects of discrimination, one-sided confidence intervals are particularly suitable for our purposes.

[19]A method related to m-out-of-n subsampling is the numerical bootstrap of Hong and Li (2020), which can be more data efficient than the subsampling method. But, we were not able to find a suitable choice of ε_n that converges to the correct coverage in our simulations using that method.

[20]The other direction may require more computational finesse, however, since it would involve taking the minimum of a maximum.

In summary, in this section we devised a new procedure for inference in optimal transport that is uniformly valid and not overly conservative for our empirical context. Our empirical analysis in Section 2.5 shows that we are able to strongly reject equality in menus for conforming mortgages using our methodology.

2.5 Empirical Estimates of Mortgage Discrimination

2.5.1 Data

We apply our methodology to a new data set constructed via matching the 2018–2019 public Home Mortgage Disclosure Act (HMDA) data to Optimal Blue rate locks. The public HDMA data contains loan level information along with indicators for borrower race and ethnicity, and we take the borrowers with an HMDA-derived race of "Black or African American" as our sample of Black borrowers, borrowers with a derived ethnicity of "Hispanic or Latino" as our sample of Hispanic borrowers, and borrowers with an HMDA-derived race of "White" along with a HMDA derived ethnicity of "Not Hispanic or Latino" as our sample of non-Hispanic white borrowers.

The HMDA data does not contain information on credit scores, and so we supplement it using data from Optimal Blue. Optimal Blue is a rate locking platform used by lenders that comprise of 40% of the US mortgage market, and has been used to study many questions regarding mortgage pricing.[21] The data construction process proceeds in two steps. First, we use year, loan type, loan purpose, occupancy, loan amount, interest rate, LTV (allowing for a 5% difference), and zip code to match each loan in Optimal Blue to originated loans in the 2018-2019 public HMDA data. For FHA mortgages, we further allow the mortgage insurance premium to be added to the loan amount. Second, we use the output of the first step to assign a de-identified set of lender IDs from HMDA to de-identified Optimal Blue lender IDs. Overall, we were able to uniquely match 55.4% of Optimal Blue rate locks to HMDA, almost all of which are unique. Our match rate is comparable to a 66% "lock pull-through rate," which is the rate at which locks turn into originated loans, that we understand is reasonable based on industry sources.

Starting from the uniquely matched HMDA-Optimal Blue matched data set, we further restrict our analysis to standard 30-year, new-purchase, fixed-rate, first-lien mortgages on owner-occupied, site-built properties without prepayment-penalties, balloon, interest-only, negative-amortization, or non-amortizing features. Table 2.2 compares the HMDA data, which include the complete set of mortgages originated in the United States with such characteristics, to our matched sample. We find that our matched sample has very similar average loan sizes, LTVs, rates, points, and the percentage composition of Black and Hispanic borrowers compared with the HMDA data, as can be seen from Table 2.2. One known

[21] For example, Bhutta, Fuster, and Hizmo (2019), Bhutta and Hizmo (2020), Bhutta and Ringo (2021), and Fuster et al. (2021).

caveat to this is that lenders using the Optimal Blue platform tend to be smaller lenders, with the larger lenders being more likely to have their own platform for rate locking.

Table 2.2: Comparison of means between the 2018–2019 HMDA data and our HMDA-Optimal Blue matched data

	Loan Size	LTV	Rate	Points	% Black	% Hispanic	*N*
Panel A: Conforming mortgages							
HMDA data	$261,566	84.4	4.50	0.08	4.5	9.0	*3,730,152*
Matched sample	$258,205	83.5	4.59	0.12	3.9	8.6	*817,588*
Panel B: FHA mortgages							
HMDA data	$215,144	96.1	4.57	0.07	14.1	19.7	*1,437,088*
Matched sample	$220,031	95.7	4.63	0.13	13.7	18.8	*360,202*

To control for the impact of lender and borrower characteristics within each loan program, we exactly match observations from Black and non-Hispanic white borrowers without replacement on groups of covariates, taking a random observation when multiple white borrowers can be matched to a Black borrower. This creates a sample containing equal numbers of Black and non-Hispanic white borrowers that are exactly matched on their covariates. If lenders offered Black and non-Hispanic white borrowers the same distribution of menus conditional on covariates, our covariate-matched sample of Black and non-Hispanic white borrowers should then have faced the same distribution of menus. The covariates that we used to match are lender, county, month of lock, four categories of loan amount, eight categories of FICO scores and nine categories of LTVs as defined in the GSE Loan-Level Price Adjustment (LLPA) matrix. Thus, we control for the effects of the interactions of all these covariates in our assessment of equality in menus. This set of controls is similar to what is used in Bartlett et al. (2021).

Summary statistics for our matched sample are in Table 2.3. Mortgage interest rates were measured based on a spread to the Freddie Mac Weekly Survey rate during the week of the rate lock. For our empirical analysis, we de-mean within each lender-county-month covariate group and round rates to the nearest eighths and points to the nearest halves, and we show in Table 2.3 that this step does not substantively change the mean differences in rates or points.[22]

As shown in Table 2.3, Black borrowers paid 4.8 basis points more in interest rate and 2.5 basis points more in points for conforming mortgages. On the other hand, they paid only 2.9 basis points more in interest rate and −3.2 basis points fewer points for FHA mort-

[22]Both the HMDA data and the Optimal Blue data contain information about discount points paid that sometimes disagrees with one another. While the literature has used both data sources, we focus on the HMDA information, because it is more precisely defined as the discount points to reduce the interest rate in Line A.01 of the Closing Cost Details page of the Closing Disclosure, according to CFPB's Regulation C. Furthermore, we find, as shown in Appendix Table A.10, that regression of the HMDA information on points on origination charges and total loan costs has a much stronger R^2, with a coefficient closer to 1, compared with the Optimal Blue information on points. In a regression controlling for the HMDA points, the effect of Optimal Blue points has minimal additional explanatory power for origination charges and total loan costs. We interpret this as suggestive evidence for there being more measurement error in the Optimal Blue definition of points. Our results using the Optimal Blue data on points are qualitatively similar.

Table 2.3: Summary statistics on the covariate matched sample

Panel A: Black and Non-Hispanic White Covariate Matched Sample						
	Conforming			FHA		
	Black	White	*Difference*	Black	White	*Difference*
Raw						
Rate Spread (bps)	36.2	31.4	*4.8*	48.3	45.4	*2.9*
Points (bps)	12.6	10.1	*2.5*	10.6	13.6	*-3.0*
De-meaned & rounded						
Rate Spread (bps)	3.1	-1.7	*4.8*	2.2	-0.7	*2.9*
Points	1.5	-0.9	*2.4*	-1.6	1.6	*-3.2*
Sample size	6,398	6,398		4,711	4,711	
Panel B: Hispanic and Non-Hispanic White Covariate Matched Sample						
	Conforming			FHA		
	Hispanic	White	*Difference*	Hispanic	White	*Difference*
Raw						
Rate Spread (bps)	38.3	34.5	*3.8*	49.3	45.4	*3.9*
Points (bps)	15.6	11.8	*3.8*	13.5	13.3	*-0.2*
De-meaned & rounded						
Rate Spread (bps)	2.5	-1.2	*3.7*	2.6	-1.2	*3.8*
Points	2.3	-1.3	*3.6*	-0.1	0.2	*-0.3*
Sample size	14,758	14,758		6,156	6,156	

gages. While this comparison of means cannot be interpreted as evidence for or against discrimination in menus (as we noted in Section 2.2), it does show that the distribution of data underlying conforming mortgages is different from the distribution of data for FHA mortgages. Appendix A.11 shows that the choice of heuristic used in analysis can lead to contradictory results in the FHA sample.

Note that because our data is limited to the set of mortgages that originated, our results should be interpreted as assessing differences in menus faced by the set of borrowers who took out a mortgage. The mortgage pricing discrimination literature including Bartlett et al. (2021) and Bhutta and Hizmo (2020) has argued that the pricing difference conditional on origination is policy relevant conditional on our observables. If borrowers that are offered worse prices are more likely to not get a mortgage at all, the extent of mortgage pricing differences at the application stage may be larger than what we detect at the origination stage. Our general methodology could be used to compare differences in menus at both stages if we had data on the chosen rates and points in the Loan Estimates and Closing Disclosures as well as at origination. HMDA does not yet collect such data on non-originated mortgage applications, and so we leave the assessment of mortgage pricing differences at the application stage for future research.

2.5.2 Analysis of differences in menus using our metrics

In this section we present our assessment of whether lenders offered minority borrowers worse menus of rates and points. Table 2.4 presents results from testing for equality between menus for Black borrowers and menus for white borrowers using our Definition 1. We find, as shown in columns (1) and (3), that our test statistic for inequality is positive and highly significant for both Black and Hispanic borrowers for conforming mortgages. More specifically, for conforming mortgages in the Black versus non-Hispanic white matched sample, our test statistic in column (1) is $\hat{T} = 2.69$, which indicates that 2.69 percent of Black borrowers' choices in the data could not have been matched to those of non-Hispanic white borrowers' choices that could have been on the same menu. Using our inference approach described in Section 2.4, we find that statistic is different from zero with $p < 0.01$. For Hispanic borrowers, our test statistic is $\hat{T} = 1.60$ in column (3), with $p < 0.05$. For FHA mortgages, on the other hand, we are unable to reject equality between the menus for Black and non-Hispanic white borrowers, with a test statistic of $\hat{T} = 0$ in column (2), but we are able to reject it at the 10 percent level for Hispanic and non-Hispanic white borrowers with a test statistic of $\hat{T} = 1.90$ in column (4).

Table 2.4: Results from our test of equality in menus ($\hat{T}$).

	Black vs Non-Hispanic White		Hispanic vs Non-Hispanic White	
	(1)	(2)	(3)	(4)
	Conforming	FHA	Conforming	FHA
Test statistic ($\hat{T}$)	2.69^{***}	0.00	1.60^{**}	1.90^{*}
95% CI	$[0.67, \infty)$	$[0.00, \infty)$	$[0.21, \infty)$	$[0.00, \infty)$
N	12,796	9,422	29,516	12,312

$^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 2.5 presents our results for our differences in menus metric as in Definition 2. Column (1) shows that for conforming mortgages, Black borrowers on average are willing to pay at least $\underline{D\hat{I}M}_{1\rightarrow 2} = 1.66$ basis points more in interest rates in order to get the non-Hispanic white borrowers' menus for conforming mortgages. This again rejects equality in menus and indicates that the distribution of menus faced by black borrowers is worse than that faced by white borrowers. Similarly, Column (3) shows that Hispanic borrowers are willing to pay at least $\underline{D\hat{I}M}_{1\rightarrow 2} = 1.46$ basis points more in order to get the non-Hispanic white borrowers' menus. Both of these statistics are significant at the 1% level. Our lower bound for how much more in interest rates non-Hispanic white borrowers would be willing to pay to switch to minority menus, on the other hand, are consistently negative. The magnitudes of these point estimates are small, even a small difference in interest rate at origination leads to a large difference in payments over the lifetime of the mortgage, as explained in Bartlett et al. (2019a).

Table 2.5: Results for our lower bound for the average interest rate increase (bps) needed for consumers to remain indifferent after switching to another group's menus ($\underline{D\hat{I}M}$)

	Black vs Non-Hispanic White		Hispanic vs Non-Hispanic White	
	Conforming	FHA	Conforming	FHA
	(1)	(2)	(3)	(4)
Minority to white ($\underline{D\hat{I}M}_{1\to 2}$)	1.66***	-1.83	1.46***	-0.91
95% CI	[1.09, ∞)	[-2.60, ∞)	[1.03, ∞)	[-1.89, ∞)
White to minority ($\underline{D\hat{I}M}_{2\to 1}$)	-6.16	-7.03	-6.04	-7.57
95% CI	[-6.82, ∞)	[-7.85, ∞)	[-6.40, ∞)	[-8.24, ∞)
N	12,796	9,422	29,516	12,312

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In summary, we find that lenders offered Black and Hispanic borrowers a distribution of menus that was different from the distribution they offered non-Hispanic white borrowers for conforming mortgages. In particular, 2.69 percent of Black borrowers' choices were unable to be matched to non-Hispanic white borrowers based on column (1) of Table 2.4, and for those unmatched Black borrowers their minimum willingness to pay to switch to unmatched non-Hispanic white borrowers' menus is at least 62 basis points (1.66/0.0269, combining with information in column (1) of Table 2.5). Similarly, column (3) of Table 2.4 shows that 1.46 percent of Hispanic borrowers' choices were unable to be matched to those of non-Hispanic white borrowers, with an average willingness to pay to switch to unmatched non-Hispanic white borrowers' menus being at least 91 basis points among those borrowers (1.46/0.0160, combining with information in column (3) of Table 2.5).

On the other hand, our results for FHA mortgages is more mixed: Column (2) of Table 2.4 shows that we cannot reject equality between the menus for Black and non-Hispanic white borrowers for FHA mortgages, and while column (4) of Table 2.4 shows that we are able to do so for Hispanic and non-Hispanic white borrowers at the 10 percent level, we are unable to reject a zero DIM metric in column (4) of Table 2.5 in terms of the average increase in rate that Hispanic borrowers would be willing to pay in order to receive the menus of non-Hispanic white borrowers.

2.5.3 Further analyses of differences in menus

To further explore where differences in menus occurs, we divide the sample into different LTV and FICO buckets. We detect more discrimination among the conforming, lower LTV and higher FICO (that is, more creditworthy) borrowers. In particular, focusing on the DIM metric, columns (1) and (3) of Table 2.6 show that we detect large differences in menus for conforming mortgages in the LTV under 75 and LTV of 75 to 80 categories, such that Black and Hispanic borrowers would be willing to pay 6-7 basis points to switch to non-Hispanic white menus. Over a 80 LTV, we can no longer say that Black and Hispanic borrowers would be willing to switch to non-Hispanic white menus. The results by FICO shown in Table 2.7 is more mixed, but suggests that our DIM metric is in general increasing in the FICO scores of conforming Black and Hispanic borrowers.

An internally consistent explanation for the fact that we primarily detect mortgage pricing discrimination among conforming, low-LTV borrowers is that lenders are more willing to offer discretionary discounts to the creditworthy non-Hispanic white borrowers, and less willing to do so for minority borrowers who are similarly creditworthy in terms of their underwriting variables. In particular, as shown in Ambrose, Conklin, and Lopez (2020), racial differences in default risk are very low among lower-LTV and higher-FICO borrowers, and in any case they are insured. Furthermore, it seems unlikely given the strict regulatory environment surrounding the mortgage market that lenders would condition their rate sheets and first offers based on race. Therefore, by process of elimination, we believe that the search and negotiation process, particularly for the more creditworthy borrowers, may play an important role in the within-lender disparate outcomes we find.

Table 2.6: Differences in menus comparing borrowers across categories of loan-to-value (LTV) ratio

	Black vs Non-Hispanic White		Hispanic vs Non-Hispanic White	
	(1)	(2)	(3)	(4)
	Conforming	FHA	Conforming	FHA
Panel A: Test of Equality in Menus ($\hat{T}$)				
LTV ≤ 75 ($\hat{T}$)	11.01**	-	8.26***	-
95% CI	[0.34, ∞)	-	[2.45, ∞)	-
$75 <$ LTV ≤ 80 ($\hat{T}$)	7.82***	-	7.30***	-
95% CI	[2.55, ∞)	-	[4.08, ∞)	-
$80 <$ LTV ≤ 90 ($\hat{T}$)	1.52	-	0.67	-
95% CI	[0, ∞)	-	[0, ∞)	-
$90 <$ LTV ≤ 95 ($\hat{T}$)	0.93	-	0.80	-
95% CI	[0, ∞)	-	[0, ∞)	-
LTV > 95 ($\hat{T}$)	1.82	0.17	0.90	1.77
95% CI	[0, ∞)	[0, ∞)	[0, ∞)	[0, ∞)
Panel B: Difference in Menus ($\hat{DIM}$) Metric				
LTV ≤ 75 ($\hat{T}$)	7.37***	-	5.80***	-
95% CI	[4.91, ∞)	-	[4.16, ∞)	-
$75 <$ LTV ≤ 80 ($\underline{\hat{DIM}}_{1\to2}$)	6.13***	-	5.16***	-
95% CI	[4.54, ∞)	-	[4.33, ∞)	-
$80 <$ LTV ≤ 90 ($\underline{\hat{DIM}}_{1\to2}$)	0.18	-	-1.30	-
95% CI	[-1.04, ∞)	-	[-2.25, ∞)	-
$90 <$ LTV ≤ 95 ($\underline{\hat{DIM}}_{1\to2}$)	-0.29	-	0.22	-
95% CI	[-0.91, ∞)	-	[-0.32, ∞)	-
LTV > 95 ($\underline{\hat{DIM}}_{1\to2}$)	-0.53	-1.76	-2.19	-1.18
95% CI	[-1.58, ∞)	[-2.45, ∞)	[-3.93, ∞)	[-1.73, ∞)
$N_{\text{LTV}\leq75}$	840	0	3,142	0
$N_{75<\text{LTV}\leq80}$	3,002	0	8,492	0
$N_{80<\text{LTV}\leq90}$	1,714	0	3,988	88
$N_{90<\text{LTV}\leq95}$	4,844	86	9,742	162
$N_{\text{LTV}>95}$	2,396	9,302	4,152	12,060

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.7: Differences in menus comparing borrowers across categories of FICO scores

	Black vs Non-Hispanic White		Hispanic vs Non-Hispanic White	
	(1)	(2)	(3)	(4)
	Conforming	FHA	Conforming	FHA
Panel A: Test of Equality in Menus ($\hat{T}$)				
FICO $\geq$ 740 ($\hat{T}$)	3.08***	9.23	2.59***	1.78
95% CI	[0.37, ∞)	[0, ∞)	[1.04, ∞)	[0, ∞)
700 $\leq$ FICO $<$ 740 ($\hat{T}$)	3.15	7.79	0.91	9.35**
95% CI	[0, ∞)	[0, ∞)	[0, ∞)	[0.87, ∞)
660 $\leq$ FICO $<$ 700 ($\hat{T}$)	4.38	2.22	3.99	6.53***
95% CI	[0, ∞)	[0, ∞)	[0, ∞)	[1.93, ∞)
620 $\leq$ FICO $<$ 660 ($\hat{T}$)	-	1.57	-	1.17
95% CI		[0, ∞)		[0, ∞)
FICO $<$ 620 ($\hat{T}$)	-	1.67	-	0.17
95% CI		[0, ∞)		[0, ∞)
Panel B: Difference in Menus ($\hat{DIM}$) Metric				
FICO $\geq$ 740 ($\underline{\hat{DIM}}_{1\to2}$)	1.84***	-0.59	2.16***	-1.33
95% CI	[1.10, ∞)	[-3.13, ∞)	[1.71, ∞)	[-3.84, ∞)
700 $\leq$ FICO $<$ 740 ($\underline{\hat{DIM}}_{1\to2}$)	2.19**	-2.97	-0.74	-1.30
95% CI	[0.90, ∞)	[-5.22, ∞)	[-1.64, ∞)	[-2.83, ∞)
660 $\leq$ FICO $<$ 700 ($\underline{\hat{DIM}}_{1\to2}$)	1.53	-2.09	1.16	0.72
95% CI	[-0.70, ∞)	[-3.30, ∞)	[-1.31, ∞)	[-0.23, ∞)
620 $\leq$ FICO $<$ 660 ($\underline{\hat{DIM}}_{1\to2}$)	-	-	-	-1.06
95% CI		[-2.36, ∞)		[-2.14, ∞)
FICO $<$ 620 ($\underline{\hat{DIM}}_{1\to2}$)	-	-2.82	-	-5.66
95% CI		[-3.89, ∞)		[-7.81, ∞)
$N_{\text{FICO}\geq 740}$	9,914	354	22,964	662
$N_{700\leq\text{FICO}<740}$	2,002	612	4,590	1,108
$N_{660\leq\text{FICO}<700}$	738	2,884	1,680	4,008
$N_{620\text{FICO}<660}$	142	4,372	282	5,188
$N_{\text{FICO}<620}$	0	1,200	0	1,346

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2.5.4 Robustness

In Section 2.5.2 we showed that lenders offered Black and Hispanic borrowers a less advantageous distribution of menus compared to observationally similar non-Hispanic white borrowers for conforming mortgages. In this section we conduct a series of robustness checks that rule out some explanations for this finding.

Multi-product discounts

One possible reason why minorities may receive worse menus of rates and points is that they are less likely to be offered multi-product discounts set by banks. For example, banks may offer discounts on rates and points based the amount of deposits a borrower hold with them. To check whether this explains our findings, Table 2.8 replicates our analysis for a sample of non-bank lenders. These are lenders that do not offer the depository services that banks do, and focus primarily on mortgage lending. As Table 2.8 shows, our results are robust to this sample restriction. Indeed, the point estimates for both our test of equality in menu ($\hat{T}$) and our differences in menus ($\underline{D\hat{I}M}$) increased under this specification. These results suggests that multi-product discounts offered by banks is not an explanation for the racial differences in menus.

Table 2.8: Analysis of differences in menus after restricting to non-bank lenders

	Black vs Non-Hispanic White		Hispanic vs Non-Hispanic White	
	(1)	(2)	(3)	(4)
	Conforming	FHA	Conforming	FHA
Panel A: Test of Equality in Menus ($\hat{T}$)				
$\hat{T}$	2.38^{***}	0.87	2.46^{***}	1.84
95% CI	$[0.77, \infty)$	$[0, \infty)$	$[0.86, \infty)$	$[0, \infty)$
Panel B: Difference in Menus ($D\hat{I}M$) Metric				
Minority to white ($\underline{D\hat{I}M}_{1\to2}$)	2.04^{***}	-1.37	1.96^{**}	-1.34
95% CI	$[1.37, \infty)$	$[-2.27, \infty)$	$[1.37, \infty)$	$[-2.03, \infty)$
White to minority ($\underline{D\hat{I}M}_{2\to1}$)	-6.57	-7.56	-6.67	-7.66
95% CI	$[-7.21, \infty)$	$[-8.36, \infty)$	$[-7.13, \infty)$	$[-8.36, \infty)$
N	10,128	8,106	24,358	10,852

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Loan originator compensation

Another possibility that may explain the differences in menus that we detect is that minority borrowers may be more likely to get mortgages from loan originators that charge more for their services. Within a mortgage bank, some loan originators may provide more services in exchange for a higher compensation, and such differences in compensation would be mechanically reflected in the menus they present to borrowers. Our Optimal Blue data contains information on loan originator compensation for a subsample of lenders. To check whether this explains our findings, we conduct a robustness check by matching on levels of loan originator compensation, rounded to the nearest 1 percent of the loan amount, in addition to our usual covariates. The results are shown in Table 2.9. In particular, Table 2.9 shows that while our $\hat{T}$ results lost significance likely due to the impact of smaller samples, the level and significance of our differences in menus $\underline{\hat{DIM}}$ results are qualitatively unchanged. In particular, the point estimate for our minority-to-white $\underline{\hat{DIM}}_{1\to2}$ measure for conforming mortgages increased from 1.66 to 2.18 for Black borrowers and from 1.46 to 1.54 for Hispanic borrowers. This suggests that differences in loan originator compensation does not explain our results.

Table 2.9: Analysis of differences in menus after further matching on loan originator compensation

	Black vs Non-Hispanic White		Hispanic vs Non-Hispanic White	
	(1)	(2)	(3)	(4)
	Conforming	FHA	Conforming	FHA
$\hat{T}$	4.62	3.48	2.13	4.83*
95% CI	[0, ∞)	[0, ∞)	[0, ∞)	[0, ∞)
Minority to white ($\underline{D\hat{I}M}_{1\to2}$)	2.18***	0.49	1.54**	-1.09
95% CI	[0.81, ∞)	[-0.94, ∞)	[0.55, ∞)	[-2.21, ∞)
White to minority ($\underline{D\hat{I}M}_{2\to1}$)	-4.86	-8.28	-5.87	-7.56
95% CI	[-6.11, ∞)	[-9.90, ∞)	[-6.83, ∞)	[-8.51, ∞)
N	1,818	1,818	5,290	2,438

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Endogeneity of LTV

Finally, we explore the possibility that the borrower may simultaneously choose their loan-to-value (LTV) ratio along with their rate/points choice. Our focus on the rate and points choice takes as given a plausible setting in which borrowers make their mortgage size choice before shopping for rates and points. Indeed, most price comparison websites require borrowers to enter their property and loan amount information before showing offers with different rate and points choices, which supports such an assumption. Nevertheless, it is possible that some borrowers may be able to jointly optimize on points and LTV. To investigate this possibility, we show in Table 2.10 the results of running our algorithm on only a sample of borrowers with an LTV of exactly 80. These borrowers are more likely to be constrained in their LTV choices. Results of Table 2.10 show that we detect large differences in menus among this sub-population, with magnitudes that are consistent with the 75-80 LTV sample in Table 2.6. Thus, we conclude that joint optimization of LTV and points is likely not a driver of the racial differences in menus that we detect.

Table 2.10: Analysis of differences in menus after restricting to borrowers with exactly 80 LTV

	Black vs Non-Hispanic White		Hispanic vs Non-Hispanic White	
	(1)	(2)	(3)	(4)
	Conforming	FHA	Conforming	FHA
$\hat{T}$	8.27^{***}	-	9.08^{***}	-
95% CI	$[2.42, \infty)$	-	$[5.25, \infty)$	-
Minority to white ($\underline{D\hat{I}M}_{1\to 2}$)	5.81^{***}	-	6.09^{***}	-
95% CI	$[4.21, \infty)$	-	$[5.13, \infty)$	-
White to minority ($\underline{D\hat{I}M}_{2\to 1}$)	-10.71	-	-5.87	-
95% CI	$[-11.90, \infty)$	-	$[-11.70, \infty)$	-
N	2,600	-	7,318	-

$^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$

2.6 Discussion

We identify a "menu problem" that confounds the estimation of mortgage pricing discrimination, which stems from potential unobserved preference heterogeneity across borrower groups. We also devise a new methodology for assessing differences in menus that is robust to such preference heterogeneity, and use it to produce new estimates of mortgage pricing discrimination. Empirically, we find that mortgage pricing differentials by race still exists, particularly among lower LTV conforming borrowers.

While the menu problem is broadly relevant, our specific methodology for addressing it is unlikely to be the final word on the problem. The main benefit of our approach is that it requires relatively few assumptions to be valid. One promising path for future methodological research is the exploration of alternative avenues of identification for assessing differences in menus.

Finally, the "menu problem" we identify suggests that the typical administrative data on borrower choices falls short of ideal in terms of assessing discrimination in menus. As such, collecting actual data on menus faced by borrowers of different races, perhaps in a field-experiment/audit study setting, would be another valuable contribution.

3

Mortgage Prepayment, Race, and Monetary Policy

[1]

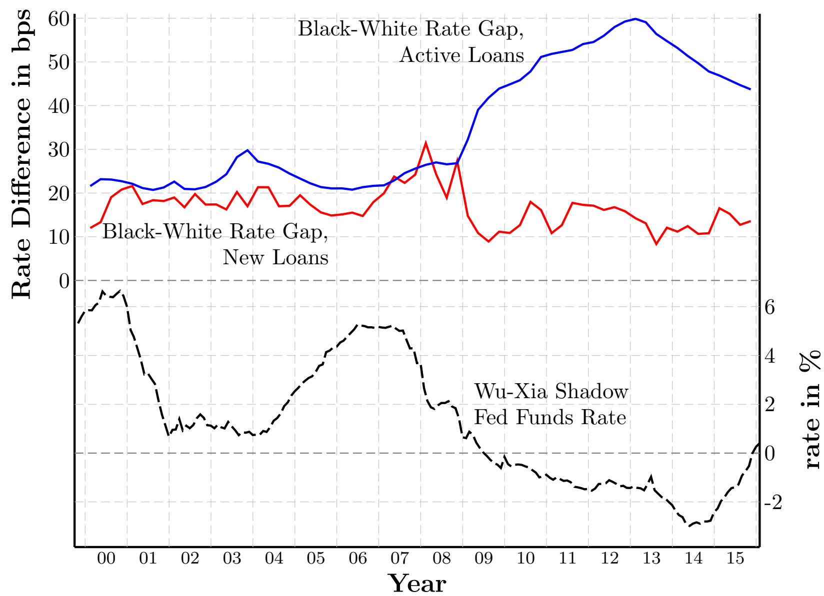
[1] This chapter is co-authored with Kris Gerardi and Paul Willen

3.1 Introduction

At the end of 2012, Black borrowers with mortgages insured by Fannie Mae or Freddie Mac (GSEs) paid interest rates that were approximately 60 basis points higher than those paid by white borrowers. This unconditional difference in the interest rates paid was not a new phenomenon, although the gap has waxed and waned over time, as depicted in Figure 3.1. What accounts for this gap? One natural explanation is that lenders charge higher rates to Black borrowers at origination, either because of racial bias or because loans to Black borrowers have characteristics associated with higher default rates, such as higher leverage or lower credit scores. A simple way to test this hypothesis is to take loan-level data and impose a rule that *all* borrowers who get a mortgage on the same day receive the nationally representative index interest rate prevailing that quarter. Two borrowers who get 30-year fixed rate mortgages in the second quarter of 2006, for example, would pay 6.6 percent regardless of race, or difference in credit score or loan-to-value ratio or any other observable measure of risk. Does this shrink the gap depicted in Figure 3.1? Yes, but not by much. Overall, we estimate that the gap would shrink by about 15 percent overall. In other words, eliminating all variation in interest rates at origination would do little to reduce racial inequality in mortgage rates. In this paper, we explain why.

We find that the key to explaining persistent interest rate gaps between Black and white borrowers has to do with refinancing. White borrowers are both much more likely to re-

Figure 3.1: Rates on outstanding mortgages insured by Fannie Mae and Freddie Mac: Black versus white Borrowers for mortgages originated from 2000–2015



finance in general and, more importantly, have a higher propensity to exploit reductions in interest rates. The quarterly hazard of prepayment due to refinance for a Black borrower with a loan from the GSEs is 0.75 percentage points lower than it is for a white GSE borrower, which corresponds to approximately 44 percent of the average quarterly refinance probability for all borrowers with GSE loans in our sample (1.71 percentage points). Given the trend decline in mortgage rates over the last 40 years, differences in refinance speeds alone would lead to lower rates for white borrowers. However, the problem is compounded by the fact that white borrowers appear to respond much more strongly to fluctuations in interest rates. In 2006 and 2007, when the Freddie Mac Primary Mortgage Market

Survey (PMMS) 30-year FRM rate averaged over 6 percent, higher than it had been since 2001, Black and white borrowers refinanced at roughly the same rate. In 2009 and 2010, when the PMMS 30-year FRM rate fell to historic lows of under 5 percent, white borrowers were almost twice as likely to refinance as Black borrowers.

What explains these differences in prepayment behavior across racial groups? Our rich data provide answers. We use the Credit Risk Insights Servicing McDash-Home Mortgage Disclosure Act (CRISM-HMDA) data set, a three-way match between administrative mortgage data from McDash, Home Mortgage Disclosure Act (HMDA) data collected by the Federal Reserve, and credit bureau data from Equifax. In contrast to data used in previous work in this area, the CRISM-HMDA data set allows us to distinguish between refinances and other mortgage prepayments, provides up-to-date information on borrower creditworthiness, and is nationally representative. We find that observable differences between Black and white borrowers account for approximately 80 percent of the difference in refinance rates. The typical Black borrower has a lower credit score, lower income, and higher leverage. A Black borrower is also more likely to be female and less likely to have a co-borrower. All of those factors lead to lower refinance propensities, regardless of race. However, a small gap remains even after controlling for these factors in addition to extremely fine geographic fixed effects. Suppose we take two borrowers living in the same ZIP code with the same credit score, income, and gender and who originated their loans in the same year and quarter. If one borrower is Black and the other is white, we show that the Black borrower is 0.15 percentage points less likely to refinance.

Further insights come from looking at responses to refinance incentives through the course of a loan. Refinance opportunities emerge for two reasons: macroeconomic and idiosyncratic. The main macro reason to refinance is to take advantage of lower interest rates. Idiosyncratic reasons stem from individual increases in creditworthiness such as a reduction in leverage from higher house prices or an increased credit score resulting from higher income and employment stability. We show that in our sample of GSE mortgages, minority borrowers are slightly *more* sensitive to idiosyncratic shocks than white borrowers. An 100 point increase in credit score leads to a 0.3 percent increase in the refinance probability for white borrowers and a 0.4 percent increase for Black borrowers and Hispanic borrowers. In short, the refinance gap that we document results entirely from differential responses to macroeconomic shocks.

The implications for monetary policy here are significant. By definition, expansionary monetary policy leads to lower interest rates and so, given the evidence we have presented, disproportionately benefits white borrowers and exacerbates mortgage rate inequality. While mortgage rates have always played a role in Federal Reserve policy, policymakers explicitly targeted mortgage rates only starting in 2008. Quantitative Easing (QE1), initiated in November of that year, consisted of large scale asset purchases (LSAPs) of mortgage-backed securities (MBS). The announcement of the LSAPs on November 25, 2008, provides a good laboratory to study the interaction between monetary policy and mortgage rate inequality. We compare the six months before with the six months after the announcement of QE1 and find that the quarterly refinance probability for white borrowers increased by 3.2 percentage points (per quarter) compared to an increase of only 1 percentage point for Black borrowers. This led to differential effects on outstanding mortgage rates, with a 21 basis point drop for the average white borrowers versus a 9 basis point drop for the average Black borrower in the six months following QE1.

Our research draws a distinction between the extensive and intensive margins of opportunity in credit markets. If we think of the intensive margin here as the mortgage rates offered to Black and white borrowers conditional on their decision to refinance, we find the intensive margin does not contribute that much to rate disparities. The extensive margin, defined here as whether Black borrowers refinance at all, appears to be more important.

We note that differences in mortgage rates paid by minority and white borrowers at origination, in the intensive margin, may not necessarily reflect lender discrimination. For FHA mortgages, Bhutta and Hizmo (2020) finds that the higher interest rate paid by minorities can be explained by their choices of upfront closing costs in the form of "points." More specifically, borrowers in the US have the option of reducing their interest rate by paying more in upfront closing costs to the lender, but minority borrowers are less likely to take up this option which explains their higher interest rate at origination. For conventional mortgages, on the other hand, Bartlett et al. (2019b) and Willen and Zhang (2021) find that Black and Hispanic borrowers pay about 2–8 basis points in higher interest rates compared to similar white borrowers in a way that is not explained by their upfront closing cost choices. For both products, we find that the racial differences in interest rate at origination is dwarfed by the differences in interest rate on active loans that emerges over time due to the refinancing differences between racial groups.

Our paper contributes to the literature on heterogeneity in monetary policy transmission in mortgage markets. Factors such as the type of mortgage contract (Calza, Monacelli, and Stracca (2013), Di Maggio et al. (2017)), house price growth (Beraja et al., 2018b), renting versus owning a home (Cloyne, Ferreira, and Surico, 2019), borrower age (Wong, 2019), income (Agarwal et al., 2020b), and lender concentration (Sunderam and Scharfstein, Agarwal et al. (2020a)) have all been found to lead to differential pass-through of monetary policy through the mortgage market across households and regions. Our finding that Black and Hispanic mortgagees benefit less from monetary policy is therefore complementary to these results.

Our paper is also related to the literature on racial differences in mortgage performance. Previous studies including Kelly (1995), Clapp et al. (2001), Deng and Gabriel (2006), Firestone, Van Order, and Zorn (2007), and Kau, Fang, and Munneke (2019b) document that minority borrowers prepay their mortgages at lower rates than white borrowers. There are some important differences between our analysis and these papers, however. First, none is able to distinguish between prepayments caused by home sales and those caused by refinances. Second, these studies use relatively narrow mortgage samples from either small geographic areas, short time periods, or individual banks/lenders. Third, previous studies focus exclusively on the pricing implications of prepayment differences and do not establish their implications for disparities in outstanding mortgage rates and the effect of monetary policy in exacerbating those differences.

Finally, our paper is related to the literature documenting that many borrowers appear to exercise their prepayment option in a suboptimal manner. Recently, Keys, Pope, and Pope (2016b) show that a significant fraction of financially unconstrained households (approximately 20 percent) do not refinance when it is optimal to do so. Johnson, Meier, and Toubia (2018b) find that more than 50 percent of borrowers neglect to refinance in a setting with zero up-front monetary costs and substantial gains in monthly payment savings. Agarwal, Ben-David, and Yao (2017b) find that many homebuyers appear to suffer from the sunk cost fallacy when deciding whether to refinance. Andersen et al. (2020b) decompose the inertia in refinancing into time and state dependence, and find significant heterogeneity in refinancing behavior by demographics in the Danish context, many of which (e.g. income, education, immigration status) can partially explain the racial differences in refinancing behavior. Earlier papers that find evidence of borrowers failing to refinance when it is likely beneficial to do so include Campbell (2006a), Chang and Yavas (2009), Deng and Quigley (2012), Green and LaCour-Little (1999b), and Schwartz (2006). Our paper focuses on documenting the large racial differences in refinancing and its implications for mortgage rate disparities, monetary policy, pricing, and mortgage contract design.

The remainder of the paper is organized as follows: Section 3.2 details our data and summary statistics. Section 3.3 contains the empirical approach we use and our results on differential prepayment tendencies across racial groups. Section 3.4 explores the implications of the differences in prepayment for the interest rate gap and the pass-through of monetary policy. Section 3.5 concludes.

3.2 Data and Summary Statistics

We use a novel data set that combines three sources of administrative data: Home Mortgage Disclosure Act (HMDA) data, Black Knight McDash mortgage servicing data (hereafter referred to as the McDash data), and credit bureau data from Equifax. The three data sources are linked together through two separate loan-level matches: a match between the HMDA and McDash databases, which we will refer to as the HMDA-McDash data set; and a match between the McDash and Equifax databases, which is referred to as CRISM (Equifax Credit Risk Insight Servicing McDash Database). We are then able to merge the two matched data sets, creating a final data set with information from all three sources, which we will refer to as the HMDA-McDash-CRISM data set. We will briefly describe each of the three sources of data below. We describe the details of the matching procedures in the Appendix (section A.1). We note that all information on borrower race and gender used in this analysis comes from the HDMA database and not from the CRISM database.

The HMDA database provides information on approximately 90 percent of US mortgage originations (see National Mortgage Database, 2017). The database contains a limited amount of information on borrower and loan characteristics at the time of mortgage origination, such as loan amount, borrower income, and borrower race and ethnicity. However, it does not contain some of the important underwriting variables, such as borrower credit scores, LTV ratios, loan maturities, and mortgage rates. In addition, since HMDA does not contain any information on mortgage performance over time, it is impossible to use the database to study prepayment and/or default behavior without merging with other datasets.

The McDash data set is constructed using information from mortgage servicers, which are financial institutions that are responsible for collecting payments from borrowers. Depending on the year, McDash covers between 60 and 80 percent of the US mortgage market and contains detailed information on the characteristics and performance of both purchase-money mortgages and refinance mortgages. For example, it includes information on borrower credit scores, LTV ratios, maturities, interest rates, documentation levels, and additional variables measured at the time of mortgage origination. Each loan is tracked at a monthly frequency from the month of origination until it is paid off voluntarily or involuntarily via the foreclosure process. The McDash database has been used by many papers in the literature to study questions around loan performance.[2] One key drawback to McDash is that, like most loan-level datasets, McDash does not provide information on the reason for prepayment of a loan. As far as McDash is concerned, sale and refinance are indistiguishable.

The CRISM data set consists of an anonymous match of credit bureau data from Equifax at the borrower level to McDash loans. The Equifax data are updated at a monthly frequency and include information on outstanding consumer loans and credit lines for the primary borrower as well as all co-borrowers associated with the McDash mortgage. We keep only observations that pertain to the primary mortgage borrower to avoid double counting. The CRISM data set provides the borrower's credit bureau information beginning six months before and ending six months after origination and termination of the McDash mortgage, respectively.The post-termination Equifax data provide two pieces of information which allows to overcome the limitations of the McDash data and identify whether a

[2]Examples include Keys, Seru, and Vig (2012), Piskorski, Seru, and Vig (2010), Jiang, Nelson, and Vytlacil (2013), Bubb and Kaufman (2014), Jiang, Nelson, and Vytlacil (2014), Kaufman (2014), Ding (2017), Fuster et al. (2018), Adelino, Gerardi, and Hartman-Glaser (2019), Agarwal, Ambrose, and Yao (2020) and Berger et al. (2020).

prepayment resulted from sale or refinanc: whether the borrower took out a new mortgage and whether the borrower moved.

Our method for identifying refinances is based on the work of Lambie-Hanson and Reid (2018) which uses similar data to study differences in refinancing behavior between subprime and prime borrowers. Specifically, we categorize a prepayment as a refinance if the borrower's address does not subsequently change and we observe a new first mortgage origination either just before or just after the time of the prepayment. We categorize a prepayment as a property sale and move if we observe the borrower's address change within a six-month window of the prepayment date. Our results are robust to narrowing the window to three months.

In addition to allowing us to distinguish between prepayments due to refinances and sales, the CRISM data set provides updated information about borrower credit scores, which we use in some of our empirical specifications to proxy for liquidity shocks. There are numerous alternative credit score measures in CRISM. Our analysis below focuses on the Equifax Risk Score 3.0 that was introduced in 2005 and predicts the likelihood of a consumer becoming seriously delinquent on any debt account. However, we have verified that our results are not sensitive to the particular credit score employed. For example, our results are virtually identical if we instead use FICO scores.

Our final HMDA-McDash-CRISM data set includes loans originated in the 2005–2015 (inclusive) period. The CRISM database begins in June 2005 but does include mortgages originated prior to 2005. However, the McDash database has poorer coverage of pre-2005 mortgage originations, and thus we include only originations on or after 2005 in our sample. Our data on loan performance extends through June 2020. In order to focus on a homogeneous mortgage product, we limit the sample to 30-year, fully amortizing, fixed-rate mortgages (FRMs) that were insured (against default risk) by the federal government. Specifically, we include loans that were acquired and insured by the GSEs (Fannie Mae and Freddie Mac) as well as loans that were insured by the Federal Housing Administration (FHA), which account for the vast majority of 30-year FRM originations during our sample period. We impose some additional sample restrictions to address outliers and missing information on key underwriting variables. Table A.4 in the Appendix lists all of the restrictions and how they impact the size of our sample. Most of the sample restrictions are adopted from Fuster et al. (2018), which uses the McDash-HMDA matched database. Finally, we include loans that were originated to Asian, Black, and white borrowers. Since HMDA provides separate identifiers for race and ethnicity, we are also able to distinguish between Hispanic/Latino white borrowers and white borrowers.[3]

[3] The race codes in HMDA are (1) American Indian or Alaska Native, (2) Asian, (3) Black or African American, (4) Native Hawaiian or other Pacific Islander, (5) white, (6) information not provided by applicant in mail, internet, or telephone application, (7) not applicable. We exclude groups 1) and 4) due to low observation counts. We also exclude groups 6) and 7). The ethnicity codes in HMDA are (1) Hispanic or Latino, (2) not Hispanic or Latino, (3) information not provided by applicant in mail, internet, or telephone application, (4) not applicable. We classify borrowers in the first group as "Hispanic," but we make the distinction only for white borrowers. We combine

Since most of our analysis is conducted on a panel data set at the quarterly frequency where the unit of observation is a loan-quarter, we work with a 7.5 percent random sample of the HMDA-McDash-CRISM data set to ease the computational burden. We also distinguish between the GSE and FHA loans in our sample and conduct our analysis on each group separately. The two loan types represent very different segments of the US mortgage market, as the FHA program typically focuses on more disadvantaged and riskier borrowers who have lower credit scores and lower down payments compared with the GSEs.

Tables 3.1 and 3.2 display summary statistics (means and standard deviations) for key observable variables in our sample of GSE and FHA loans, respectively. The top panel in each table displays mortgage and borrower characteristics at origination where the unit of observation is a loan (that is, one observation per loan), while the bottom panels display summary statistics of the time-varying variables included in our analysis where the unit of observation is a loan-quarter (that is, multiple observations per loan). In both tables we display statistics for the pooled sample of borrowers as well as separately for Black, Hispanic, and white borrowers. Asian borrowers are included in the pooled sample, but due to space constraints we do not include separate statistics for them in the table. The characteristics of Asian borrowers look very similar to white borrowers across most observable variables. There are large differences across the racial/ethnic categories for many of the observable

Hispanic and non-Hispanic Black borrowers into the single "Black" category.

variables in both tables. Focusing on the GSE sample, for example, white borrowers have significantly higher average credit scores and household incomes compared with Black and Hispanic borrowers (752 versus 715 and 730 and $97.6K versus $81.6K and $79.1K, respectively). White borrowers obtain significantly lower mortgage rates on average (5.18 versus 5.64 and 5.45, respectively).

Table 3.1: Summary Statistics: GSE Sample

Panel A: Fixed Characteristics

	All		Black		Hispanic		white	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Equifax Risk Score (100s points)	7.50	0.59	7.15	0.72	7.30	0.63	7.52	0.57
LTV (%)	72.6	15.9	75.6	15.4	74.0	15.9	72.5	16.0
Loan Amount ($100k)	2.12	1.13	1.84	1.00	1.98	1.03	2.10	1.11
Interest Rate (ppts)	5.20	1.02	5.64	1.09	5.45	1.06	5.18	1.01
Income ($1k)	97.6	64.0	81.6	51.9	79.1	51.5	98.6	64.7
Borrower Age (years)	46.3	13.4	48.4	13.2	45.2	12.8	46.5	13.5
Refinance (d)	0.538	0.499	0.588	0.492	0.513	0.500	0.543	0.498
Condo (d)	0.140	0.347	0.149	0.356	0.141	0.348	0.134	0.340
2-4 Family (d)	0.018	0.133	0.039	0.194	0.040	0.197	0.015	0.121
Low Documentation (d)	0.308	0.462	0.325	0.468	0.308	0.462	0.309	0.462
Non-Occupant Owner (d)	0.140	0.347	0.160	0.367	0.144	0.351	0.137	0.344
Female (d)	0.294	0.455	0.478	0.500	0.312	0.463	0.284	0.451
Co-applicant (d)	0.505	0.500	0.278	0.448	0.357	0.479	0.531	0.499
# Loans	800,806		32,753		43,269		676,986	

Panel B: Time-Varying Characteristics

	All		Black		Hispanic		white	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Qtrs since Orig	12.5	9.9	14.2	11.1	13.5	10.6	12.3	9.8
Interest Rate (ppts)	5.15	1.01	5.61	1.06	5.45	1.04	5.12	1.00
Call Option (ppts)	4.77	6.40	7.16	6.87	6.38	6.71	4.57	6.31
ADL (ppts)	-0.30	0.78	-0.20	0.87	-0.24	0.84	-0.31	0.78
SATO (ppts)	0.150	0.411	0.281	0.478	0.235	0.445	0.139	0.403
LTV Change	-4.60	14.70	-1.92	16.96	-1.74	20.38	-4.84	13.96
Negative Equity (d)	0.045	0.207	0.090	0.287	0.104	0.306	0.038	0.192
Risk Score Change (100s points)	0.070	0.530	-0.030	0.686	0.005	0.659	0.077	0.510
Prepay Refinance (ppts)	1.71	12.95	1.21	10.95	1.21	10.91	1.74	13.09
Prepay Sale (ppts)	0.96	9.76	0.54	7.35	0.63	7.93	1.02	10.03
Default (ppts)	0.35	5.90	0.87	9.28	0.80	8.92	0.30	5.44
# Loan-Quarters	15,460,588		730,648		924,765		12,970,785	

Table 3.2: Summary Statistics: FHA Sample

Panel A: Fixed Characteristics

	All		Black		Hispanic		white	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Equifax Risk Score (100s points)	6.84	0.67	6.53	0.71	6.77	0.64	6.89	0.66
LTV (%)	93.6	7.5	93.1	8.2	94.1	7.2	93.5	7.4
Loan Amount ($100k)	1.73	0.91	1.68	0.90	1.67	0.88	1.72	0.89
Interest Rate (ppts)	4.93	1.00	5.10	1.05	4.87	0.98	4.92	0.99
Income ($1k)	65.8	37.5	61.0	33.3	56.2	30.3	67.6	38.5
Borrower Age (years)	38.5	11.9	41.9	12.1	37.8	11.2	38.2	11.9
Refinance (d)	0.294	0.456	0.310	0.462	0.181	0.385	0.312	0.463
Condo (d)	0.115	0.318	0.155	0.362	0.110	0.312	0.106	0.308
2-4 Family (d)	0.014	0.119	0.024	0.154	0.031	0.174	0.010	0.101
Low Documentation (d)	0.190	0.393	0.207	0.405	0.164	0.370	0.192	0.394
Non-Occupant Owner (d)	0.033	0.178	0.033	0.179	0.026	0.158	0.034	0.181
Female (d)	0.353	0.478	0.530	0.499	0.318	0.466	0.333	0.471
Co-applicant (d)	0.414	0.493	0.248	0.432	0.367	0.482	0.445	0.497
# Loans	295,487		31,764		33,717		222,236	

Panel B: Time-Varying Characteristics

	All		Black		Hispanic		white	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Qtrs since Orig	13.3	10.2	15.0	11.1	13.7	10.3	12.9	10.0
Interest Rate (ppts)	4.93	0.98	5.11	1.01	4.90	0.96	4.92	0.97
Call Option (ppts)	4.77	6.53	5.68	6.68	4.88	6.50	4.64	6.50
ADL (ppts)	-0.60	0.87	-0.53	0.90	-0.61	0.88	-0.61	0.86
SATO (ppts)	0.116	0.346	0.165	0.376	0.158	0.356	0.104	0.338
Equity (%)	-9.25	14.85	-8.84	16.42	-12.38	16.86	-8.70	14.10
Negative Equity (d)	0.117	0.322	0.146	0.353	0.105	0.307	0.115	0.319
Risk Score Change (100s points)	0.016	0.697	-0.109	0.778	0.002	0.727	0.036	0.676
Prepay Refinance (ppts)	1.33	11.47	0.89	9.40	1.03	10.10	1.44	11.93
Prepay Sale (ppts)	0.94	9.67	0.47	6.87	0.62	7.86	1.08	10.33
Default (ppts)	0.89	9.41	1.58	12.47	0.90	9.42	0.81	8.94
# Loan-quarters	6,184,502		765,502		749,691		4,518,876	

Interestingly, Black borrowers are much more likely to be female (47.8 percent) compared with both Hispanic (31.2 percent) and white (28.4 percent) borrowers, while white borrowers are much more likely to have a co-applicant on the mortgage (53.1 percent) compared with Black (27.8 percent) and Hispanic (35.7 percent) borrowers. While we see similar discrepancies between the racial/ethnic groups in the FHA sample, the values of the group averages are quite different. For example, average credit scores and household income levels are significantly lower for all groups in the FHA sample compared with the GSE sample. In addition, LTV ratios are much higher in the FHA sample (93.6 percent versus 72.6 percent).

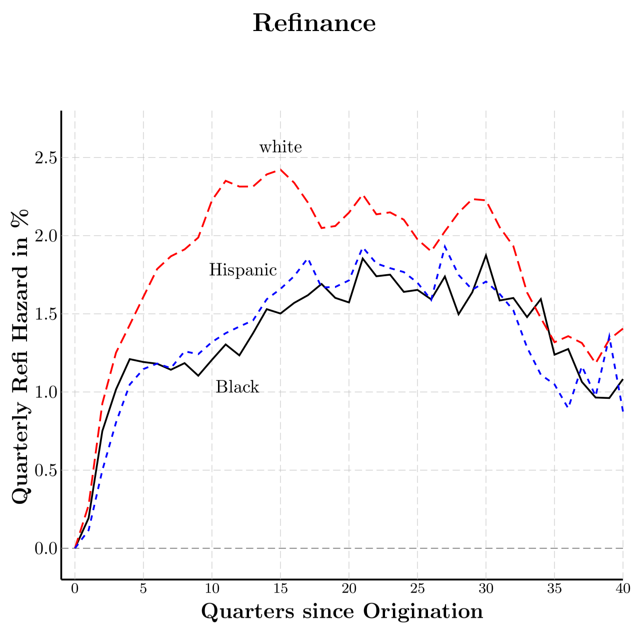
The bottom panel of Table 3.1 shows that the average prepayment rate due to refinancing is 1.71 percent per quarter in our GSE sample, while the average prepayment rate due to selling and moving is 0.96 percent per quarter. The average quarterly default rate, which we define as 90 days or more past due to be consistent with the previous literature, is only 0.35 percent. The average refinance rate is slightly lower in the FHA sample (1.33 percent) while the average sale hazard is virtually identical. The FHA default rate is more than twice as high (0.89 percent) as the GSE rate, which is unsurprising since the FHA program is characterized by mostly first-time homebuyers with low income and low credit scores. There are large differences in average refinance rates across racial/ethnic groups in both loan samples. In the GSE sample, white borrowers refinance at an average rate of 1.74 percent per quarter compared to only 1.21 percent for Black and Hispanic borrowers. There are similar differences between white and Black refinance rates in the FHA sample (1.44 percent versus 0.89 percent). There are also fairly large differences across racial/ethnic groups in both quarterly default rates as well as quarterly sale rates in both mortgage samples.

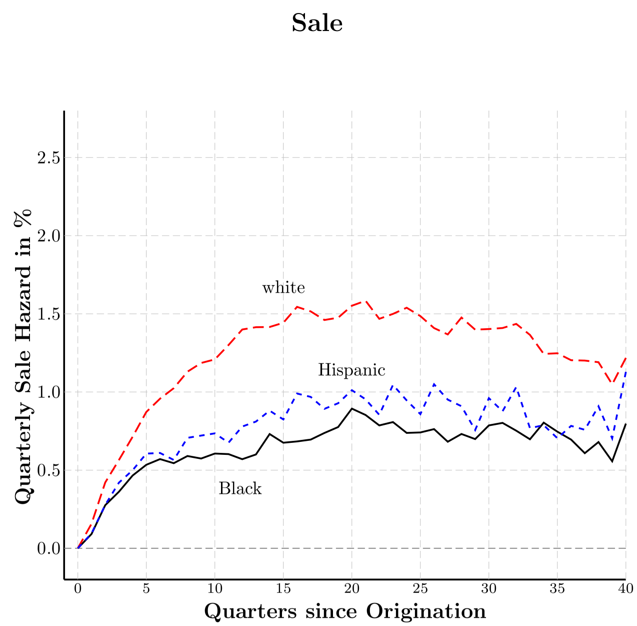
The left panel in Figure 3.2 plots Kaplan-Meier estimates of the hazard rates of prepayment due to refinancing by racial/ethnic group. These are unconditional, average quarterly rates as a function of duration that account for right censoring. Specifically, the Kaplan-Meier estimates are calculated as follows: Assuming that hazards occur at discrete times t_j where $t_j = t_{0+j}, j = 1, 2, ..., J$, if we define the number of loans that have reached time t_j without being terminated or censored as n_j, and the number of terminations due to refinancing at t_j as d_{pj}, then the Kaplan-Meier estimate of the hazard function is: $\lambda_p(t_j) = \frac{d_{pj}}{n_j}$. The figure shows that the unconditional hazard estimates of refinancing for white borrowers are approximately 1 to 1.5 percentage points higher than those for Black borrowers, and that difference is fairly constant over the first 10 years of the mortgage life cycle. Hispanic borrowers also have considerably lower refinance hazards compared with white borrowers, although the difference is not as large as it is for Black borrowers.

Figure 3.2: Kaplan Meier unconditional refinance and sale hazard rates

Refinance

Sale





The right panel in Figure 3.2 displays the Kaplan-Meier estimates of the sale hazards by racial/ethnic group. Consistent with the summary statistics discussed above, the level of the sale hazards is significantly lower than those of the refinance hazards. However, similar to the refinance estimates, we see large gaps between the hazards for white borrowers and our two minority borrower groups, as white households are much more likely to sell and move each quarter compared with Black and Hispanic households.

There are also significant differences in quarterly default rates across the racial/ethnic groups. Table 3.1 shows that in the GSE sample, Black borrowers are almost three times as likely to default as white borrowers (0.30 percent versus 0.87 percent per quarter). Hispanic borrowers are also characterized by relatively high default hazards (0.80 per per quarter). These differences are similar in the FHA sample.

3.3 Prepayment Results

In this section we present our main empirical results. We start by showing estimates of the gap between minority and white households in voluntary prepayments due to both refinancing and selling. Next, we test for differences in default behavior across the racial/ethnic borrower groups. We then show that differences in refinancing propensities are primarily due to differences in the extent to which borrowers refinance when their prepayment options are in the money, which are in turn mostly explained by observables such as income, credit scores, and loan-to-value ratios. Finally, we provide evidence that monetary policy has exacerbated the gaps in refinance propensities.

3.3.1 Empirical Setup

We examine differences in mortgage prepayment behavior due to refinance and home sale as well as differences in the propensity to default across racial/ethnic groups. For the bulk of our analysis we will focus on linear probability models (LPMs) that are estimated at a quarterly frequency, which is the frequency of our HMDA-McDash-CRISM data set. While linear probability models have some notable drawbacks,[4] they allow us to work with relatively large sample sizes and easily incorporate multiple levels of fixed effects, including highly disaggregated geographic fixed effects. We also consider logit models and show that the estimated average marginal effects are very similar to the LPM coefficient estimates.

Our primary specifications take the following general form:

$$Outcome_{it} = \beta_1 * Black_i + \beta_2 * Hispanic_i + \beta_3 * Asian_i + \gamma * X_{ijt} + \nu_g + \mu_v + \varepsilon_{it}, \quad (3.1)$$

where i indexes the individual mortgage and t indexes the year-quarter. We focus on three

[4] For example, Horrace and Oaxaca (2006) prove that the LPM can lead to biased and inconsistent estimates of structural parameters when the predicted values from the regression falls outside of the [0,1] interval. On the other hand, Jörn-Steffen Pischke notes that if marginal effects are of interest, the linear probability model will be a good approximation to the conditional expectation function: http://www.mostlyharmlesseconometrics.com/2012/07/probit-better-than-lpm/.

mortgage outcomes: the likelihood of voluntary prepayment due to refinance, prepayment due to home sale, and finally, the likelihood of default. Specifically, $Prepay_{it}^{refi}$ is an indicator variable that takes a value of 1 if loan i prepays due to the borrower refinancing in year-quarter t, and $Prepay_{it}^{sale}$ takes a value of 1 if loan i prepays due to the borrower selling the house and moving in year-quarter t. $Default_{it}$ is an analogous indicator variable that identifies when a loan defaults. Our focus will be on testing for differences in mortgage outcomes across the racial/ethnic borrower groups, which will include Black, Hispanic, Asian, and white borrowers. We specify indicator variables for each group in equation (3.1) with white borrowers representing the omitted category. Thus, the β coefficients will tell us how much more or less likely Black, Hispanic, and Asian borrowers are to prepay/default compared with white borrowers. X_{it} is a vector of control variables that include numerous mortgage and borrower characteristics, which we describe in detail below. Most of the control variables are time-invariant, but a few vary at the quarterly frequency. In some specifications we will include geographic fixed effects, ν_g, typically at the state level or ZIP code level, as well as vintage year-quarter fixed effects, μ_v. The standard errors are heteroskedasticity robust and are double clustered by county and year-quarter of origination.

Since the LPMs are estimated at a quarterly frequency, we are working in a hazard framework in which we model the likelihood of prepayment/default in year-quarter t conditional on the loan surviving through $t-1$. For example, if a loan is active for three years, at which point it prepays due to the borrower refinancing into a new loan, it will contribute 12 observations, with the $Prepay_{it}^{refi}$ indicator taking a value of 0 for the first 11 observations and a value of 1 for the final observation. Hazard models are commonly employed in the mortgage literature due to their ability to account for right-censored data (that is, loans that do not prepay during the sample period and are either still active at the end of the sample or exit the data set prior to the end of the sample period for other reasons). In Appendix A.5 we estimate our main LPM specifications assuming that 90-day delinquency, rather than foreclosure, is a competing risk in order to address the fact that borrowers who have missed mortgage payments are typically ineligible to refinance.

3.3.2 Prepayment due to Refinancing

We begin by estimating the LPM model in equation (3.1) for prepayment due to borrowers refinancing into new loans. Table 3.3 contains the results. Columns (1) through (6) report estimates for the GSE sample, while columns (7) through (9) show estimates for the FHA sample. In all columns, we have multiplied the dependent variable (refinance indicator) by 100 so that the coefficients can be interpreted in terms of percentage points. Column (1) reports estimates from our simplest specification, which includes vintage year-quarter fixed effects to control for unobservable changes in underwriting standards over time and a control for mortgage age (third-order polynomial). Our results are robust to higher order polynomials as well as one-year bins for loan age. Black (Hispanic) borrowers refinance at a rate that is 0.75 (0.69) percentage point lower than white borrowers on average, while Asian borrowers refinance at a rate that is 0.44 percentage point higher than white borrowers on average. These differences are all statistically significant as well as economically meaningful. The gap between Black and white borrowers is approximately 44 percent of the average quarterly refinance hazard among all GSE loans (1.71 percentage points).

To examine the extent to which lower prepayment likelihood of minority borrowers can be explained by their observable characteristics, in column (2) of Table 3.3 we include controls for some basic underwriting characteristics at origination, such as the borrower's credit score (Equifax risk score), LTV ratio, loan size, and indicator variables for loans that are refinances, less than full documentation of income/assets, and different property types (condominiums and 2 to 4 units). We also include indicators for missing information about documentation and property type. In addition, we include an estimate for the borrower's change in LTV over time, which we calculate by updating the mortgage balance based on the amortization schedule and the value of the property using the change in the county-level house price index since the quarter of origination. Finally, we add state fixed effects to the specification. The underwriting coefficient estimates are consistent with our expectations and with previous findings in the prepayment literature. Borrowers with higher credit scores and larger loan sizes refinance at faster rates. The differences in refinancing propensities between racial/ethnic groups decrease significantly with the addition of these controls. The difference between Black and white borrowers drops by almost 50 percent, from 0.75 to −0.38 percentage point per quarter. The differences between white borrowers and the other minority groups also decline (in absolute magnitude) with the addition of the underwriting controls. These results suggest that about half of the difference in refinance behavior can be attributed to differences in basic underwriting variables.

In column (3) we add more information about the borrower. First, we add three variables from the HMDA database: the borrower's reported income at the time of loan origination, an indicator for female borrowers, and an indicator for the presence of a co-applicant. Borrowers with higher income are more likely to refinance, while female borrowers are slightly less likely to do so. Borrowers with a co-applicant are more likely to prepay. We also control for three additional variables in column (3). We control for borrower age (second order polynomial), which we obtain from the CRISM data set. We control for the "moneyness" of the refinance option using a measure constructed by Deng, Quigley, and Van Order (2000b) that compares the present discounted value of the remaining stream of mortgage payments discounted at the borrower's current mortgage rate and the remaining stream discounted at the prevailing market rate. Specifically, the "Call Option" measure of Deng, Quigley, and Van Order (2000b) is calculated as:

$$Call\ Option_{i,k} = \frac{V_{i,m} - V_{i,r}}{V_{i,m}}$$

where

$$V_{i,m} = \sum_{s=1}^{TM_i - k_i} \frac{P_i}{(1 + m_i t)^s}$$
$$V_{i,r} = \sum_{s=1}^{TM_i - k_i} \frac{P_i}{(1 + r_i)^s}$$

and r_i is borrower i's mortgage rate, TM_i is the mortgage term, k_i is the age/seasoning of the mortgage, $m_i t$ is the prevailing market rate, and P_i is the mortgage payment. The larger the value of the "Call Option," the more the borrower would benefit from refinancing into a new loan with a lower rate and payment. We consider two versions of the call option variable that differ based on the assumed market rate faced by borrowers, $m_i t$. The first call option variable, which is included in column (3), assumes that all borrowers can refinance at the Freddie Mac PMMS rate. The second call option variable, which we introduce below, adjusts for the fact that borrowers often face different rates at which they can refinance due to the GSE loan-level pricing adjustments. "Call Option V2" uses the borrower's current credit score and an estimate of current LTV to make the adjustment. The FHA program does not employ risk-based pricing and thus, we do not make this adjustment in our FHA sample.

The third variable that we add to the specification in column (3), "SATO" (spread at origination), is the difference between the borrower's mortgage rate and the value of the

PMMS index in the year-quarter of origination. *SATO* is often included in prepayment models to proxy for unobserved constraints that may prevent a borrower from being able to obtain the prevailing market rate. Both *Call Option* and *SATO* are strong predictors of refinance propensities as a one standard deviation increase in "Call Option" (6.4 percentage points) is associated with a 1.48 percentage point increase in the refinance hazard, while a one standard deviation increase in SATO (0.41 percentage points) is associated with a -0.61 percentage point decrease in the refinance hazard. Finally, we specify credit score, LTV, and loan size in small, discrete bins, rather than as continuous variables in column (3), in order to allow for any non-linearities that might exist in their relationship with the propensity to refinance. The inclusion of all these additional controls and the more flexible functional forms has a small, but non-trivial effect on the prepayment gaps between racial/ethnic groups relative to basic underwriting variables.

Comparing the coefficients associated with the minority groups and the white group in columns (1) and (3), we see that approximately 41 percent of the gap remains for Black borrowers, while two-thirds of the gap remains for Hispanic borrowers. One possibility is that minority borrowers are more likely to experience adverse income or liquidity shocks that make it difficult to qualify for a new loan. While we do not have direct information on income or wealth over time, the CRISM data include updated information about borrower credit scores over the life of the mortgage. Since income and wealth shocks are correlated with the likelihood of debt repayment, updated credit scores should serve as a proxy for such shocks. In column (4) of Table 3.3 we use this information and include the change in the borrower's credit score between the current year-quarter and the quarter of origination. The change in the Risk Score is highly correlated with the likelihood of refinancing. A 100 point increase is associated with a 0.77 percentage point increase in the quarterly refinance hazard. In Appendix A.5 we also control for borrowers that have recently missed a mortgage payment since many lenders will not allow a borrower to refinance in such a situation. The addition of the variable also has a significant impact on the difference in refinance propensities between Black borrowers and white borrowers, as the gap declines by approximately 23 percent (0.071 percentage points). Therefore, the results displayed in columns (1)–(4) suggest that a majority of the refinancing gap between white and minority borrowers can be attributed to differences in underwriting variables and time-varying credit scores.

In column (5) we substitute the alternative call option variable, "Call Option V2", that accounts for the fact that borrowers looking to refinance into a GSE loan face different mortgage rates depending on their credit score and LTV ratio. The call option coefficient is unaffected, but the coefficient associated with the change in the Equifax risk score significantly decreases, from 0.769 to 0.348. This suggests that part of the reason that an increase in credit score increases the probability of refinancing is due to the lower rate that a higher credit score allows borrowers to obtain.

Next, we examine whether refinancing differences are more correlated with race or the neighborhoods that minorities live in. The specification reported in column (6) of Table 3.3 includes ZIP code fixed effects, so that differences in refinance hazards between groups in column (6) are estimated using variation only within a fairly small geographic area. This specification has the virtue of accounting for many sources of time-invariant, unobserved heterogeneity, such as the demographic composition of the ZIP code area as well as the average income/wealth of the area. Controlling for the ZIP code significantly narrows the gap between the racial/ethnic groups. Both the Black and Hispanic coefficients decline significantly. In Appendix A.4 we show results from specifications that include a full set of ZIP-code-by-year and ZIP-code-by-year-quarter fixed effects. These specifications control for time-varying, unobserved heterogeneity at the ZIP code level, and thus account for local economic shocks as well as local house price dynamics. The results are similar to those reported in column (6) of Table 3.3. Comparing columns (1) and (6), controlling for all observable variables at the time of mortgage origination, in addition to the change in credit scores, LTV, and ZIP code level differences, we can explain approximately 81 percent of the gap between the refinance behaviors of Black and white borrowers and about 61 percent of the gap between Hispanic and white borrowers. This again suggests that a race-neutral policy based on addressing refinancing gaps by neighborhood and borrower characteristics would resolve most of the gap in refinancing. Columns (7) through (9) in Table 3.3 display results corresponding to three LPM specifications estimated on our sample of FHA loans. Column (7) is analogous to column (1) and includes only vintage effects and controls for mortgage age, while column (8) is the same specification displayed in column (2), which includes basic underwriting controls such as credit score and LTV. Column (9) is the same specification as column (6) and includes ZIP code fixed effects. The differences in refinance hazards across the racial/ethnic groups in the FHA sample and the patterns across the different specifications are similar to what we found in the GSE sample. Notably, similar to the results that we obtained from the GSE sample, comparing columns (7) and (9), controlling for observable borrower and mortgage characteristics and geographic differences, explains a large fraction (about 64 percent) of the differences in refinance propensities between Black and white borrowers. Interestingly, this is not the case for Hispanic borrowers. Observables can explain only a trivial amount of the gap in refinance behavior between Hispanic and white borrowers in the FHA sample.

In Table A.19 in the Appendix we show that the results in Table 3.3 are not sensitive to our choice of the LPM, which assumes that the refinance hazard is a linear function of the covariates. The table contains estimated average marginal effects from logit models corresponding to each specification in Table 3.3, with the exception of the specification with ZIP code fixed effects which is not feasible in a logit model. The average marginal effects associated with the logits in all specifications are very close to the corresponding LPM coefficients.

3.3.3 Prepayment due to Selling

In Table 3.4 we test for prepayment differences across racial and ethnic groups due to home sales rather than refinancing activity. Our dependent variable in the LPM regressions is an indicator that takes a value of 1 if mortgage *i* voluntarily prepays in year-quarter *t* and we see that the borrower has moved and changed addresses (and 0 otherwise). We multiply the sale indicator by 100 so that the coefficients can be interpreted in terms of percentage points. The table is structured similarly to Table 3.3. The big difference is that we do not show specifications for both call option variables for the GSE sample. For the remainder of the paper we use the "Call Option V2" variable that depends on the borrower's updated credit score and LTV ratio in all GSE regressions.

Columns (1) and (5) show that there are large differences in the propensity to sell between minority and white households, controlling for only vintage effects and the age of the loan in both the GSE and FHA samples. Black borrowers are approximately 0.52 (0.64) percentage points less likely to sell their homes in a given quarter compared with white borrowers in the GSE (FHA) sample, which corresponds to about 54 percent (68 percent) of the quarterly sample average (0.96 and 0.94 percentage points, respectively). In contrast to our analysis of prepayment due to refinancing, adding detailed controls for borrower and mortgage characteristics in columns (2) and (8) does not have a large effect on the minority coefficients. The gap between sale hazards for Black borrowers and Hispanic borrowers decreases (in absolute magnitude) by approximately 20 percent in the GSE sample and even less in the FHA sample.

The addition of the HMDA variables (income, gender, and co-applicant indicator), updated credit score information, our proxy for the incentive to refinance (*Call Option*), and geographic fixed effects (state and ZIP code) does further attenuate the gaps between the sale propensities of the racial/ethnic groups. However, controlling for our detailed observable borrower and loan characteristics does not have as large of an effect on the differences in sale hazards as it did on the differences in refinance hazards that we see in Table 3.3. Comparing the simplest specification in column (1) with our most sophisticated specification in column (4), we can explain approximately one-third of the differences between sale hazards of minority and non-white Hispanic borrowers in our GSE loan sample. We find similar patterns in our FHA sample in columns (5)–(8).

3.3.4 Default

In this section we present results on differences in default hazards across racial/ethnic groups. We assume that borrowers default when they miss at least three payments (that is, 90-plus days past due), to be consistent with the recent mortgage default literature. Table 3.5 presents estimation results for the same LPM specifications in Tables 3.3 and 3.4, with one exception. We do not include a separate specification in which we add a control for changes in borrowers' credit scores since credit scores mechanically decline when borrowers miss mortgage payments. Again, we multiply the default indicator by 100 so that the coefficients can be interpreted in terms of percentage points.

In column (1) we see large differences between the default hazards of minority borrowers compared with white borrowers. Black borrowers with GSE loans are 0.44 percentage points more likely to default on their payments each quarter, which is more than 125 percent of the average default hazard in the GSE sample (0.35 percentage points). The addition of basic controls attenuates this difference, as the Black coefficient declines to 0.27 percentage points in column (2). Further controlling for our HMDA variables and ZIP code fixed effects reduces the coefficient to 0.11 percentage points. Comparing columns (1) and (4), we are able to explain more than 75 percent of the differences in Black versus white default hazards by controlling for observable borrower and loan characteristics and highly disaggregated geographic-by-time fixed effects. The pattern is similar for the estimated differences between Hispanic and white borrowers.

The default patterns are largely similar for Black borrowers in the FHA sample, but they are different for Hispanic borrowers. The gap for Hispanic borrowers of 0.165 percentage points is much smaller in column (5) (only 17 percent of the FHA sample average), and it becomes only marginally insignificant in column (7) when we add our controls and the

ZIP-code-by-year-quarter fixed effects.

These results are consistent with previous studies documenting that Black borrowers tend to have higher cumulative default probabilities than white borrowers, which has also been found in Canner, Gabriel, and Woolley (1991), Berkovec et al. (1994), and Berkovec et al. (1998). However, it is important to note that they are quite sensitive to the definition of default that one employs. In Table A.22 in the Appendix we estimate the same specifications but use a default definition that is based on involuntary prepayments due to foreclosure or pre-foreclosure distressed sales (that is, short sales) rather than serious delinquency. The table shows that minority loans are significantly more likely to end in involuntary prepayment when we do not control for borrower and mortgage characteristics. However, when those controls are included (in both the GSE and FHA samples), minority loans are *less* likely to involuntarily prepay. This pattern suggests that minority borrowers are more likely to miss payments, but are less likely to actually lose their homes to foreclosure. One possibility is that minority households are more likely to obtain loan modifications and avoid foreclosure.

3.3.5 Racial Differences in the Sensitivity of Refinancing to Mortgage Rates

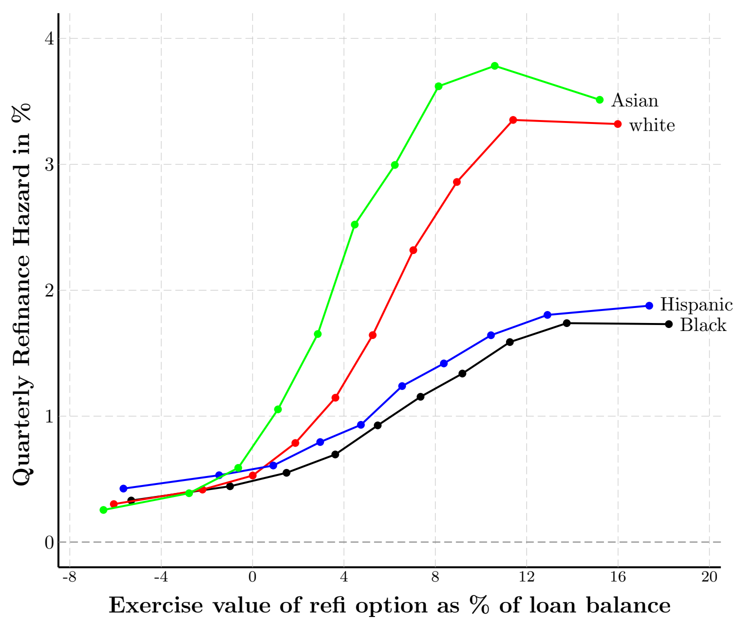
In this section we dig a bit deeper into the results on refinance disparities that we documented in section 3.3.2. The most common reason for borrowers to refinance is to take advantage of lower market rates and save on interest payments. In Table 3.3 we found that the *Call Option* variable, which proxies for the "moneyness" of the prepayment option and is driven by movements in market rates relative to the borrower's current rate, is an important predictor of the propensity to refinance. One possible explanation for the large disparities in refinancing behavior between our racial/ethnic groups is that minority borrowers are less likely or less able to refinance to take advantage of lower rates. We test this hypothesis by estimating a version of equation (3.1) in which we interact our race/ethnicity variables with *Call Option*:

$$Prepay_{it} = \beta * Black_i + \eta * Call\ Option_{it} + \delta * (Black_i * Call\ Option_{it}) + \gamma * X_{ijt} + \nu_g + \mu_v + \varepsilon_{it}, \tag{3.2}$$

If minority borrowers refinance less in response to increases in the exercise value of the prepayment option due to declines in rates, then we should expect to find $\delta < 0$.

Before discussing the results from estimating equation (3.2), we present a simple binned scatter plot in Figure 3.3 that shows the unconditional relationship between the propensity to refinance and *Call Option* for each of our racial/ethnic groups. Specifically, in Figure 3.3 we group the *Call Option* variable into deciles (separately for each racial/ethnic group) and then plot the average value of *Call Option* against the average quarterly refinance rate within each decile. The chart shows that all borrowers are more likely to refinance when the *Call Option* variable increases in magnitude, which corresponds to the prepayment option

Figure 3.3: Responses to gain from exercising the refinance option.



being deeper in the money. However, the figure clearly shows that white and Asian borrowers are much more likely to refinance compared with Black and Hispanic borrowers when their prepayment options are deeper in the money. When market rates are either higher or about the same as the borrowers' coupon, so that *Call Option* is negative or close to zero, all borrowers have a similarly low propensity to refinance. When market rates are lower relative to the rates on outstanding loans and *Call Option* becomes more positive, the refinance hazard for white and Asian borrowers increases by more than a factor of five to approximately 5 percentage points. In contrast, Black and Hispanic borrowers' average refinance hazard approximately doubles.

These patterns are confirmed in Table 3.6, which displays the results from estimating equation (3.2) separately for GSE and FHA loans. We start by displaying results for the LPM model without any interactions in columns (1) and (4). These specifications closely correspond to the specifications in columns (6) and (9) in Table 3.3, which include all of our controls as well as ZIP code fixed effects, but do not include Asian borrowers. In columns (2) and (5) we add the interactions between the Black and Hispanic dummies and the *Call Option* variable. The addition of the *Call Option* interaction results in the Black and Hispanic coefficients flipping from negative to positive, which implies that conditional on observable characteristics, minority borrowers are actually more likely to refinance when the value of the refinance option is out of the money. Both columns (2) and (5) show that Black and Hispanic borrowers are significantly less likely to refinance in response to market rates declining and the prepayment option becoming more valuable. In the GSE sample, a one standard deviation increase in *Call Option* (6.40 percentage points) increases the likelihood of refinancing by 1.5 percentage points for white borrowers but only 1.0 percentage points for Black and Hispanic borrowers. The patterns are similar in the FHA sample, as Black and Hispanic borrowers are about one-third less sensitive to changes in the call option variable compared to white borrowers. In Appendix Table A.23 we show that these results are robust to an alternative measure of the moneyness of the prepayment option, the one derived by Agarwal, Driscoll, and Laibson (2013a) that accounts for mobility, the volatility of interest rates, closing costs, and inflation.

The change in a borrower's credit score is another time-varying factor that we found to be a strong predictor of refinance behavior in Table 3.3 and that has an important effect on the estimated disparities in refinance hazards between minority and white borrowers. Our contention is that changes in credit scores over time likely reflect liquidity/income shocks that are impacting a borrower's ability to repay debt. In columns (3) and (6) we interact the change in credit score with the Black and Hispanic dummies to see if there are heterogeneous effects across racial/ethnic groups in their propensity to refinance in response to credit score changes. In the GSE sample, we find that the refinancing behavior of minority borrowers is *more* sensitive to credit scores changes compared to white borrowers. In contrast, minority FHA borrowers are statistically significantly less likely to refinance in response to credit score improvements compared with white borrowers.

3.3.6 The Effect of Monetary Policy on Refinance Gaps

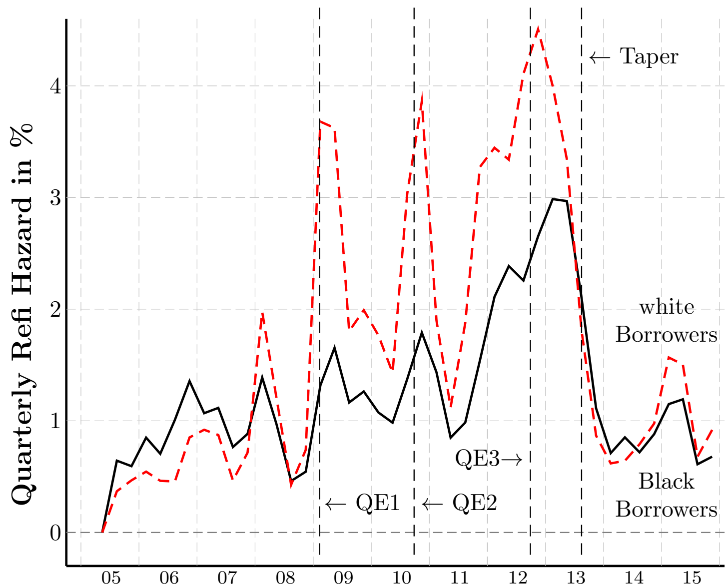
In the previous section we found that minority borrowers respond significantly less to changes in market rates that make their prepayment options more valuable compared with white borrowers. This suggests that expansionary monetary policy that lowers mortgage rates could exacerbate the refinancing disparities that we have documented. In this section we take a closer look at this issue.

Figure 3.4 displays unconditional, quarterly refinance rates for Black (solid black line) and white (dashed red line) GSE loans in calendar time over the course of our sample period. The figure shows that the refinance gap is relatively small in the first few years of the sample period, but then it increases dramatically beginning in early 2009, right about the time of the announcement of the Federal Reserve's first large-scale asset purchase program (LSAP), which is commonly referred to as quantitative easing (QE1). The gap closes in late 2009/early 2010, but then grows again in the third quarter of 2010, which coincides with the first Federal Reserve discussions of the second LSAP, QE2.[5] Finally, the third increase in the refinance gap in the figure occurs around the time of the announcement of the Fed's final LSAP, QE3, in the third quarter of 2012.

While Figure 3.4 is consistent with the hypothesis that the Federal Reserve's unconventional monetary policies played an important role in generating large differences in refinancing behavior between minority and white borrowers, it is not definitive. The post-crisis period was extremely turbulent, with many other policies and shocks impacting the mortgage market. For example, the Home Affordable Refinance Program (HARP) was initiated by the Federal Housing Finance Agency in March 2009 and was reformed and expanded in December 2011. For that reason, we implement a more direct test for monetary policy effects on the gap between the refinance behaviors of minority and white households. We

[5] On August 27, 2010, Fed Chairman Ben Bernanke stated in his speech at the Jackson Hole monetary policy conference, "A first option for providing additional monetary accommodation if necessary, is to expand the Federal Reserve's holdings of longer-term securities."

Figure 3.4: Unconditional quarterly refinance hazards for Black and white borrowers.



focus exclusively on our GSE sample since we showed in the previous section that the racial gaps in refinance behavior among FHA borrowers are not as sensitive to fluctuations in market rates. We also explicitly focus on the first LSAP, QE1. Beraja et al. (2018b) show that mortgage rates fell significantly and refinancing activity expanded considerably when QE1 was announced. Furthermore, the paper argues that unlike later LSAPs, QE1 was unanticipated by mortgage borrowers and thus provides for a fairly clean source of identification for the monetary policy effects on refinancing behavior.

QE1 was announced by the Federal Reserve on November 25, 2008, and initially called for purchases of as much as $500 billion in MBS guaranteed by the GSEs. It also announced purchases of as much as $100 billion in debt obligations of Fannie Mae, Freddie Mac, Ginnie Mae, and the Federal Home Loan Banks. In March 2009, the Federal Reserve announced that it would expand the program by purchasing $750 billion more in MBS. QE1 terminated at the end of the first quarter of 2010 with the Federal Reserve having purchased a total of $1.25 trillion in MBS. Fuster and Willen (2010) describes QE1 and its effect on the mortgage market in more detail.

We test whether QE1 exacerbated the gap between the refinance rates for minority and white borrowers by estimating the following difference-in-differences regression:

$$Prepay_{it} = \beta * Black_i + \eta * postQE1_t + \delta * (Black_i * postQE1_t) + \gamma * X_{ijt} + \nu_g + \mu_v + \varepsilon_{it}, \quad (3.3)$$

where *postQE1* is an indicator variable that equals 1 for the period after QE1 and 0 for the period before QE1 as well as the quarter in which QE1 was announced (2008:Q4). Since QE1 was announced at the end of November, refinances driven by QE1 are unlikely to be originated and show up in our data until the beginning of 2009:Q1. We consider two different sample windows around the QE1 announcement: a one-year window that consists of the two quarters before and after the announcement as well as a two-year window that consists of the 4 quarters before and after the announcement.

Table 3.7 displays the estimation results. In columns (1) through (3) we restrict the sample to a one-year window around QE1, and in columns (4) through (6) we expand the sample to a two-year window. For each window we estimate three specifications. First, we estimate an unconditional regression with no additional controls. Second, we estimate our preferred specification from above that includes all of our loan and borrower underwriting variables as well as ZIP code and origination year-quarter fixed effects (the specification in column (5) in Table 3.3). Finally we estimate a specification that adds interaction terms between our *postQE*1 dummy and credit scores as well as LTV ratios. This is a more flexible specification that allows QE1 to differentially impact borrowers with different credit scores and LTVs, and it is motivated by anecdotal evidence that suggests the refinancing boom that followed QE1 was driven mainly by borrowers with high credit scores and low LTVs.

The estimation results in Table 3.7 suggest that QE1 had a large effect on the racial gap in refinance propensities. According to column (1), Black borrowers were about 0.1 percentage point less likely to refinance in the six months prior to QE1 compared with white borrowers, and the gap increases by an order of magnitude to approximately 2.3 percentage points after QE1. While refinance propensities for white borrowers increased by 3.2 percentage points, an increase of approximately 520 percent of their rate prior to QE1 (0.6 percent points), Black borrowers increased their refinance rates by approximately 1.0 percentage point, an increase of approximately 200 percent of their pre-QE1 rate (0.5 percent points). Including our controls and fixed effects slightly changes the magnitudes, but the large effect of QE1 on refinance gaps remains. In column (2) Black and Hispanic conditional prepayment rates are actually significantly higher than those of white borrowers in the six months before QE1, but afterwards their rates fall more than 2.6 percentage points below the rates for white borrowers.

In column (3) the addition of the interactions between the *postQE*1 dummy and credit scores and LTVs slightly attenuates the gaps between refinances by minority and white borrowers that emerged after QE1, but the differences remain large and statistically significant. The interactions with credit score, which are displayed in the table, are striking. High-credit-score borrowers (Risk Score $>$ 740) increased their refinance rates by more than 3.7 percentage points after QE1 compared with an increase of about 0.77 percentage points for low-credit-score borrowers (Risk Score $\leq$ 600). Since the refinance differences across credit score bins are small in the period before QE1, these findings are consistent with the claim that the refinancing boom from QE1 was disproportionately driven by borrowers with high credit scores.

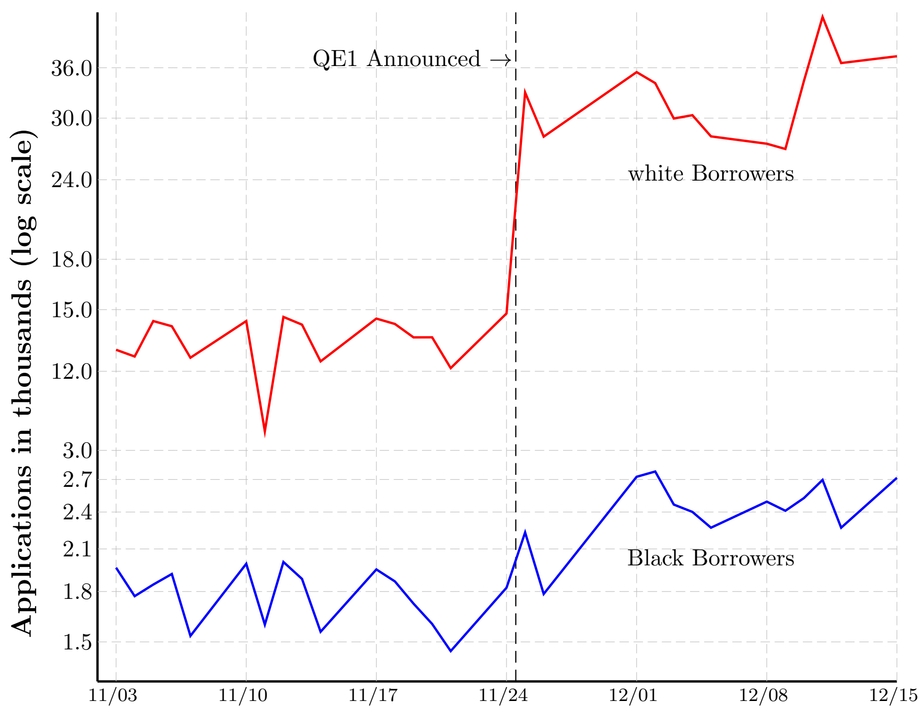
Columns (4) through (6) show that expanding the window size to one year slightly changes the estimated magnitudes, but does not alter the main patterns. QE1 appears to have generated a much larger increase in refinancing behavior by white borrowers compared with minority borrowers as well as high-credit-score borrowers compared with those with lower credit scores.

While the results in Table 3.7 strongly suggest that QE1 significantly exacerbated refinance disparities between minority and white borrowers, there were other major policies enacted around the same time as QE1, which could confound inference from our difference-in-differences estimator. For example, the Home Affordable Refinance Program (HARP) and the Home Affordable Modification Program were both enacted in March 2009, and may have had an impact on refinancing disparities across racial/ethnic groups. To address this issue and increase our confidence that QE1 really drove the differential changes in refinancing behavior in the relevant window, we zero in on the day of the announcement. To do this, we use confidential HMDA data, which provide information on the exact day on which a borrower applied for a mortgage. Figure 3.5 shows that from November 24 to 25, refinance applications by white borrowers increased from 15,000 to more than 30,000, an increase of over 100 percent. Over those same days, applications by Black borrowers increased from 1,800 to 2,100, a gain of a little over 15 percent. Black borrowers did make further gains over the next week, but overall, over the next few weeks, the maximum increase relative to November 24 was about 50 percent, whereas for white borrowers the increase rarely fell below 100 percent.

3.4 Implications for Mortgage Rate Disparities

The literature on discrimination in mortgage market pricing focuses almost exclusively on the flow of mortgage rates—the difference in rates obtained by minority and white borrowers at the time of origination. In this section we show that the large differences across

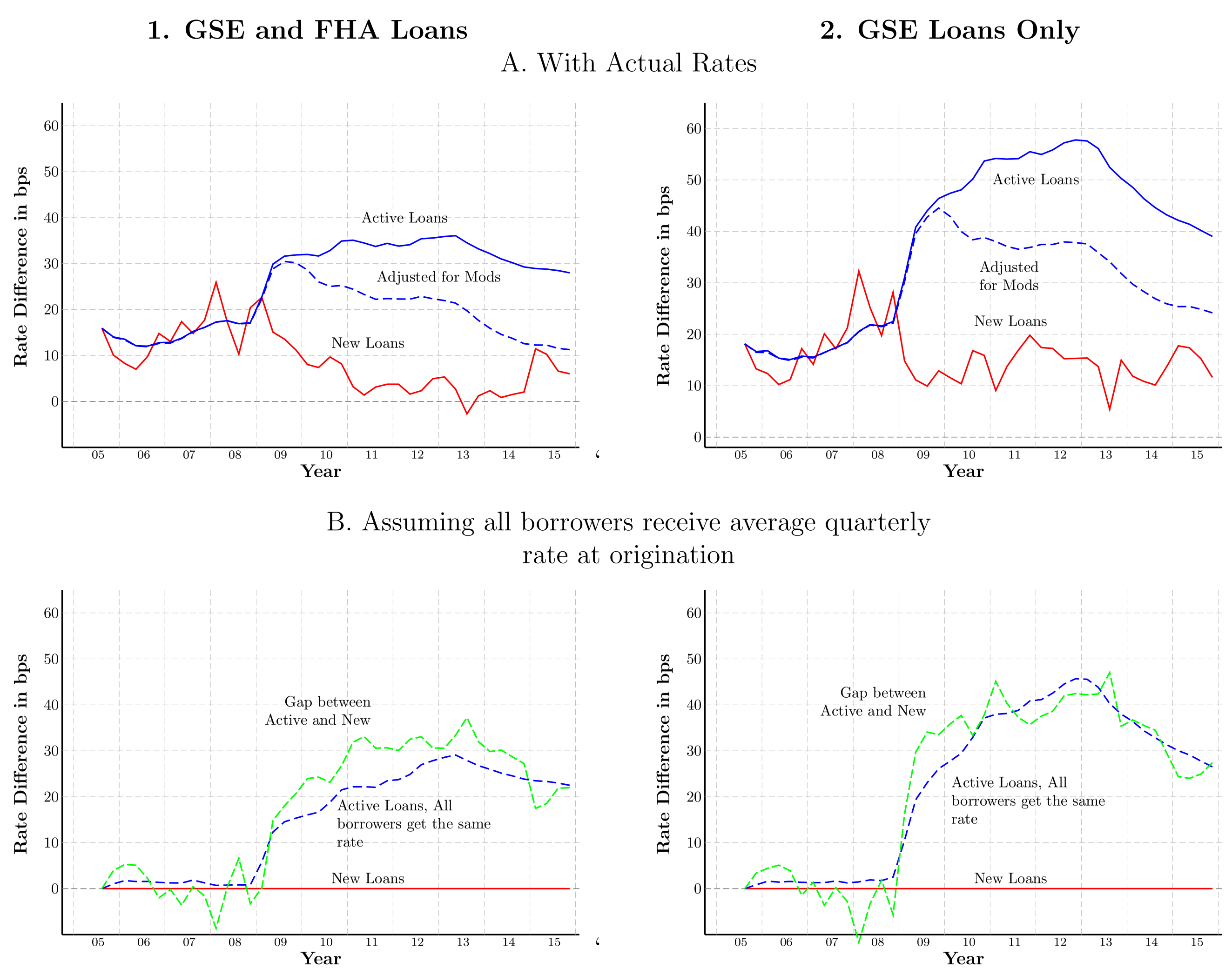
Figure 3.5: Event study of the announcement of first quantitative easing (QE1) on November 25, 2008.



groups in prepayment behavior drives large disparities in the *stock* of mortgage rates across racial/ethnic groups—the difference in rates associated with outstanding mortgages. While there are certainly good reasons to focus on the *flow* of rates, as we will show, the disparities in the stock of rates are significantly larger than the flow differences. Furthermore, we will show that monetary policy appears to have driven disparities in the stock of rates while having little impact on flow disparities.

The top panel of Figure 3.6 displays the difference in the flow of average mortgage rates (solid red line) for Black and white borrowers during our sample period and the difference in the stock of average rates (solid blue line). The left panel pools together FHA and GSE loans, while the right panel focuses on only GSE mortgages. These graphs are very similar to Figure 3.1, with the only difference being that they are constructed using our estimation sample of loans originated during the 2005–2015 period.

Figure 3.6: Gap between interest rates for Black and white borrowers for mortgages originated from 2005–2015



The flow gap fluctuates between 10 and 25 basis points over the first few years of the pooled sample before falling to zero in 2011 and remaining below 10 basis points through the end of the sample period. In the GSE sample, the flow gap falls from just over 30 basis points in 2008 to 10 basis points in 2010 and then fluctuates between 5 and 20 basis points for the remainder of the period. In contrast to the gap in the flow of rates, the gap in the stock of mortgage rates rises substantially after 2008 in both graphs. In the pooled sample it peaks at about 35 basis points in 2013, while it climbs to almost 60 basis points in the GSE sample.

We include a third series in each panel (dotted blue line) that adjusts the gap in outstanding rates to account for loan modifications. As we discussed above, HAMP was introduced in early 2009 and provided loan modifications to many borrowers in distress. One of the common types of modifications was interest rate reductions. Our McDash data provide information on interest rate changes over time, which we use to adjust the gap in the stock of rates to account for modifications that reduced borrower rates.

Interestingly, modifications appear to have had a significant impact on the rate gaps. In both panels, we can see that the difference between the average outstanding rate for Black versus white borrowers is significantly reduced when we account for rate-reducing modifications. This suggests that broad-based modification programs disproportionately affected minority borrowers and helped alleviate rate disparities in the aftermath of the crisis.

In Appendix A.17 we repeat the exercise with Survey of Consumer Finances (SCF) data as a robustness check. Although the data are much more noisier due to a smaller sample size and an inability to control for the quarter of origination, we do find a similar pattern in that the rate difference by race is larger in the stock of mortgages than at origination for new mortgages.

To isolate the disparities in the stock of rates that is due only to prepayment behavior (as opposed to differences in pricing at origination) in the bottom panel of Figure 3.6, instead of using actual interest rates paid by borrowers, we assume that every mortgage origination receives that quarter's PMMS value. Thus, by construction, there are no disparities in the rate of mortgage flows for Black and white borrowers, so that the disparities in the stock of rates are driven only by the differences in prepayment propensities. The bottom panel of Figure 3.6 shows that beginning in 2009, the tendency of Black borrowers to pay higher than market rates for longer than white borrowers increases the rate gap by more than 35 basis points in the pooled sample and by almost 50 basis points in the GSE sample.

If we go back to Figure 3.1, where we have a longer time series that goes back to 2000, we can see the obvious correlation between refinance waves and the differences in the stock of rates. The gap spikes during the refinance wave in the early 2000s and then again during the 2009–2015 period when unconventional monetary policy, largely through the purchases of trillions of dollars in mortgage-backed securities (MBS), drove down mortgage rates and spurred another refinance boom.

We now look further into the role played by unconventional monetary policy in driving the large increase in the gap in outstanding mortgage rates that we see in Figure 3.6 by estimating a difference-in-differences specification that is similar to equation 3.3 above. Specifically we estimate the following regression:

$$R^M_{it} = \beta * Black_i + \eta * postQE1_t + \delta * (Black_i * postQE1_t) + \varepsilon_{it}, \quad (3.4)$$

where the dependent variable, R^M_{it} is the current mortgage interest rate paid by borrower i (which is the same as the rate at origination, since all loans in our sample are fixed rate).

Table 3.8 displays the estimation results for three windows around the announcement of QE1: one year, two years, and four years. For each window we display two different specifications. In columns (1), (3), and (5) we estimate specifications with no additional controls, while in columns (2), (4), and (6) we add a set of vintage year-quarter fixed effects. Adding vintage year-quarter fixed effects means that only loans originated in the same year-quarter identify the QE1 coefficients, and thus, it eliminates all variation due to prepayment differences.

The unconditional regression estimates are consistent with Figure 3.6. Rates paid by white borrowers drop significantly after QE1—21 basis points in the one-year window and 46 basis points in the four-year window. At the same time, rates paid by minority borrowers also decline, but by much smaller magnitudes. For the one-year window, average rates paid by black borrowers drop by 11.5 basis points after QE1 and by about 23 basis points in the four-year window. This causes the gap in outstanding rates to grow from 21 basis points in the two years before QE1 to 44 basis points in the two years after the policy.

The addition of vintage year-quarter fixed effects completely eliminates the positive post-QE1 estimates on mortgage rates for all borrowers. This confirms that it is loans originated in different periods that drive the unconditional results, which is consistent with differential refinancing behavior driving the large divergence in mortgage rates for minority and white borrowers in the period after QE1.

3.5 Conclusion

In this paper we have shown that Black and Hispanic borrowers refinance their fixed-rate mortgages at a significantly lower rate compared with white and Asian borrowers, and that expansionary monetary policy appears to have exacerbated these differences. In turn, the large differences in refinance propensities have resulted in significant disparities in the average interest rate that minority borrowers pay on the stock of outstanding mortgages compared with their white counterparts. These differences in the stock of rates are much larger in magnitude than the corresponding differences in the rates paid on newly originated loans.

To be clear, our analysis does not suggest that policies that drive down mortgages rates are harmful to minority borrowers. To the contrary, minority borrowers do benefit from lower mortgage rates. However, our analysis suggests that they benefit much less than white borrowers.

Our research leads to two important questions. First, why do Black and Hispanic borrowers refinance less frequently? In particular, why are they so much less responsive to variation in interest rates? As we have shown, observable differences across borrowers can explain about 80 percent of the difference, but a nontrivial gap remains. The remaining gap could be explained by numerous factors that are omitted from our analysis including different levels of education and/or financial literacy, differential exposure to negative income/employment shocks that may inhibit the ability to refinance into low rates and that are not reflected in updated credit scores, or even heterogeneous social networks, which have been shown to be important transmitters of information about refinancing opportunities (Maturana and Nickerson (2019)).

The second question is, what can policymakers do to reduce racial differences? The prepayable, fixed-rate mortgage plays a central role in the story. Many commentators argue that the FRM offers the best of both worlds. Essentially, the prepayment option enables the borrower to take advantage of falling rates while providing insurance against rising rates. But the value of the option to get a lower rate, in the real world, depends on both the willingness and ability of borrowers to exercise the option. The data show systematic variation across racial groups in refinancing and moving propensities, and thus, in a sense, the value of the option.

How could a policymaker enable Black and Hispanic borrowers to exploit rate reductions more effectively? One way would be to expand the use of adjustable-rate mortgages (ARMs). The United States is almost unique in its reliance on FRMs. In many countries, the mortgage ecosystem is largely populated with ARMs, and those countries enjoy high home-ownership rates and have foreclosure problems that are no worse than in the United States. Another would be to encourage the mortgage industry to develop products that combine the benefits of FRMs and ARMs. For example, for many years, market participants have discussed “ratchet” mortgages, which adjust down but not up. These alternative mortgage contract designs may lead to a more equitable distributional impact of monetary policy. Finally, complementary, race-neutral policies that make it easier and less costly to refinance such as streamlined refinancing programs could also be effective in closing these rate disparities.

Table 3.3: Baseline Prepayment due to Refinance Results

Dependent Variable: Prepay Refinance (d)	GSE Loans						FHA Loans		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black (d)	-0.746***	-0.376***	-0.308***	-0.237***	-0.232***	-0.142***	-0.600***	-0.350***	-0.216***
	(0.086)	(0.048)	(0.036)	(0.032)	(0.032)	(0.026)	(0.053)	(0.030)	(0.029)
Hispanic (d)	-0.687***	-0.447***	-0.448***	-0.409***	-0.401***	-0.266***	-0.401***	-0.372***	-0.385***
	(0.118)	(0.064)	(0.060)	(0.055)	(0.060)	(0.040)	(0.076)	(0.040)	(0.045)
Asian (d)	0.436***	0.261***	0.277***	0.265***	0.309***	0.221***	0.417***	-0.011	0.088
	(0.143)	(0.092)	(0.097)	(0.097)	(0.100)	(0.068)	(0.088)	(0.059)	(0.072)
Equifax Risk Score		0.382***						0.445***	
		(0.068)						(0.060)	
LTV Origination		-0.011***						-0.009***	
		(0.002)						(0.002)	
Loan Amount		0.566***						0.794***	
		(0.058)						(0.053)	
LTV Change		-0.002	-0.045***	-0.042***	-0.022***	-0.021***		-0.013***	-0.030***
		(0.004)	(0.004)	(0.003)	(0.004)	(0.004)		(0.003)	(0.002)
Refinance (d)		-0.203***	-0.240***	-0.210***	-0.213***	-0.222***		-0.114***	-0.137***
		(0.060)	(0.050)	(0.048)	(0.048)	(0.049)		(0.033)	(0.040)
Female (d)			-0.063***	-0.062***	-0.062***	-0.084***			-0.077***
			(0.012)	(0.012)	(0.012)	(0.013)			(0.016)
Call Option			0.232***	0.237***					0.139***
			(0.016)	(0.016)					(0.014)
Call Option V2					0.232***	0.234***			
					(0.017)	(0.017)			
SATO			-1.492***	-1.444***	-1.310***	-1.293***			-0.196*
			(0.119)	(0.113)	(0.104)	(0.107)			(0.115)
Risk Score Change				0.769***	0.348***	0.336***			0.835***
				(0.084)	(0.067)	(0.067)			(0.084)
Loan Age	X	X	X	X	X	X	X	X	X
Underwriting Vars		X	X	X	X	X		X	X
HMDA Vars			X	X	X	X			X
Vintage Year-Qtr FE	X	X	X	X	X	X	X	X	X
State FE		X	X	X	X			X	
ZIP Code FE						X			X
# Observations	15,460,588	12,572,069	12,114,409	12,032,408	10,960,104	10,960,090	6,184,502	4,563,659	3,953,353
# Loans	792,823	652,106	629,224	629,154	600,792	600,792	291,587	221,036	192,645
R^2	0.007	0.012	0.018	0.019	0.020	0.023	0.004	0.011	0.017

Table 3.4: Baseline Prepayment due to Sale Results

Dependent Variable: Prepay Sale (d)	GSE Loans				FHA Loans			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black (d)	-0.524***	-0.413***	-0.397***	-0.373***	-0.644***	-0.536***	-0.431***	-0.413***
	(0.019)	(0.018)	(0.019)	(0.022)	(0.030)	(0.036)	(0.035)	(0.032)
Hispanic (d)	-0.430***	-0.333***	-0.330***	-0.280***	-0.515***	-0.550***	-0.514***	-0.472***
	(0.028)	(0.019)	(0.022)	(0.020)	(0.029)	(0.036)	(0.035)	(0.035)
Asian (d)	-0.185***	-0.193***	-0.214***	-0.216***	-0.233***	-0.347***	-0.327***	-0.345***
	(0.031)	(0.026)	(0.030)	(0.032)	(0.041)	(0.033)	(0.037)	(0.036)
Equifax Risk Score		0.034**				0.125***		
		(0.016)				(0.011)		
LTV Origination		-0.001				-0.002**		
		(0.001)				(0.001)		
Loan Amount		0.136***				0.195***		
		(0.015)				(0.013)		
LTV Change		-0.015***	-0.019***	-0.020***		-0.021***	-0.022***	-0.023***
		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.002)
Refinance (d)		-0.199***	-0.144***	-0.124***		-0.228***	-0.115***	-0.096***
		(0.019)	(0.019)	(0.018)		(0.023)	(0.020)	(0.021)
Female (d)			0.027***	0.020**			0.025**	0.010
			(0.008)	(0.009)			(0.011)	(0.010)
Call Option			0.035***	0.036***			0.004	0.004
			(0.002)	(0.002)			(0.003)	(0.003)
SATO			-0.127***	-0.112***			0.112**	0.165***
			(0.028)	(0.029)			(0.048)	(0.051)
Risk Score Change			-0.065**	-0.079***			0.278***	0.267***
			(0.028)	(0.028)			(0.014)	(0.014)
Loan Age	X	X	X	X	X	X	X	X
Underwriting Vars		X	X	X		X	X	X
HMDA Vars			X	X			X	X
Vintage Year-Qtr FE	X	X	X	X	X	X	X	X
State FE		X	X			X	X	
ZIP Code FE				X				X
# Observations	15,460,588	12,572,069	10,960,104	10,960,090	6,184,502	4,563,659	3,953,353	3,953,352
# Loans	792,823	652,106	600,792	600,792	291,587	221,036	192,645	192,645
R^2	0.002	0.003	0.004	0.006	0.003	0.004	0.006	0.012

Table 3.5: Baseline Default Results

Dependent Variable: Default (d)	GSE Loans				FHA Loans		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Black (d)	0.443***	0.270***	0.147***	0.105***	0.733***	0.447***	0.401***
	(0.076)	(0.050)	(0.024)	(0.017)	(0.057)	(0.036)	(0.031)
Hispanic (d)	0.422***	0.268***	0.058***	0.054***	0.165***	0.148***	0.077*
	(0.097)	(0.063)	(0.015)	(0.015)	(0.044)	(0.044)	(0.043)
Asian (d)	0.026	0.026**	0.021*	0.010	-0.125***	-0.065***	-0.056**
	(0.018)	(0.012)	(0.010)	(0.009)	(0.023)	(0.024)	(0.026)
Equifax Risk Score		-0.452***				-0.903***	
		(0.062)				(0.076)	
LTV Origination		0.009***				0.014***	
		(0.001)				(0.001)	
Loan Amount		0.048***				0.124***	
		(0.012)				(0.026)	
LTV Change		0.031***	0.022***	0.024***		0.029***	0.031***
		(0.003)	(0.003)	(0.003)		(0.003)	(0.004)
Refinance (d)		0.128***	0.070***	0.066***		0.243***	0.139***
		(0.018)	(0.012)	(0.012)		(0.040)	(0.032)
Female (d)			-0.013***	-0.013***			-0.028**
			(0.004)	(0.004)			(0.012)
Call Option			-0.028***	-0.028***			-0.000
			(0.005)	(0.005)			(0.003)
SATO			0.551***	0.553***			0.446***
			(0.086)	(0.085)			(0.061)
Loan Age	X	X	X	X	X	X	X
Underwriting Vars		X	X	X		X	X
HMDA Vars			X	X			X
Vintage Year-Qtr FE	X	X	X	X	X	X	X
State FE		X	X			X	
ZIP Code FE				X			X
# Observations	14,883,532	12,125,625	10,684,132	10,684,116	5,484,924	4,057,993	3,524,527
# Loans	792,823	652,106	600,795	600,795	291,587	221,036	192,655
R^2	0.006	0.011	0.008	0.012	0.006	0.011	0.012

Table 3.6: Prepayment due to Refinance with Interaction Effects

Dependent Variable: Prepay Refinance (d)	GSE Loans			FHA Loans		
	(1)	(2)	(3)	(4)	(5)	(6)
Black (d)	-0.170***	0.339***	0.337***	-0.160***	0.223**	0.216**
	(0.027)	(0.075)	(0.076)	(0.026)	(0.098)	(0.098)
Hispanic (d)	-0.280***	0.172***	0.166**	-0.297***	0.041	0.050
	(0.040)	(0.062)	(0.064)	(0.039)	(0.076)	(0.077)
Call Option	0.231***	0.239***	0.239***	0.139***	0.151***	0.151***
	(0.017)	(0.018)	(0.018)	(0.015)	(0.016)	(0.016)
Risk Score Change	0.329***	0.330***	0.315***	0.828***	0.825***	0.878***
	(0.067)	(0.067)	(0.070)	(0.082)	(0.082)	(0.086)
Black *Call Option		-0.083***	-0.083***		-0.051***	-0.053***
		(0.007)	(0.007)		(0.007)	(0.007)
Hispanic * Call Option		-0.078***	-0.079***		-0.049***	-0.051***
		(0.007)	(0.007)		(0.006)	(0.006)
Black * Risk Score Change			0.109**			-0.202***
			(0.043)			(0.025)
Hispanic * Risk Score Change			0.089**			-0.167***
			(0.037)			(0.037)
Loan Age	X	X	X	X	X	X
Underwriting Vars	X	X	X	X	X	X
HMDA Vars	X	X	X	X	X	X
Vintage Year-Qtr FE	X	X	X	X	X	X
ZIP Code FE	X	X	X	X	X	X
# Observations	10,354,221	10,354,221	10,354,221	3,856,796	3,856,796	3,856,796
# Loans	563,995	563,995	563,995	187,528	187,528	187,528
R^2	0.023	0.023	0.023	0.023	0.024	0.024

Table 3.7: Effect of QE1 on Differences in Refinance Propensities

Dependent Variable: Prepay Refinance (d)						
Window:	1-Year			2-Year		
	(1)	(2)	(3)	(4)	(5)	(6)
postQE1 (d)	3.221***	3.004***	0.771*	1.762***	2.361***	0.970***
	(0.496)	(0.219)	(0.414)	(0.438)	(0.222)	(0.283)
Black * postQE1	-2.236***	-2.671***	-1.798***	-1.241***	-1.593***	-1.037***
	(0.417)	(0.271)	(0.205)	(0.338)	(0.187)	(0.142)
Hispanic * postQE1	-2.288***	-2.624***	-2.119***	-1.263***	-1.515***	-1.191***
	(0.413)	(0.275)	(0.239)	(0.346)	(0.216)	(0.191)
Black (d)	-0.112**	1.056***	0.595***	-0.268***	0.608***	0.305***
	(0.040)	(0.113)	(0.090)	(0.092)	(0.105)	(0.094)
Hispanic (d)	-0.311***	0.914***	0.641***	-0.480***	0.423***	0.241**
	(0.051)	(0.093)	(0.082)	(0.103)	(0.094)	(0.089)
$600 \leq$ Equifax Risk Score < 740 (d)		0.511***	-0.185**		0.546***	0.074
		(0.062)	(0.072)		(0.060)	(0.059)
Equifax Risk Score ≥ 740 (d)		1.643***	-0.243**		1.386***	0.186
		(0.122)	(0.094)		(0.128)	(0.109)
postQE1 * ($600 \leq$ Equifax Risk Score < 740)			1.484***			1.033***
			(0.162)			(0.119)
postQE1 * (Equifax Risk Score ≥ 740)			3.754***			2.377***
			(0.331)			(0.227)
Constant	0.615***	-3.886***	-3.312**	1.095***	0.256	0.912**
	(0.066)	(1.275)	(1.271)	(0.177)	(0.261)	(0.343)
Loan Age		X	X		X	X
Underwriting Vars		X	X		X	X
HMDA Vars		X	X		X	X
Vintage Year-Qtr FE		X	X		X	X
ZIP Code FE		X	X		X	X
# Observations	1,066,525	782,523	782,523	2,129,912	1,563,213	1,563,213
R^2	0.012	0.055	0.058	0.004	0.038	0.039

Table 3.8: Effect of QE1 on Differences in the Stock of Outstanding Mortgage Rates

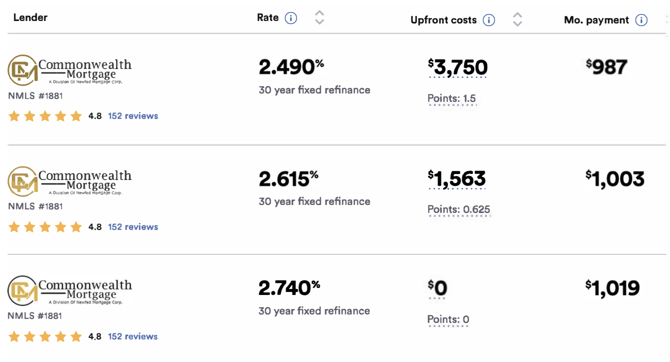
Dependent Variable: Mortgage Rate						
Window:	1-Year		2-Years		4-Years	
	(1)	(2)	(3)	(4)	(5)	(6)
Black (d)	0.224***	0.189***	0.222***	0.187***	0.210***	0.176***
	(0.014)	(0.012)	(0.014)	(0.012)	(0.014)	(0.012)
Hispanic (d)	0.135***	0.112***	0.132***	0.109***	0.123***	0.101***
	(0.019)	(0.016)	(0.019)	(0.016)	(0.019)	(0.016)
postQE1 (d)	-0.209***	-0.006***	-0.313***	-0.004***	-0.463***	-0.008***
	(0.004)	(0.000)	(0.005)	(0.000)	(0.007)	(0.001)
Black * postQE1	0.115***	-0.005***	0.162***	-0.009***	0.226***	-0.007*
	(0.004)	(0.002)	(0.005)	(0.002)	(0.007)	(0.003)
Hispanic * postQE1	0.111***	0.004**	0.157***	0.003	0.214***	0.006*
	(0.004)	(0.002)	(0.006)	(0.002)	(0.010)	(0.004)
Constant	6.239***	6.142***	6.245***	6.090***	6.252***	6.000***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Vintage Year-Qtr FE		X		X		X
# Observations	1,066,525	1,066,525	2,129,912	2,129,912	4,085,825	4,085,825
R^2	0.044	0.528	0.070	0.588	0.114	0.660

A

Supplemental Material

A.1 Upfront closing cost choices example

Figure A.1: Rate and upfront closing costs trade-offs facing mortgage borrowers



Lender	Rate	Upfront costs	Mo. payment
Commonwealth Mortgage, A Division Of Newfed Mortgage Corp. NMLS #1881, 4.8, 152 reviews	2.490% 30 year fixed refinance	$3,750 Points: 1.5	$987
Commonwealth Mortgage, A Division Of Newfed Mortgage Corp. NMLS #1881, 4.8, 152 reviews	2.615% 30 year fixed refinance	$1,563 Points: 0.625	$1,003
Commonwealth Mortgage, A Division Of Newfed Mortgage Corp. NMLS #1881, 4.8, 152 reviews	2.740% 30 year fixed refinance	$0 Points: 0	$1,019

A.2 Data Construction and Summary Statistics

A.2.1 Optimal Blue-HMDA sample

I constructed the Optimal Blue-HMDA sample by merging the Optimal Blue rate locks from 2018–2019 with the public HMDA data. Because Optimal Blue contains a lender identifier number but no lender names, the merge proceeds in two steps: (1) an initial match based on loan characteristics, and (2) a second filtering based on a correspondence between the lender identifier in Optimal Blue and an anonymized version of HMDA lender IDs implied by the first step.

The initial match was made using loan amount, rate, year, loan type, loan purpose, loan term, ZIP code (with all ZIP codes corresponding to an HMDA census tract included), and up to a 5% difference in LTV with all matches kept in the data set. Then, for the second step I impose the requirement that the lender identifier in Optimal Blue is matched to an anonymized version of

HMDA lender ID at least 10% of the time.[1] Overall, this two-step procedure uniquely matches 1,186,906 out of 2,318,940 locks for 30-year, conforming fixed-rate mortgages, implying a match rate of 51%. The match rate is comparable to a 66% "lock pull-through rate," which is the percent of rate locks that turn into originated loans, that I understand to be reasonable based on industry sources.

In terms of variable definitions, I construct a Black dummy equal to one if the mortgage has a HMDA-derived race variable of "Black or African American." The Hispanic dummy is equal to one if the mortgage has a HMDA derived ethnicity variable of "Hispanic or Latino." The Single Male and Single Female dummies are inferred from the HMDA-derived gender. Summary statistics for these samples are shown in the table below.

---

[1] The 10% requirement was set purposefully low to include cases where the Optimal Blue lender ID may not correspond to a HMDA reporter for example in the case of correspondent lending. It is sufficient to reduce the percent of matches that are non-unique from 49.6% to 3.9%.

Table A.1: Summary statistics for the 2018–2019 Optimal Blue-HMDA sample

	Mean	Std. Dev.	25th Pctile	Median	75th Pctile	N
Loan amount ($'000s)	256.7	117.8	166.3	240.0	333.8	1186906
Origination cost ($)	1818.7	1643.6	995.0	1370.0	2155.8	1154700
Credit Score	747.9	44.5	717	755	784	1186906
LTV (%)	80.3	15.0	75	80	95	1186906
DTI (%)	35.0	9.7	28.2	36.2	42.9	1171555
Interest rate (%)	4.54	0.58	4	4.625	4.875	1186906
First-time Home Buyer (d)	0.307	0.461	0	0	1	1186551
Black (d)	0.041	0.199	0	0	0	1071773
Hispanic (d)	0.096	0.294	0	0	0	1063925
Single Female (d)	0.252	0.434	0	0	1	1140692
Single Male (d)	0.330	0.470	0	0	1	1140692

Notes: This table reports summary statistics from the 2018–2019 Optimal Blue-HMDA merged sample. Loan amount is expressed in thousands of dollars, origination costs are expressed in dollars, credit score is the borrower's Optimal Blue credit score at origination, and LTV, interest rate are expressed in percentage points. The label (d) denotes dummy variables.

A.2.2 Optimal Blue-HMDA-CRISM sample

I also construct a merge between Optimal Blue, HMDA, and CRISM data sets for mortgages originated between 2013–2019, with loan performance until April 2021. The CRISM data set is an anonymous credit file match from Equifax consumer credit database to Black Knight's Mcdash loan-level Mortgage Data set. My Optimal Blue-HMDA-CRISM sample was constructed by joining together three merges, (i) the 2018–2019 Optimal Blue and HMDA merge described in Section A.2.1, (ii) a 2013–2017 Optimal Blue and HMDA merge, and (iii) the 2013–2019 Optimal Blue and CRISM merge.

Similar to the 2018-2019 Optimal Blue and HMDA merge, the 2013–2017 Optimal Blue and HMDA merge was also conducted in two steps, with an initial step based on loan characteristics, and a second step based on a correspondence between the Optimal Blue lender ID and an anonymized HMDA lender ID. A separate merge was conducted because the data fields in 2013–2017 HMDA are different than those in 2018–2019 HMDA: the interest rate, loan term, and LTV fields were not available, while loan amount was given in finer detail.

The first step for the 2013–2017 Optimal Blue to HMDA match was made using loan amount, year, loan type, loan purpose, occupancy, ZIP code (with all ZIP codes corresponding to an HMDA census tract included) with all matches kept in the data set. Then, for the second step I impose the requirement that the lender identifier in Optimal Blue is matched to an HMDA respondent ID at least 10% of the time.[2] Overall, this two-step procedure uniquely matches 1,382,057 out of

[2] The 10% requirement was set purposefully low to include cases where the Optimal Blue lender

2,563,550 locks for 30-year, conforming fixed-rate mortgages, implying a match rate between locks to originated mortgages of 54%. The match rate is again comparable to a 66% "lock pull-through rate," which I understand to be reasonable based on industry sources.

The 2013–2019 Optimal Blue to CRISM match was made in one step. The variables used for matching are the loan amount, ZIP code, month of origination (which I require to lie within the date of the lock and the date of the lock plus the lock term), loan type, loan term, loan purpose, Equifax Risk Score (within 20 points of the Optimal Blue credit score), LTV (within 5%), and the rate. The more detailed loan-level information enabled the match to proceed despite not having lender information. Overall, I uniquely matched 617,058 out of 5,269,107 locks for 30-year, conforming fixed-rate mortgages, implying a match rate between locks to originated mortgages in the CRISM data set of 12%. The lower match rate is reasonable because neither the CRISM data nor the Optimal Blue data covers all US mortgage originations, so the overlap between the two must be smaller than the overlap between Optimal Blue and HMDA as the HMDA does provide essentially complete coverage of all US mortgage originations.

Combining the three merges, I get an Optimal Blue-HMDA-CRISM sample with 360,291 loans. In terms of variable definitions, I construct a Black dummy equal to one if the mortgage has a 2018–2019 HMDA derived race variable of "Black or African American." The Hispanic dummy is equal to one if the mortgage has a HMDA derived ethnicity variable of "Hispanic or Latino." In the case of 2013–2017 HMDA, these dummies are defined using the algorithm of Bhutta and Can-

ID may not correspond to an HMDA reporter for example in the case of correspondent lending. It is sufficient to reduce the percent of matches that are non-unique from 75.2% to 11.8%.

ner (2013). The Single Male and Single Female dummies are inferred from the 2018–2019 HMDA derived gender or the applicant gender when no co-applicant is present in the case of 2013–2017 HMDA. Finally, the Credit Card Revolver dummy is set equal to 1 if the primary borrower on the mortgage has a credit card balance of greater than or equal to $10,000 at the time of origination while also having a credit card utilization of greater than 40%.

Summary statistics on this sample are as follows:

Table A.2: Summary statistics for the Optimal Blue-HMDA-CRISM sample

	Mean	Std. Dev.	25th Pctile	Median	75th Pctile	*N*
Loan amount ($'000s)	252.7	114.7	166.2	234.0	325.5	360291
Credit score	748.5	46.0	717.0	757.0	786.0	360291
LTV (%)	79.8	15.0	75.0	80.0	92.0	360291
DTI (%)	34.4	9.4	27.8	35.6	42.0	360291
First-time home buyer (d)	0.200	0.400	0	0	1	351955
Black (d)	0.030	0.171	0	0	0	340665
Hispanic (d)	0.076	0.265	0	0	0	340665
Single Female (d)	0.249	0.433	0	0	1	346222
Single Male (d)	0.322	0.467	0	0	1	346222
Credit Card Revolver (d)	0.110	0.313	0	0	1	346222

Notes: This table reports summary statistics from the Optimal Blue-HMDA-CRISM merged sample. Loan amount is expressed in thousands of dollars, origination costs are expressed in dollars, credit score is the borrower's Optimal Blue credit score at origination, and LTV, interest rate are expressed in percentage points. The label (d) denotes dummy variables. CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data.

A.2.3 The LoanSifter data

The LoanSifter data contains information about rate and upfront closing cost (i.e., points) trade-offs in rate sheets, which are prices that loan originators and mortgage brokers can offer to clients in locking the loan. Because these are actual available prices within a lender, they allow me to observe the rate and point menus that borrowers face. The sample period runs from September 9, 2009 to December 31, 2014 and consists of rate sheets from a sample of lenders from 50 metropolitan areas. Rate sheets observations are at the lender-day level, and in rare cases where a lender issues more than one rate sheet on a given day the observations with the best prices are kept. Linear interpolation was used to estimate the rate at various levels of points, following Fuster, Lo, and Willen (2017). To compare the rate and points menus in the lender rate sheets to the MBS TBA prices, I focus on rate sheets for conforming, 30-year, fixed-rate mortgages with a loan-to-value ratio of 80% and a loan amount of greater than or equal to $300k.

Summary statistics for this data are shown in Table A.3.

Table A.3: Summary statistics for the LoanSifter data

Year	No. of Lenders	Rate at 0 points	Rate at 2 points	N lender-days obs
2009	93	5.01	5.42	3923
2010	93	4.70	5.10	16025
2011	83	4.46	4.82	16589
2012	86	3.67	4.07	18105
2013	126	4.07	4.42	19993
2014	103	4.21	4.52	19446

Note: This table contains information on the number of distinct lenders, mean rate at 0 points, mean rate at 2 points, and number of distinct lender-day observations by year. The data set comes from LoanSifter. The interest rates at 0 points and at 2 points are estimated through linear interpolation for lenders that do not offer mortgages at exactly those points.

A.3 When are closing costs added to the rate?

This paper focuses on the cross-subsidization of mortgage closing costs to the extent that they are added to the rate of the mortgage. I refer to mortgages with closing costs "added to the rate" as mortgages with a high enough interest rate c such that secondary marketing income(c) in Equation (1.1) is positive.[3] While intuitive, this definition is most sensible in a world in which lenders pass through their secondary marketing income as lower upfront closing costs to borrowers, for example in a model with a perfectly competitive supply side. Otherwise, the positive secondary marketing in-

[3] This turns out to be true for most mortgages, as I show in Section 1.4.1.

come may reflect not only closing costs added to the rate but also an additional cost that only some borrowers pay. Empirically, my analysis of US mortgage pricing finds this pass-through to be nearly complete which makes my definition sensible.

To assess this pass-through, I examine how the secondary marketing income-interest rate trade-off matches the retail interest rate and upfront closing costs trade-off in the cross-section, with results in Figure A.2. I use data LoanSifter matched with MBS TBA pricing data from 2009Q3 to 2014. Following the methodology of Fuster, Lo, and Willen (2017), which estimates the price of intermediation as the premium of the mortgage over par on the secondary market, I estimate (i) the secondary marketing revenue generated by lenders in the secondary market as implied by MBS TBA prices, and (ii) the sum of the revenue generated by lenders in the secondary market and the upfront closing costs they charge in the form of points, for borrowers with a $300k conforming mortgage, 700 LoanSifter credit score, 80% LTV, and 30% DTI.

Then, with the interest rate spread to the Freddie Mac Primary Mortgage Market Survey (PMMS) rate[4] rounded to the nearest 1/8th $\tilde{c}$, I run a linear regressions of the form:

$$\varphi_{ijt} = \sum_{l=1}^{N} \gamma_l \mathbb{1}(c = c_l) + \xi_{jt} + \varepsilon_{ijt}, \tag{A.1}$$

where c_l are the categorical variables of interest rate spread rounded to the nearest 1/8th, ξ_{jt} are lender-day fixed effects, and ε_{ijt} is the error term. φ_{ijt} is either the secondary marketing revenue

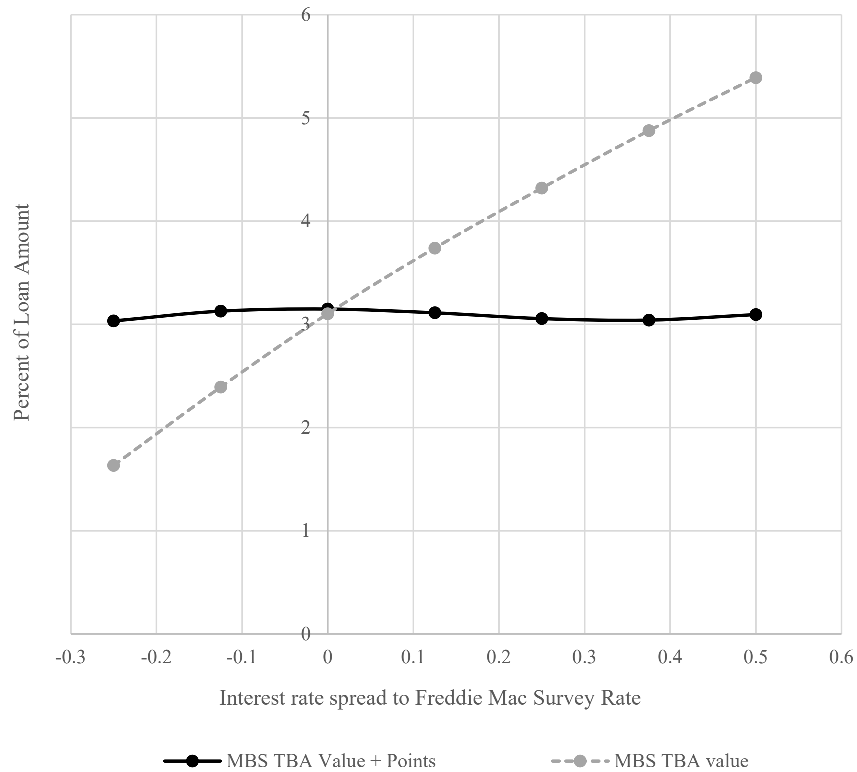
[4]The Freddie Mac Primary Mortgage Market Survey rate is obtained from https://fred.stlouisfed.org/series/MORTGAGE30US.

generated the lender or sum of the revenue generated by lenders in the secondary market and the upfront closing costs, both expressed as a percentage of the loan amount.

Results are presented in Figure A.2, which shows that mortgages that are originated at a higher spread to the Freddie Mac Survey rate tend to command higher valuations in the secondary market but generate almost exactly the same lender total income. This suggests that higher secondary marketing income is almost entirely passed through to consumers in the form of lower upfront lender fees/points.[5] Given the near complete pass-through of secondary marketing income to primary market upfront closing costs on average, it is economically meaningful to say that mortgages with positive secondary marketing income have a part of their upfront closing costs "added to the rate" which is then subject to cross-subsidization.

[5] The same patterns also exist in the time series, as I illustrate in Appendix Figure A.3. In Figure A.3, there is some evidence that in more recent years the interest rate is slightly lower on low upfront closing cost mortgages than what would be implied by secondary marketing income, perhaps suggesting a role for markups. I abstract from markups that vary by points in this paper as the magnitude of the cross-subsidization I study is significantly larger than the differences shown in Figure A.3.

Figure A.2: Secondary marketing income and total lender revenues



In addition to cross-section, I also examine the relationship between the rate and upfront closing cost trade-off in the time series in Figure A.3. Using the LoanSifter data, I estimate the rate increase from paying 1 less point (i.e., 1% of the loan amount less) in upfront closing costs as the interest rate increase from going from a mortgage with 1 point in upfront closing costs to a mortgage with 0 points within each lender rate sheet. To get the corresponding exchange rate in the MBS TBA data, I take the mortgage rate at 0 points (net of the g-fees or the price of GSE guarantee) and compute the increase in rate that would imply a 1% increase in the MBS TBA value of the mortgage, with interpolated values for coupon rates in between eighths. I then take the mean of the exchange rate implied by the LoanSifter data and the MBS TBA data by month, with results plotted in Figure A.3.

Figure A.3: The interest rate increase from paying 1 less point in upfront closing cost over time, lender ratesheets (green) versus MBS TBA implied (red)

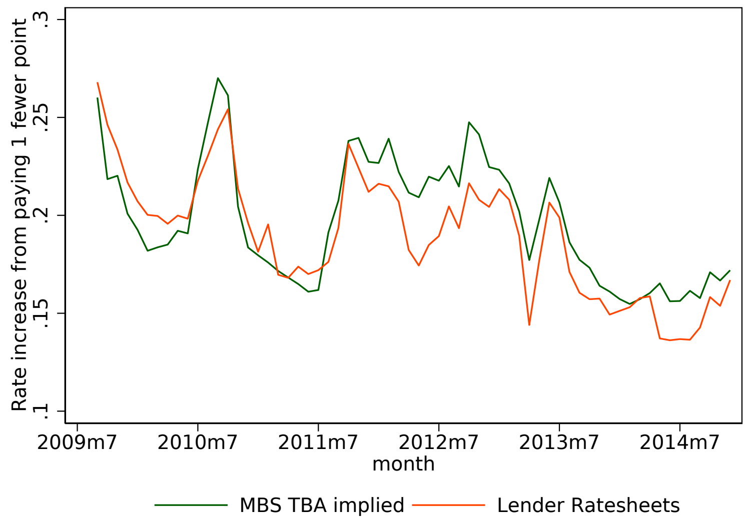
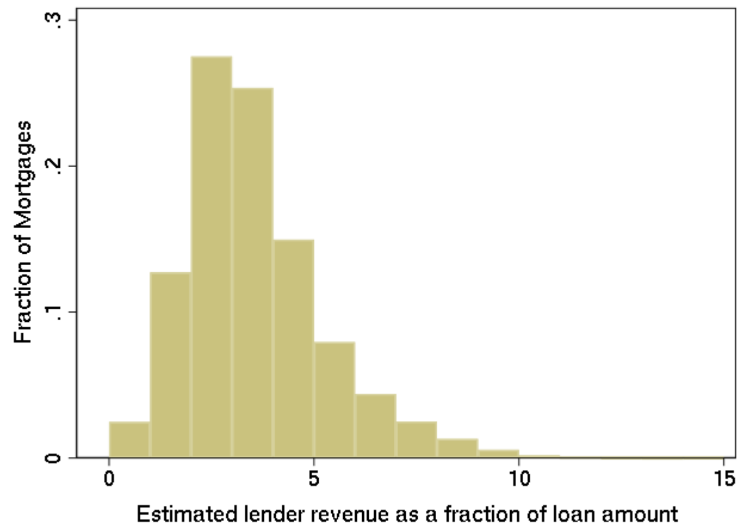


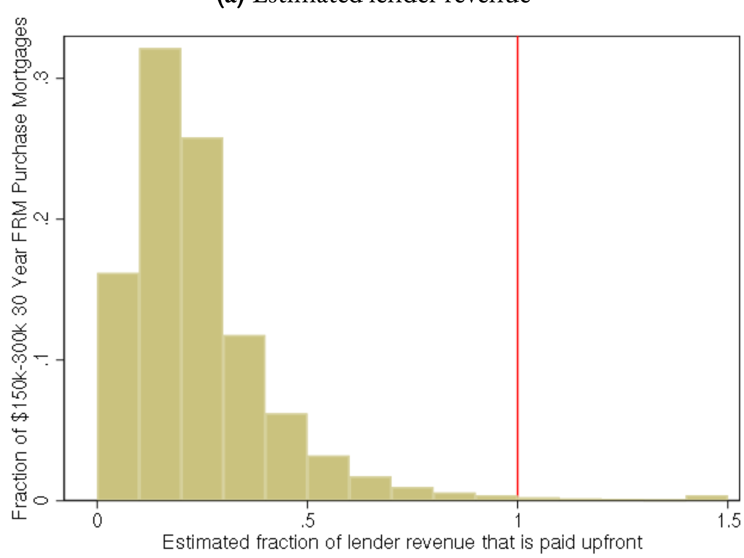
Figure A.3 shows that the exchange rate implied by the the LoanSifter data and the MBS TBA data are fairly close to each other, with the MBS TBA implied exchange rate being slightly larger near the end of the sample. This is consistent with near complete pass-through of secondary marketing revenue to upfront closing costs, with a small discount to lower closing cost mortgages in the retail market as compared to the secondary market near the end of the sample.

A.4 Robustness check on the proportion of closing costs paid upfront

Figure A.4: Lender revenue and percent paid as upfront closing costs, net of mortgage servicing revenue



(a) Estimated lender revenue



(b) Fraction of lender revenue paid upfront

A.5 Model details

A.5.1 Exogenous states

The risk-free rate follows the Cox, Ingersoll, and Ross (1985) model which has a natural zero lower bound:

$$dr_{1t} = a(b - r_{1t})dt + \sigma\sqrt{r_{1t}}dW_t. \tag{A.2}$$

I estimate the evolution of exogenous states in the model via maximum likelihood[6] using the three-month Treasury bill data from January 1987 to January 2021.[7] The results for the risk-free rate are as follows:

Table A.4: Estimation of the CIR model of interest rates

Parameter	Estimate	Standard Error
a	0.0910	0.0506
b	1.2649	0.7209
σ	0.4930	0.0175

[6] The program was based on Kladivko (2021), with some modifications to obtain standard errors.

[7] Board of Governors of the Federal Reserve System (US), 3-Month Treasury Bill Secondary Market Rate [TB3MS], retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/TB3MS.

I model the average mortgage rate $\bar{c}_t$, changes in log real house prices ΔH_t, and changes in log real personal income ΔL_t and as a vector autoregression (VAR) with r_{1t} as an exogenous dependent variable. I use two lags in the VAR, with the constraint that the matrix of coefficients on first lag is identity and on the second lag is positive only for the house price coefficient to reduce dimensionality.[8] More specifically, with $\boldsymbol{s_t} = \begin{bmatrix} \bar{c}_t \\ 100 * \Delta H_t \\ 100 * \Delta L_t \end{bmatrix}$, the VAR equation is as follows:

$$\boldsymbol{s_t} = \mu + r_{1t}\beta_{r_{1t}} + \Phi_1 \boldsymbol{s'_{t-1}} + \Phi_2 \Delta H_{t-1} + \boldsymbol{e_t}, \qquad \text{(A.3)}$$

where $\boldsymbol{e_t} \sim N(0, \hat{\Sigma}_s)$ and $\mu, \beta_{r_{1t}}, \Phi_2$ are the coefficients to be estimated. In terms of the state variables, data on $\bar{c}_t$ is obtained as the Primary Mortgage Market Survey (PMMS) rate,[9] H_t is obtained from the Case-Shiller National House Price Index,[10] and L_t is obtained from the US Personal Income[11] divided by the US population.[12] Furthermore, H_t and L_t are converted to real terms using

[8]The second lag on the house price variable is added to capture momentum and mean reversion as in Glaeser and Nathanson (2017).

[9]Freddie Mac, 30-Year Fixed Rate Mortgage Average in the United States [MORTGAGE30US], retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/MORTGAGE30US.

[10]S&P Dow Jones Indices LLC, S&P/Case-Shiller U.S. National Home Price Index [CSUSHPINSA], retrieved from FRED, Federal Reserve Bank of St. Louis;,https://fred.stlouisfed.org/series/CSUSHPINSA.

[11]U.S. Bureau of Economic Analysis, Personal Income [PI], retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/PI

[12]U.S. Bureau of Economic Analysis, Population [POPTHM], retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/POPTHM.

the Consumer Price Index for All Urban Consumers.[13] The results of the VAR estimation are as follows:

Table A.5: VAR estimates of state transitions

Parameter	μ	$\beta_{r_{1t}}$	Φ_1			Φ_2	$\hat{\Sigma}_s$		
$\bar{c}_t$	.093 (.051)	.024 (.010)	.972 (.012)	0	0	0	.050		
$100 * \Delta H_t$	.051 (.028)	-.008 (.007)	0	1.060 (.047)	0	-.286 (.047)	.009	.126	
$100 * \Delta L_t$	.182 (.079)	-.007 (.021)	0	0	-.232 (.053)	0	-.006	.041	1.030

The estimates from Tables A.4 and A.5 are then used to simulate the transitions of the exogenous states in my model in Section 2.3.

A.5.2 OAS

An empirical model of prepayment behavior combined with my model of interest rates is needed to estimate the OAS in Section 1.5.1. For my empirical model of prepayment, I use my panel data to estimate a logit regression of an indicator variable for borrower prepayment on the spread of the mortgage interest rate to the Freddit Mac survey rate at origination (SATO) as well as categories of the interest rate incentive defined as the current mortgage interest rate minus the Freddit Mac survey rate. To maintain comparability to the TBA market from which I derive the market exchange rate between the interest rate and upfront closing costs, I further restrict my analysis to 30 year purchase

[13] U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/CPIAUCSL.

mortgages with a balance above \$150k, FICO above 680, and LTV below 85% following Fusari et al. (2020). Results of this regression are shown in Table A.6, which is used for my model of $\hat{p}_t$ as in Equation (1.19).

Table A.6: Logit model of prepayment

	(1)	
	Prepaid	
SATO	-0.538***	(-11.22)
SATO Sq	-0.325***	(-6.72)
Rate Incentive $\geq$ 0%	0.482***	(28.16)
Rate Incentive $\geq$ 0.5%	1.061***	(68.46)
Rate Incentive $\geq$ 1%	0.768***	(46.27)
Constant	-5.281***	(-398.27)
Observations	4081108	

t statistics adjusted for clusters at the lender by county level in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Using the prepayment model from Table A.6 and the interest rate model of Section A.5.1, with the risk-free rate r_{tf} being given as the implied 10 year rate under the Cox, Ingersoll, and Ross (1985) model, I estimate a $\hat{OAS} = 0.22\%$ by minimizing the equally-weighted difference between the observed MBS TBA price for the nearest two coupons above and below the Freddie Mac survey rate - gfees - servicing fees with the implied NPV given by Equation (1.19). The MBS TBA price is inclusive of the new production pay-up for a coupon (with data from Morgan Markets). The gfee is assumed to be 0.42% and servicing fee 0.25% following Fuster, Lo, and Willen (2017).

A.5.3 Model hyper-parameter estimates

Table A.7: Estimated hyper-parameters and their standard errors

Parameter	Value	Standard Error
μ_{p^a}	-2.941	(0.244)
σ_{p^a}	0.879	(0.097)
μ^{β}	2.322	(1.034)
σ_{β}	3.950	(0.045)
ρ	0.956	(0.018)
μ_{p^m}	-2.103	(0.092)
σ_{p^m}	0.190	(0.039)
μ_{κ}	3.551	(0.051)
σ_{κ}	2.108	(0.023)
$\mu^b_{p^a}$	-0.626	(0.326)
$\mu^b_{p^m}$	-0.851	(0.253)
μ^b_{κ}	-0.132	(0.080)
$\mu^h_{p^a}$	-0.520	(0.200)
$\mu^h_{p^m}$	-0.655	(0.153)
μ^h_{κ}	0.059	(0.057)

A.5.4 Sanity check on model

Table A.8: Regression of model-implied optimal upfront closing cost choices by borrower characteristics

	(1)
	optimal_points
beta	3.173***
	(0.309)
gamma	-0.748***
	(0.0192)
pa	-2.56***
	(0.290)
log_kappa	0.0547***
	(0.00447)
pm	-5.572**
	(2.239)
savings_cat	0.0821***
	(0.00435)
_cons	-0.617***
	(0.1213)
N	24930

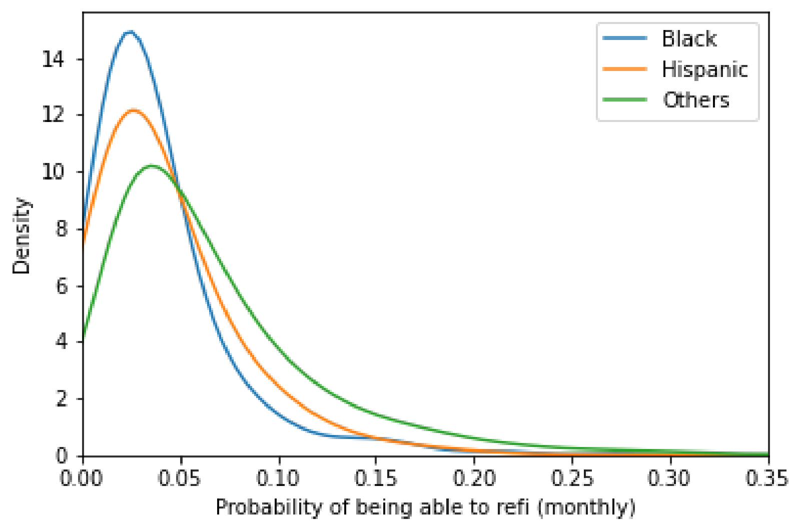
Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.6 Estimates by race

In the main text of the paper, many results were aggregated across borrower racial groups. This Appendix section presents some estimates by race.

Figure A.5: Distribution of borrower refinancing types by race

(a) Probability of being able to refi



(b) Hassle cost for refinancing

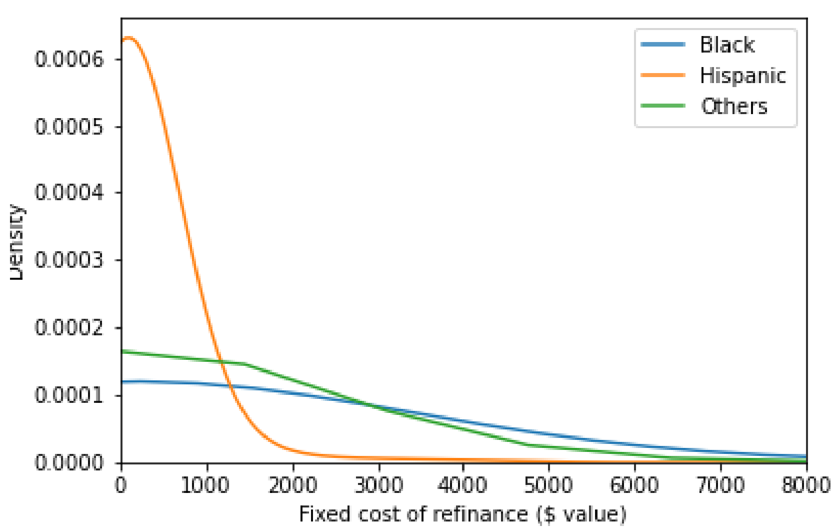


Figure A.6: Moving probability

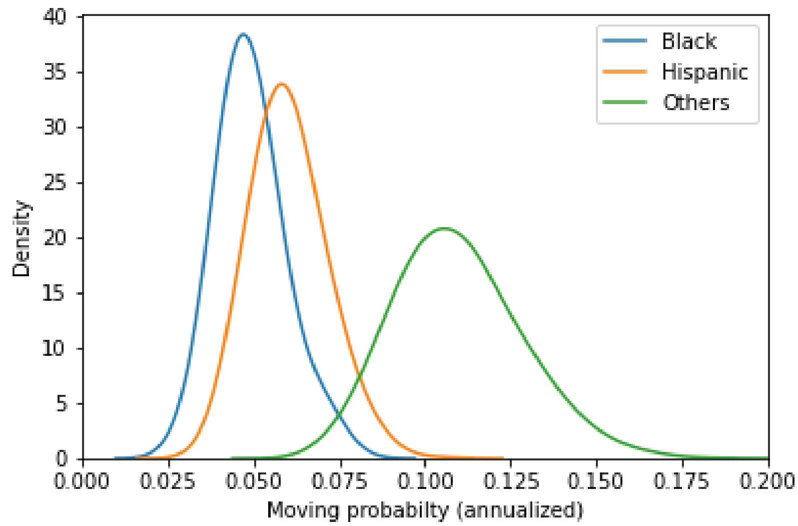


Figure A.7: Counterfactual change in utility from adding closing costs to balance of the loan, by racial group

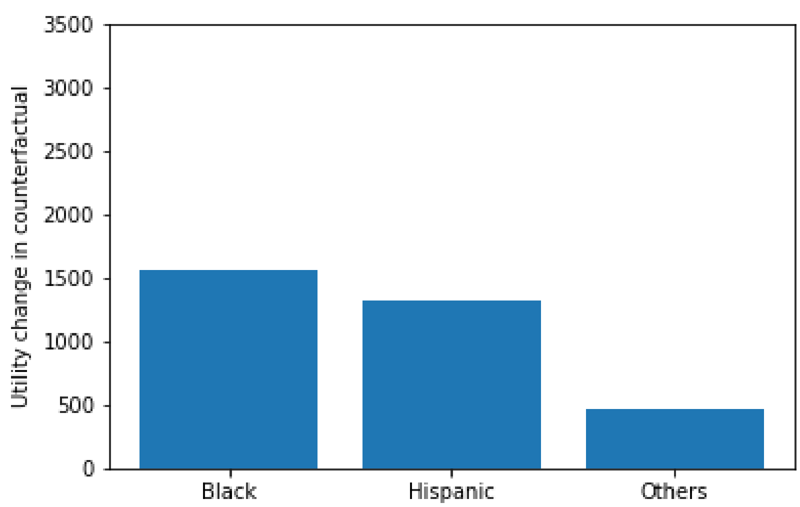
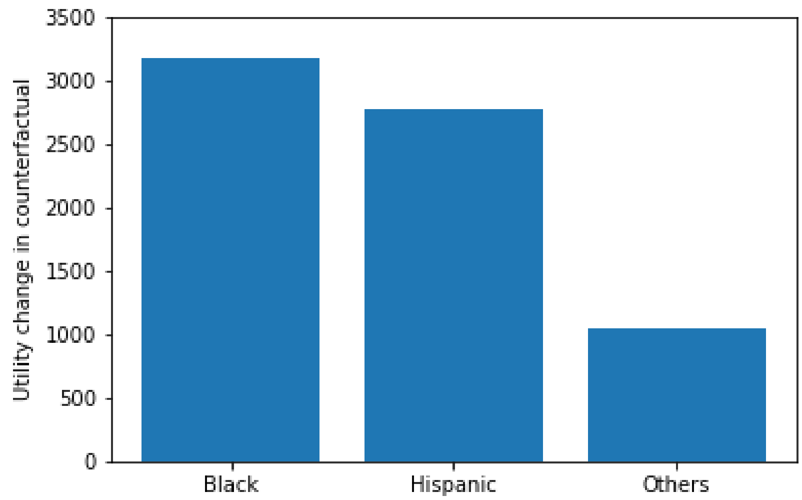


Figure A.8: Counterfactual change in utility from automatically refinancing, by racial group



A.7 ADL

A.7.1 Implications of market pricing of low upfront closing cost mortgages

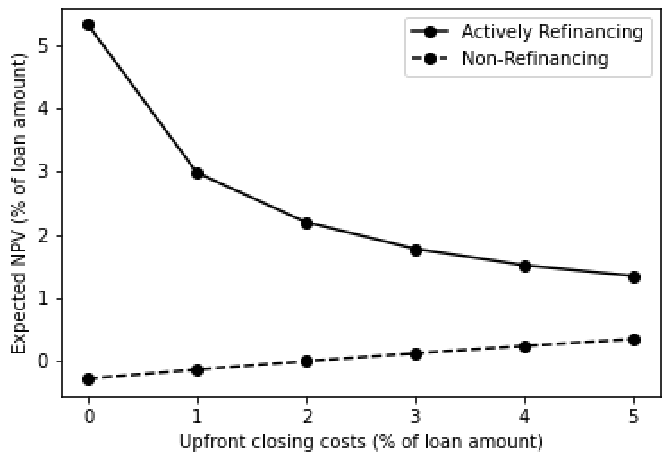
Next, I show that the pricing of lower upfront closing cost mortgages is especially favorable to the benchmark optimally refinancing borrowers of Agarwal, Driscoll, and Laibson (2013b) but unfavorable to non-refinancing borrowers. Figure A.9 compares the expected NPV of an optimally refinancing borrower from the model of Agarwal, Driscoll, and Laibson (2013b) under a variety of choices of upfront closing costs with that of a non-refinancing borrower. Both borrowers have a moving hazard of 10% per year, and interest rate movements are assumed to be Brownian motion with a calibrated annualized standard deviation from the post-crisis period.[14] The borrowers are assumed to always pick the same upfront closing costs, which is justifiable as a benchmark by the stationarity of the problem.

Figure A.9a shows that the expected NPV of the optimally refinancing borrower increases significantly with lower upfront closing costs at a market price of +0.17% in higher interest rate per percentage of the loan amount reduction in upfront closing costs. In particular, the expected NPV for the optimally refinancing borrower goes from 1.3% of the loan amount if the borrower were to always choose a mortgage with an upfront closing cost of 5% to 5.3% if the borrower were to always choose a "no closing cost" mortgage, a fourfold increase. The reason for this increase is that the interest rate increase for lower upfront closing cost mortgages is too low for the optimally refinancing borrower: it is significantly cross-subsidized, and the optimally refinancing borrower is able to derive significant value by choosing a lower upfront closing cost and then refinancing. Indeed, the pattern is opposite, though less drastic, for the non-refinancing borrower: their expected NPV is 0.3% with a 5% upfront closing cost and -0.3% with no upfront closing costs. The fact that non-refinancing borrowers (indeed, almost all borrowers) choose lower upfront closing cost mortgages implies that they cross-subsidize the optimally refinancing borrowers' mortgages. Comparing across the two types of borrowers, Figure A.9b shows that the prevalence of low to zero upfront closing cost mortgages increases the NPV advantage to optimally refinancing by a factor of five.

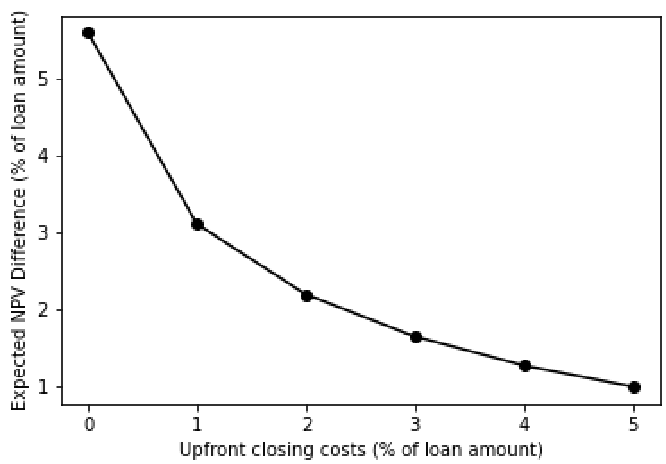
---

[14]For the period of January 2010 to January 2021, I find the annualized standard deviation of monthly average mortgage interest rates from the Freddie Mac Mortgage Survey to be $\sigma = 0.4597\%$.

Figure A.9: Expected NPV by Upfront Closing Costs



(a) NPV by upfront closing costs

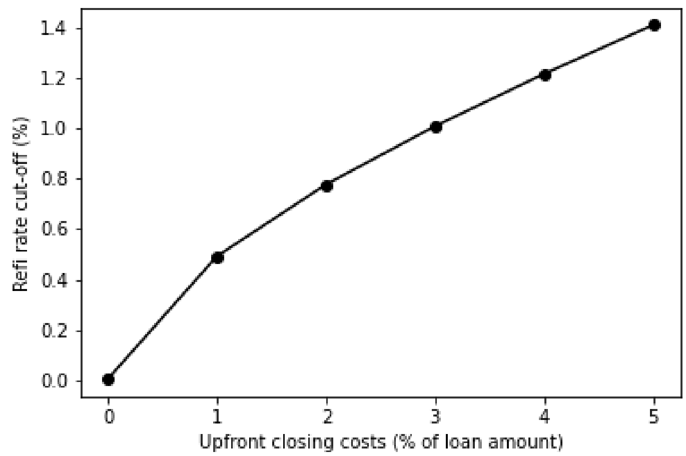


(b) NPV advantage to actively refinancing

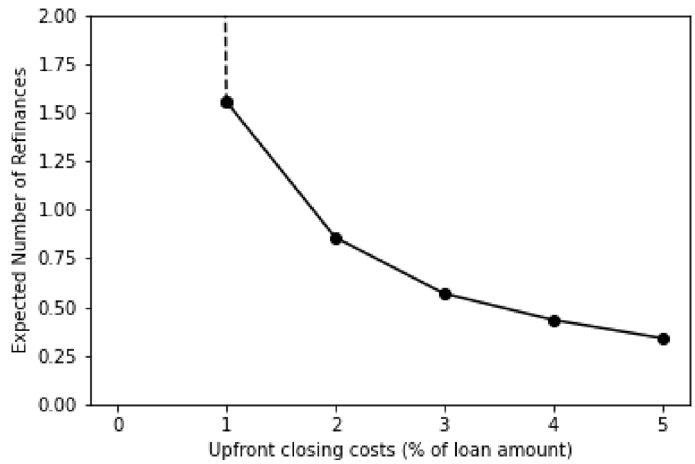
Figure A.10 shows the implications of lower upfront closing cost mortgages for the optimally refinancing borrower's "refi cutoff" and expected number of refinances per mortgage. The "refi cutoff" is computed directly from the formula of Agarwal, Driscoll, and Laibson (2013b), and as shown in Figure A.10a it is significantly declining in the level of upfront closing cost chosen, from 1.4% for mortgage with all closing costs paid upfront to 0% for a no upfront closing cost mortgage. Correspondingly, the expected number of refinances per mortgage, which I compute by simulating Brownian motion paths for interest rates and counting the number of refinances, is significantly increasing in upfront closing cost choice as shown in Figure A.10b. It goes from 0.25 per mortgage with 5% in closing costs, to 1.25 refinances per mortgage for a 1% in closing costs, to a degenerate number with zero closing cost mortgage.[15]

[15] Because I simulate the Brownian motion of Agarwal, Driscoll, and Laibson (2013b) in monthly intervals, I get 9.3 expected refinances per new mortgage origination with zero closing cost mortgages.

Figure A.10: Optimal refinancing strategy with upfront closing cost choice



(a) Interest rate decrease needed for refinance



(b) Expected number of refinances per mortgage

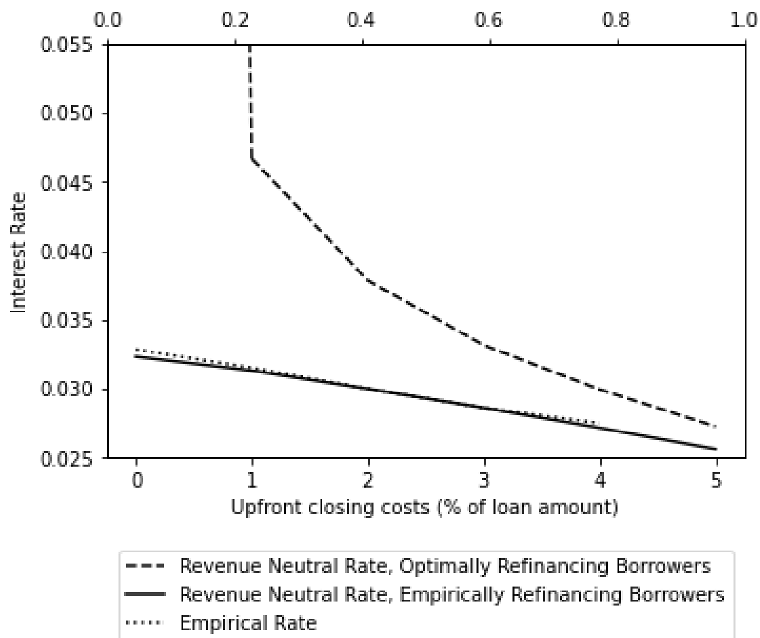
A.7.2 Counterfactual pricing of low upfront closing cost mortgages

Next, I investigate the cross-subsidization of low upfront closing cost mortgages from the perspective of lender rate-setting under my model of the competitive supply side. Instead of using my model, I use the model of Agarwal, Driscoll, and Laibson (2013b) for my model of actively refinancing borrower behavior.

Results are shown in Figure A.11. In particular, I find that the revenue neutral rate and upfront closing cost trade-off matches the empirically observed rate and upfront closing cost well. Furthermore, Figure A.11 suggests that actively refinancing borrowers (in the sense of Agarwal, Driscoll, and Laibson (2013b)) receive a substantially lower rate than what they would have received under perfect information: if all borrowers were optimally refinancing, lenders would charge substantially higher interest rates particularly for low upfront closing cost mortgages, on the order of 1.49% more with a 1% upfront closing cost mortgage compared to only 0.13% with a 5% upfront closing cost mortgage.[16] In other words, the interest rates on lower upfront closing cost mortgages appear to be substantially discounted for optimally refinancing borrowers due to the presence of non-refinancing borrowers in the market.

[16]With zero upfront closing costs and optimally refinancing borrowers, I find that lenders would have to charge a rate of 57% to remain revenue-neutral.

Figure A.11: Pricing of mortgages mortgages by upfront closing cost choice and borrower type



A.8 Proofs of results of Chapter 2

A.8.1 Proof of Theorem 1

Proof. In the forward direction, if such a $\pi(x_1, x_2)$ exists, then it is possible to construct a series of menus $\mathbf{m} = \{x_1, x_2\}$ for each $\{x_1, x_2\}$ such that $\varphi(x_1, x_2) = 1$, each appearing with probability $g_1(\mathbf{m}) = g_2(\mathbf{m}) = \pi(x_1, x_2)$, from which group 1 consumers chose x_1 and group 2 consumers chose x_2. Under this construction, the distributions of menus across the two groups are equal such that $\mathbf{M}_1 = \mathbf{M}_2$, and the choice probabilities are rationalized. The reverse direction follows from the fact that, denoting $c(x_1, x_2|\mathbf{m})$ by the probability that group 1 consumers choose $x_1 \in \mathbf{m}$ and group 2 consumers choose $x_2 \in \mathbf{m}$ given a menu $\mathbf{m}$, we can construct the required $\pi(x_1, x_2) =$

$\sum_{\mathbf{m}} c(x_1, x_2|\mathbf{m})g(\mathbf{m})$ for any $c(x_1, x_2|\mathbf{m}), g(\mathbf{m})$ where $g(\mathbf{m}) = g_1(\mathbf{m}) = g_2(\mathbf{m})$. □

A.8.2 Proof of Theorem 2

Proof. Under constant marginal utility in interest rate, Equations (2.11) and (2.12) imply that $u_i(\mathbf{m}_j) - u_i(\mathbf{m}_i) = \beta d_{i\to j}(\mathbf{m}_i, \mathbf{m}_j) \geq \beta \underline{d}_{1\to 2}(x_1, x_2)$ for some constant β. Then, by property of minimum we must have $\beta \underline{DIM}_{1\to 2} \leq \Delta W_{\mathcal{I}_1\to\mathcal{I}_2,\pi}, \forall \pi$, with $\beta \equiv 1$ if utility were measured in terms of interest rate paid. □

A.8.3 Proof of Theorem 3

Proof. First, we show that ϕ is Gâteaux directionally differentiable in the sense that the limit:

$$\phi'_{\mathbf{p}_1,\mathbf{p}_2}(\mathbf{h}_1, \mathbf{h}_2) = \lim_{t\to 0^+} \frac{\phi(\mathbf{p}_1 + t\mathbf{h}_1, \mathbf{p}_2 + t\mathbf{h}_2) - \phi(\mathbf{p}_1, \mathbf{p}_2)}{t} \tag{A.4}$$

exists for all and is equal to that given by Equation (2.22) for $\{\mathbf{p}_1, \mathbf{p}_2\} \in \mathbb{D}_\phi$ and $\{\mathbf{h}_1, \mathbf{h}_2\} \in \mathbb{D}_0$. To do this, without loss of generality letting $\varphi^* = \varphi + M, M = \sup|\varphi|$ such that $\varphi^* \geq 0$ and $\phi = \phi^* - M$, we transform the problem to standard linear programming form:

$$\phi^* = \min_{\pi} E_\pi \varphi^* \text{ s.t. } E_{x_2}\pi \geq \mathbf{p}_1, E_{x_1}\pi \geq \mathbf{p}_2, \pi \geq 0 \tag{A.5}$$

and then use Theorem 3.1 of Gal and Greenberg (2012), who set out conditions for the Gâteaux directional differentiability of standard linear programs with inequality constraints. In particular, we need to check primal and dual stability in the sense that, if we let $\Pi^*(p_1, p_2)$ be the set of primal solutions to the linear programming problem in Equation (A.5), then the set of primal solutions reachable from perturbations in the direction $\{h_1, h_2\}$ is non-empty, such that:

$$\Pi^{\infty}(\{\mathbf{p}_1, \mathbf{p}_2\}, \{\mathbf{h}_1, \mathbf{h}_2\}) = \{\pi : \{\pi^k\} \to \pi \text{ for some } \{\varepsilon_k \to 0^+\}, \tag{A.6}$$

$$\text{with } \pi^k \in \Pi^*(\mathbf{p}_1 + \varepsilon_k \mathbf{h}_1, \mathbf{p}_2 + \varepsilon_k \mathbf{h}_2)\} \neq \emptyset, \tag{A.7}$$

and analogously for dual solutions. This can be done by referencing existing results:

1. Since $\{\mathbf{p}_1 + \varepsilon_k \mathbf{h}_1, \mathbf{p}_2 + \varepsilon_k \mathbf{h}_2\}$ are a series of probability measures for $\varepsilon_k \leq 1$, primal stability in the sense of Equation (A.6) is guaranteed by Theorem 5.19 in Villani (2008).
2. Similarly, dual stability is guaranteed by Theorem 1.52 in Santambrogio (2015), in particular by taking the sequence of c-concave Kantorovich potentials corresponding to $(\mathbf{p}_1 + \varepsilon_k \mathbf{h}_1, \mathbf{p}_2 + \varepsilon_k \mathbf{h}_2)\}$.

Therefore, ϕ^* is Gâteaux differentiable. Second, we show that ϕ^* is Lipschitz, such that the Gâteaux directionally differentiability of ϕ^* is equivalent to Hadamard directionally differentiability following Shapiro (1990). For the l_1 norm $l_1(\{\mathbf{p}_{1,1}, \mathbf{p}_{1,2}\}, \{\mathbf{p}_{2,1}, \mathbf{p}_{2,2}\}) = \sum_x |\mathbf{p}_{1,1}(x) - \mathbf{p}_{1,2}(x)| + |\mathbf{p}_{2,1}(x) - \mathbf{p}_{2,2}(x)|$, we will show that:

$$|\phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}) - \phi^*(\mathbf{p}_{2,1}, \mathbf{p}_{2,2})| \leq M l_1(\{\mathbf{p}_{1,1}, \mathbf{p}_{1,2}\}, \{\mathbf{p}_{2,1}, \mathbf{p}_{2,2}\}) \tag{A.8}$$

More specifically, let $\mathbf{p}_1^- = \min\{\mathbf{p}_{1,1}, \mathbf{p}_{1,2}\}$, $\mathbf{p}_1^+ = \max\{\mathbf{p}_{1,1}, \mathbf{p}_{1,2}\}$, and analogously for $\mathbf{p}_2^-, \mathbf{p}_2^+$. By construction we know that $\phi^*(\mathbf{p}_1^-, \mathbf{p}_2^-) \leq \phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}), \phi^*(\mathbf{p}_{2,1}, \mathbf{p}_{2,2})$ and $\phi^*(\mathbf{p}_1^+, \mathbf{p}_2^+) \geq$

$\phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}), \phi^*(\mathbf{p}_{2,1}, \mathbf{p}_{2,2})$, and therefore:

$$\phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}) - \phi^*(\mathbf{p}_1^+, \mathbf{p}_2^+) \leq \phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}) - \phi^*(\mathbf{p}_{2,1}, \mathbf{p}_{2,2}) \leq \phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}) - \phi^*(\mathbf{p}_1^-, \mathbf{p}_2^-) \tag{A.9}$$

Furthermore, we know that $\phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}) - \phi^*(\mathbf{p}_1^-, \mathbf{p}_2^-) \leq Ml_1$ since taking the optimal plan from $\pi^- = \phi^*(\{\mathbf{p}_1^-, \mathbf{p}_2^-\})$ and then constructing a plan $\pi^{-,*} = \pi^- + (\mathbf{p}_{1,1} - \mathbf{p}_1^-)(\mathbf{p}_{1,2} - \mathbf{p}_2^-)$ yields an upper bound for the value value $\phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}) \leq E_{\pi^{-,*}}\varphi = \phi^*(\{\mathbf{p}_1^-, \mathbf{p}_2^-\}) + E_{(\mathbf{p}_{1,1}-\mathbf{p}_1^-)(\mathbf{p}_{1,2}-\mathbf{p}_2^-)}\varphi \leq \phi^*(\{\mathbf{p}_1^-, \mathbf{p}_2^-\}) + M\sum(\mathbf{p}_{1,1} - \mathbf{p}_1^-)(\mathbf{p}_{1,2} - \mathbf{p}_2^-) \leq \phi^*(\{\mathbf{p}_1^-, \mathbf{p}_2^-\}) + Ml_1$. Similarly, we have that $\phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}) - \phi^*(\mathbf{p}_1^+, \mathbf{p}_2^+) \geq -Ml_1$. Substituting into Equation (A.9) yields:

$$-Ml_1 \leq \phi^*(\mathbf{p}_{1,1}, \mathbf{p}_{1,2}) - \phi^*(\mathbf{p}_{2,1}, \mathbf{p}_{2,2}) \leq Ml_1 \tag{A.10}$$

which implies Equation (A.8) and that the mapping ϕ^* is Lipschitz.

□

A.8.4 Proof of Corollary 1

Proof. Suppose we knew the true $\mathbf{p}_1, \mathbf{p}_2$. Then, the "oracle" analogue of the critical value at level $1 - \alpha + \beta$ conditional on the true $\mathbf{p}_1, \mathbf{p}_2$ is:

$$\hat{c}^o_{1-\alpha+\beta} = \inf\{c \in \mathbb{R} : \Pr(\hat{\psi}'_{\mathbf{p}_1,\mathbf{p}_2}(\mathbf{h}_1, \mathbf{h}_2) \leq c) \geq 1 - \alpha + \beta\} \tag{A.11}$$

By construction, $\hat{c}_{1-\alpha+\beta} \geq \hat{c}^{o}_{1-\alpha+\beta}$ whenever $[\mathbf{p}_1, \mathbf{p}_2] \in \hat{\mathcal{P}}_{\beta}$. Therefore, if we let the event $[\mathbf{p}_1, \mathbf{p}_2] \in \hat{\mathcal{P}}_{n,\beta}$ be E and $[\mathbf{p}_1, \mathbf{p}_2] \notin \hat{\mathcal{P}}_{n,\beta}$ be $\neg E$, it follows that:

$$\Pr(\hat{\psi} \geq \hat{c}_{1-\alpha+\beta}) = \Pr(\hat{\psi} \geq \hat{c}_{1-\alpha+\beta}|E)\Pr(E) + \Pr(\hat{\psi} \geq \hat{c}_{1-\alpha+\beta}|\neg E)\Pr(\neg E) \tag{A.12}$$

$$\leq \Pr(\hat{\psi} \geq \hat{c}^{o}_{1-\alpha+\beta})\Pr(E) + \Pr(\hat{\psi} \geq \hat{c}_{1-\alpha+\beta}|\neg E)\Pr(\neg E) \tag{A.13}$$

$$\leq \Pr(\hat{\psi} \geq \hat{c}^{o}_{1-\alpha+\beta}) + \Pr(\neg E) \tag{A.14}$$

By property of limsup, we know that:

$$\limsup_{n\to\infty} \Pr_{\mathbf{p}_1,\mathbf{p}_2}(\hat{\phi}_n \geq \hat{c}_{n,1-\alpha+\beta}) \leq \limsup_{n\to\infty} \Pr(\hat{\psi}_n \geq \hat{c}^{o}_{n,1-\alpha+\beta}) + \limsup_{n\to\infty} \Pr([\mathbf{p}_1, \mathbf{p}_2] \notin \hat{\mathcal{P}}_{n,\beta}) \tag{A.15}$$

By Equation (2.24) and the Portmanteau Theorem, we have:

$$\limsup_{n\to\infty} \Pr_{\mathbf{p}_1,\mathbf{p}_2}((\hat{\phi}_n \geq \hat{c}^{o}_{n,1-\alpha+\beta})) \leq \alpha - \beta \tag{A.16}$$

By Equation (2.25) for the uniform confidence band, we know that:

$$\limsup_{n\to\infty} \Pr_{\mathbf{p}_1,\mathbf{p}_2}([\mathbf{p}_1, \mathbf{p}_2] \notin \hat{\mathcal{P}}_{n,\beta}) \leq \beta \tag{A.17}$$

Combining the two parts, we have:

$$\limsup_{n\to\infty} \Pr_{\mathbf{p}_1,\mathbf{p}_2}(\hat{\phi}_n \geq \hat{c}_{n,1-\alpha+\beta}) \leq \beta + (\alpha - \beta) = \alpha \tag{A.18}$$

as required. □

A.9 Additional Tables and Figures for Chapter 2

Figure A.12: Screenshot of advertised rate and closing cost table

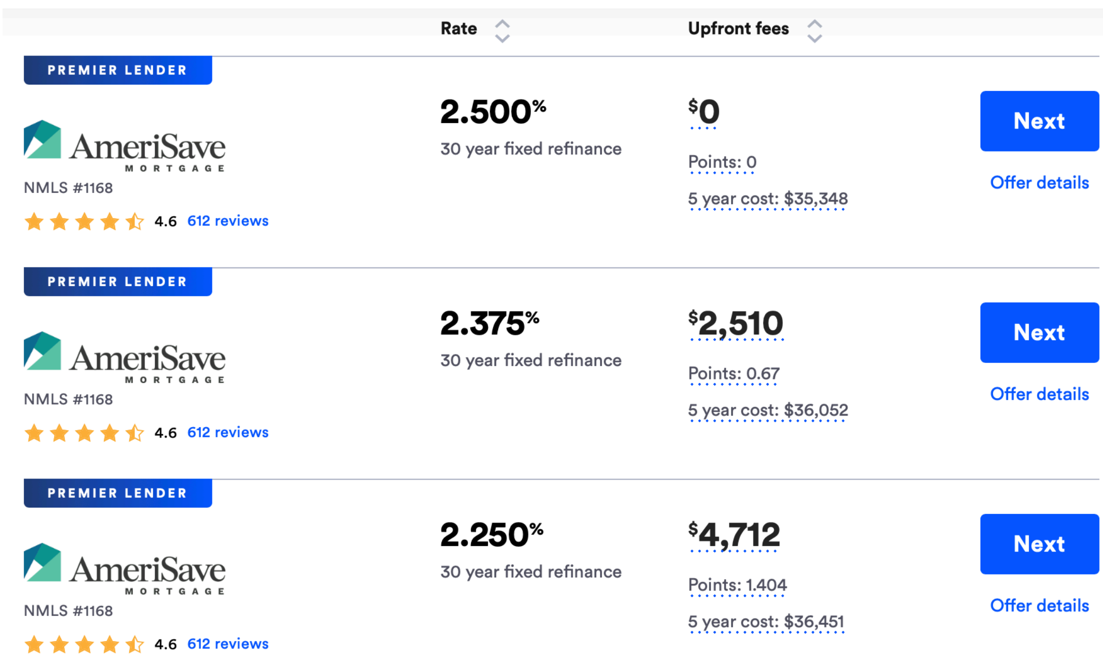


Figure A.13: How the choice of which menu dimension to condition on can yield contradictory results

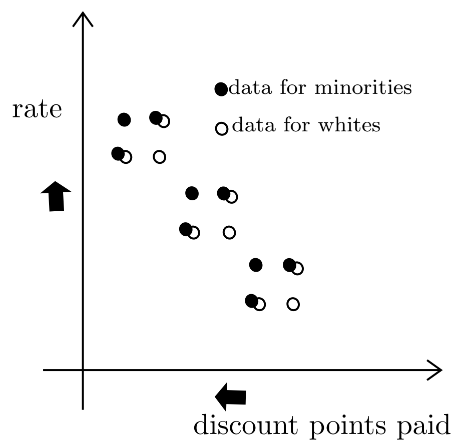
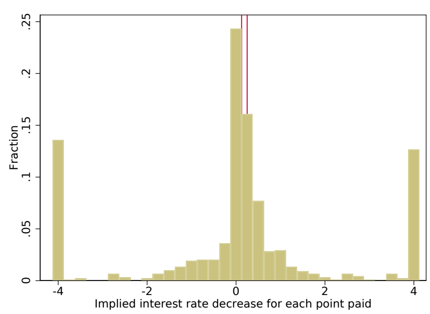


Figure A.14: Implied interest rate decrease per point paid, based on randomly sampled pairs of borrowers

(a) Conforming mortgages



(b) FHA mortgages

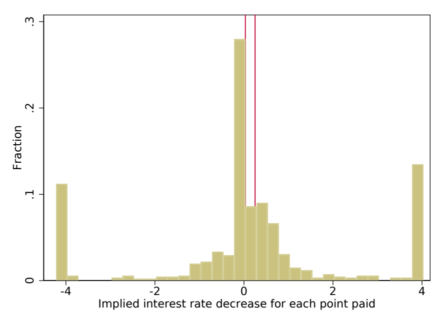


Table A.9: Other approaches to inference in optimal transport, compared with our size-corrected directional derivatives approach

	HSS (2021)	m out of n subsampling			Size-corrected directional derivatives
		$m = n^{2/3}$	$m = n^{1/2}$	$m = n^{1/3}$	
$n_1 = n_2 = 500$	0.0	31.8	21.9	16.2	4.4
$n_1 = n_2 = 1000$	0.0	28.3	18.4	12.5	4.7
$n_1 = n_2 = 5000$	0.0	20.0	11.3	8.6	4.8
$n_1 = n_2 = 10000$	0.1	18.3	10.3	8.4	5.8
$n_1 = n_2 = 50000$	0.0	13.7	8.6	6.5	5.0

Note: Entries represent the probability of rejecting the null at the 5 [percent level. The control of the subsampling approach is taken with subsample size m as indicated, with $n = \frac{n_1 n_2}{n_1 + n_2} = \frac{1}{2} n_1$. The control of our size-corrected directional derivatives approach was computed via 2000 sample draws and 500 draws of h_1, h_2 from the estimated asymptotic multivariate Normal distribution for p_1, p_2 within each sample draw.

Table A.10: Regressions of origination costs and total loan costs as a percentage of the loan amount on HMDA's information on points (hmda_points) versus Optimal Blue's information on points (ob_points)

	orig_charge_frac			tot_loan_costs_frac		
	(1)	(2)	(3)	(4)	(5)	(6)
hmda_points	0.915***		0.902***	0.940***		0.941***
	(0.00277)		(0.00308)	(0.00372)		(0.00463)
ob_points		0.469***	0.0200***		0.474***	0.00531
		(0.00377)	(0.00220)		(0.00434)	(0.00361)
_cons	0.580***	0.560***	0.575***	2.194***	2.174***	2.190***
	(0.000432)	(0.00129)	(0.000591)	(0.000577)	(0.00149)	(0.000927)
N	1224911	1221338	1221208	1224417	1220921	1220715
R^2	0.553	0.231	0.553	0.233	0.093	0.232

Standard errors in parentheses

$^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Note: The sample consists of the conforming and FHA purchase mortgages originated within the retail channel within our 2018–2019 HMDA-Optimal Blue matched sample. In each regression, we excluded observations with extreme outliers for points (below −4 or above 4) and for origination costs and total loan costs as a percentage of the loan amount below −3 percent or above 10 percent. All regressions include lender by county by year by product type fixed effects. Standard errors were also clustered at the lender by county by year by product type level.

Table A.11: Assessments of lender discrimination using two heuristic approaches in the Black and non-Hispanic white matched sample

	Heuristic 1		Heuristic 2				Alternate Heuristic 1	
	Conforming	FHA	Conforming		FHA		Conforming	FHA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	points	points	rate_1/8	rate_1/4	rate_1/32	rate_1/4	rate	rate
black	4.482***	-0.908	2.905***	3.309***	3.032***	2.519***	2.627***	2.983***
	(0.994)	(1.319)	(0.327)	(0.376)	(0.436)	(0.509)	(0.322)	(0.437)
Rate Decile FE	Yes	Yes	No	No	No	No	No	No
Points Decile FE	No	No	No	No	No	No	Yes	Yes
N	12278	9208	12278	12278	9208	9208	12278	9208
R^2	0.041	0.025	0.006	0.006	0.005	0.003	0.023	0.025

Standard errors in parentheses

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

A.10 HMDA-McDash and HMDA-McDash-CRISM Match Rates

As we discussed in section 3.2, our analysis employs a novel data set that combines three sources of administrative data: Home Mortgage Disclosure Act (HMDA) data, Black Knight McDash mortgage servicing data, and credit bureau data from Equifax. The three data sources are linked together through two separate loan-level matches: a match between the HMDA and McDash databases, which we will refer to as the HMDA-McDash dataset, and a match between the McDash and Equifax databases, which is referred to as CRISM (Equifax Credit Risk Insight Servicing Mc-

Dash Database). We are then able to merge the two matched data sets, creating a final data set with information from all three sources, which we will refer to as the HMDA-McDash-CRISM data set. Below we will discuss some of the details of both matches and show match rates by loan vintage (year) to provide information on the quality and scope of the final data set used in the analysis.

A.10.1 HMDA-McDash Database

The HMDA-McDash matched data set is available to users within the Federal Reserve System and includes more than 93 million loans originated from 1992 through2015 (inclusive). The matching algorithm was written by the Risk Assessment, Data Analysis and Research (RADAR) group at the Federal Reserve Bank of Philadelphia and matches HMDA and McDash loans by the origination date, origination amount, property Zip code, lien type, loan purpose (that is, purchase or refinance), loan type (for example, conventional or FHA), and occupancy type. Tables A.12 and A.12 display match rates by origination year;the former table calculates rates by dividing by the number of McDash loans, while the latter table divides by the total number of HMDA loans. Overall, approximately two-thirds of McDash loans are successfully matched to HMDA, while almost 40 percent of HMDA loans are successfully matched to loans in McDash. Since the HMDA database covers a greater fraction of the mortgage market, the match rates normalized by HMDA loans are significantly lower than the rates normalized by McDash loans.

Our sample includes only loans originated in 2005 and later due to lower coverage in the pre-2005 McDash database. In 2005 McDash added a large servicer to its database, which substantially increased the overall coverage of the database. The last column in Table A.12 shows that the coverage (relative to the total number of HMDA loan originations) goes from 65 percent in 2004 to 81 percent in 2006. When servicers are added to the McDash database, they typically provide information on only their active loans. This raises concerns of attrition bias, and thus we focus only on loans originated in 2005 and later.

The matching algorithm is based on the following logic:

- Origination date (McDash) and action date (HMDA) must be within five days of each other.
- Origination amounts must be within $500.
- Property Zip codes must match.
- Lien types must match.
- Loan purposes (purchase, refinance) must match.
- Loan types (conventional, jumbo, etc.) must match.
- Occupancy types must match.

In our analysis, we use only loans that were uniquely matched. The last column in Table A.13 shows that during our sample period (2005 through 2015) our sample covers from 34 percent to 47 percent of all loan originations in HMDA.

Table A.12: Match Rate by Origination Year (Matched McDash Mortgages/All McDash Mortgages)

Origination Year	McDash Loans Matched	Only 1 HMDA Candidate	McDash Loans Uniquely Matched	McDash Coverage
1992	51%	48%	20%	58%
1993	55%	50%	19%	70%
1994	58%	53%	24%	52%
1995	61%	57%	29%	46%
1996	63%	58%	33%	42%
1997	62%	58%	35%	39%
1998	65%	60%	36%	52%
1999	65%	60%	35%	46%
2000	64%	61%	50%	31%
2001	64%	60%	49%	44%
2002	65%	59%	50%	50%
2003	71%	64%	53%	67%
2004	69%	64%	55%	65%
2005	**67%**	**61%**	**51%**	**73%**
2006	**63%**	**59%**	**49%**	**81%**
2007	**63%**	**59%**	**50%**	**87%**
2008	**65%**	**62%**	**54%**	**79%**
2009	**67%**	**64%**	**59%**	**79%**
2010	**69%**	**67%**	**61%**	**77%**
2011	**69%**	**67%**	**61%**	**73%**
2012	**73%**	**71%**	**64%**	**67%**
2013	**75%**	**74%**	**67%**	**62%**
2014	**77%**	**76%**	**71%**	**48%**
2015	**79%**	**78%**	**75%**	**45%**
Total	66%	62%	49%	61%

Notes: Match rates are calculated by the Risk Assessment, Data Analysis and Research (RADAR) group. McDash coverage is estimated by dividing the number of originations in the McDash database by the number of originations in HMDA.

Table A.13: Match Rate by Origination Year (Matched HMDA Mortgages/All HMDA Mortgages)

Origination Year	HMDA Loans Matched	Only 1 McDash Candidate	HMDA Loans Uniquely Matched
1992	21%	14%	12%
1993	27%	16%	13%
1994	22%	15%	12%
1995	22%	15%	13%
1996	21%	16%	14%
1997	21%	16%	14%
1998	30%	23%	19%
1999	25%	19%	16%
2000	19%	17%	16%
2001	27%	24%	22%
2002	33%	30%	25%
2003	48%	43%	36%
2004	45%	41%	36%
2005	**48%**	**43%**	**37%**
2006	**50%**	**45%**	**40%**
2007	**53%**	**48%**	**43%**
2008	**49%**	**46%**	**43%**
2009	**53%**	**50%**	**47%**
2010	**53%**	**50%**	**47%**
2011	**49%**	**47%**	**45%**
2012	**47%**	**45%**	**42%**
2013	**46%**	**44%**	**42%**
2014	**37%**	**35%**	**35%**
2015	**36%**	**35%**	**34%**
Total	38%	34%	30%

Notes: Match rates are calculated by the Risk Assessment, Data Analysis and Research (RADAR) group.

A.10.2 CRISM Database

CRISM is a data set that consists of McDash mortgages matched to credit bureau data from Equifax at the borrower level. The Equifax credit bureau data are updated at a monthly frequency and include information on outstanding consumer loans and credit lines for the primary borrower as well as all co-borrowers associated with the McDash mortgage. The matching process was conducted by Equifax using confidential and proprietary data. The exact matching algorithm is proprietary, but according to Equifax, anonymous fields such as the original and current mortgage balance, date of origination, ZIP code, and monthly payment history are all used in the algorithm.

CRISM coverage begins in June 2005, and according to Equifax, approximately 90 percent of McDash mortgages were matched to a credit bureau account with high confidence.[17] We keep only observations that pertain to the primary mortgage borrower to avoid double counting. Borrower credit information is included in the data set for the life of each loan as well as for the six months preceding origination and the six months following termination.

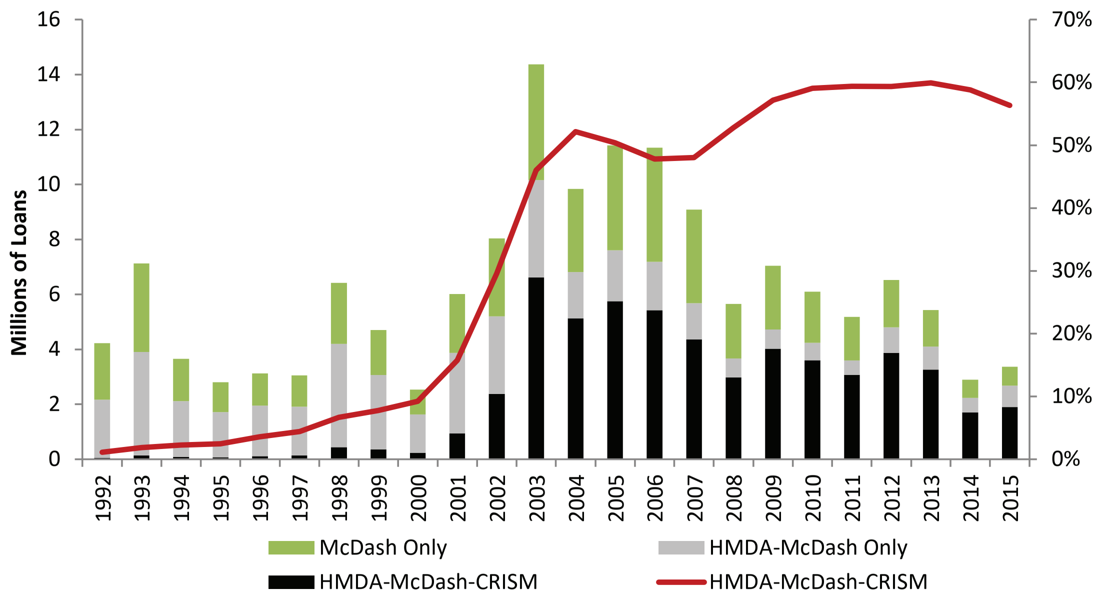
Figure A.15 displays the match rate by vintage for the HMDA-McDash-CRISM matched data set as a ratio of the total number of McDash originations (solid red line). For 2005–2015 originations, the match rate is between 50 percent and 60 percent. The figure also shows the total number of mortgage originations for the McDash data set, the HMDA-McDash matched data set, and the

[17] Equifax provides a "Match Confidence Score" that is based on a scale of 0 to 0.9, where a higher score indicates that the McDash and Equifax data align better on the matching fields. Approximately 90 percent of McDash loans have a match confidence score of 0.8 or higher. Equifax recommends using 0.8 as a threshold for modeling purposes, and we follow this advice, keeping only matches with scores above 0.8.

HMDA-McDash-CRISM matched data set. The largest decline in the sample occurs when the McDash database is matched to HMDA. The addition of CRISM data results in only a small decline in loan originations during our sample period.

Finally, in Table A.14 we compare summary statistics for the HMDA-McDash and HMDA-McDash-CRISM GSE (Panel A) and FHA (Panel B) samples, respectively. The tables show that the summary statistics are almost identical across the two samples, which suggests that the addition of the Equifax credit bureau data does not significantly alter the composition of mortgages.

Figure A.15: Loans in the HMDA-McDash-CRISM Match, HMDA-CRISM Match, and McDash Data Sets by Vintage



Notes: This figure shows the number of loans in the McDash data set, the HMDA-McDash data set, and the HMDA-McDash-CRISM data set by vintage (bars). In addition, the solid red line shows the match rate for the HMDA-McDash-CRISM data set, calculated as a percentage of the number of loans in the McDash database by vintage. This figure was created by the Risk Assessment, Data Analysis and Research (RADAR) group, which conducted the matching exercise.

Table A.14: Comparison of Summary Statistics: HMDA-McDash vs. HMDA-McDash-CRISM Databases

Panel A: GSE Loans

	HMDA-McDash-Equifax		HMDA-McDash	
	Mean	Std. Dev.	Mean	Std. Dev.
FICO Origination (100s points)	7.44	0.54	7.45	0.53
LTV (%)	72.6	15.9	72.7	15.9
Loan Amount ($100k)	2.12	1.13	2.12	1.13
Interest Rate (ppts)	5.20	1.02	5.20	1.02
Income ($1k)	97.6	64.0	97.5	63.9
Refinance (d)	0.538	0.499	0.539	0.498
Condo (d)	0.140	0.347	0.139	0.346
2-4 Family (d)	0.018	0.133	0.018	0.133
Low Documentation (d)	0.308	0.462	0.309	0.462
Non-Occupant Owner (d)	0.140	0.347	0.140	0.347
Female (d)	0.294	0.455	0.294	0.456
Co-applicant (d)	0.505	0.500	0.503	0.500
# Loans	800,806		1,076,117	

Panel B: FHA Loans

	HMDA-McDash-Equifax		HMDA-McDash	
	Mean	Std. Dev.	Mean	Std. Dev.
FICO Origination (100s points)	6.85	0.60	6.88	0.59
LTV (%)	93.6	7.5	93.6	7.4
Loan Amount ($100k)	1.73	0.91	1.73	0.91
Interest Rate (ppts)	4.93	1.00	4.93	1.00
Income ($1k)	65.8	37.5	65.8	37.3
Refinance (d)	0.294	0.456	0.295	0.456
Condo (d)	0.115	0.318	0.114	0.317
2-4 Family (d)	0.014	0.119	0.015	0.120
Low Documentation (d)	0.190	0.393	0.191	0.393
Non-Occupant Owner (d)	0.033	0.178	0.033	0.178
Female (d)	0.353	0.478	0.352	0.478
Co-applicant (d)	0.414	0.493	0.415	0.493
# Loans	295,487		397,686	

Notes: This table reports summary statistics from a 7.5% random sample of GSE loans originated between 2005 and 2015 (inclusive) from a matched HMDA-McDash-CRISM data set and a 10% random sample of GSE and FHA loans originated between 2005 and 2015 (inclusive) from a matched HMDA-McDash data set. The label (d) denotes dummy variables.

A.11 Sample Restrictions

Table A.15 below displays all of the restrictions that we impose in constructing our 7.5 percent random sample of the HMDA-McDash-CRISM data set. We adopt most of the restrictions implemented in Fuster et al. (2018). We implement most of our restrictions while querying the database (and thus, we do not know how many loans are lost as a result of those restrictions).[18] For the restrictions that we implement while applying code to clean and create our variables, we display the number of loans that are dropped.

Table A.15: Sample Restrictions

Sample Restriction:	# Loans Lost	# Loans Remaining
Originations between 01/2005 and 12/2015		
Loans with "conf" $\geq$ 0.80		
Fixed Rate Loans		
First Liens		
Fully Amortizing Loans No Prepayment Penalty		
$20 \leq$ LTV ≤ 100		
Occupancy Non-missing		
Loan Amount $\leq$ \$1m		
Income $\leq$ \$500k		
Term = 30 years		
No Home Improvement Loans		1,681,252
Seasoning $\leq$ 6 Months	193,898	1,487,354
Black, Hispanic, Asian, and White Borrowers	208,817	1,278,537
GSE and FHA Loans	179,810	1,098,727
3% $\leq$ Mortgage Rate $\leq$ 8%	2,434	1,096,293

[18] Because the HMDA-McDash-CRISM database is a monthly panel and extremely large, we were unable to download more than a 7.5 percent sample given computing constraints.

A.12 LPM Estimates with Zip Code-by-Time Fixed Effects

In Table A.16 below we display refinance regression estimates for specifications that include Zip code-by-year and Zip code-by-year-quarter fixed effects. The specifications are otherwise identical to those displayed in columns (6) and (9) in Table 3.3 for GSE and FHA loans respectively.

Table A.16: Refinance Results with Zip Code-by-Time Fixed Effects

Dependent Variable: Prepay Refinance (d)	GSE Loans		FHA Loans	
	(1)	(2)	(3)	(4)
Black (d)	-0.166***	-0.158***	-0.174***	-0.164***
	(0.026)	(0.028)	(0.027)	(0.029)
Hispanic (d)	-0.299***	-0.295***	-0.328***	-0.313***
	(0.041)	(0.041)	(0.042)	(0.041)
Asian (d)	0.200***	0.204***	0.022	0.035
	(0.067)	(0.067)	(0.072)	(0.072)
LTV Change	-0.026***	-0.024***	-0.016***	-0.012***
	(0.002)	(0.003)	(0.002)	(0.002)
Refinance (d)	-0.253***	-0.259***	-0.201***	-0.215***
	(0.049)	(0.048)	(0.043)	(0.045)
Female (d)	-0.083***	-0.082***	-0.101***	-0.098***
	(0.014)	(0.014)	(0.017)	(0.017)
Call Option			0.198***	0.478***
			(0.024)	(0.075)
Call Option V2	0.271***	0.336***		
	(0.027)	(0.023)		
SATO	-1.634***	-2.210***	-0.719***	-3.409***
	(0.188)	(0.174)	(0.169)	(0.592)
Risk Score Change	0.340***	0.236***	0.839***	0.829***
	(0.059)	(0.044)	(0.081)	(0.083)
Loan Age	X	X	X	X
Underwriting Vars	X	X	X	X
HMDA Vars	X	X	X	X
Vintage Year-Qtr FE	X	X	X	X
Zip Code-by-Year FE	X		X	
Zip Code-by-Year-Qtr FE		X		X
# Observations	10,954,271	10,714,510	3,947,819	3,713,341
# Loans	600,758	584,786	192,615	176,908
R^2	0.043	0.087	0.067	0.149

A.13 Delinquency as a Competing Risk and Controlling for Missed Payments

In our empirical analysis of refinance disparities (Section 3), we treat other types of loan termination as competing risks to refinances. These include prepayment due to home sale, involuntary termination due to foreclosure (or other type of distressed sale), and loan transfers to servicers that are not in the McDash dataset. We do not treat default, which we define to be when a borrower first becomes 90-days delinquent, as a competing risk. This means that quarterly observations after the first 90-day delinquency remain in the sample until the loan terminates. This could be problematic and introduce bias in our estimates of racial disparities in refinancing since borrowers who are behind on their payments or who are current but have missed a payment in the recent past, are usually ineligible to refinance.

In this section we will address this issue by implementing two exercises. First, we re-estimate our primary refinance regressions (Table 3.3) assuming that 90-day delinquency is a competing risk to refinancing. That is, we eliminate from our sample all observations after the first 90-day delinquency, so that a loan in default is no longer at-risk to refinance (i.e. it is treated as a right-censored observation). We report the results of this exercise in Table A.17 below. The table includes the exact same specifications as Table 3.3. The estimated racial refinance disparities are very similar to those documented in Table 3.3, with only minor differences in the Black, Hispanic, and Asian coefficients.

While the regressions reported in Table A.17 address the fact that borrowers in default are unlikely to be eligible to refinance into a new mortgage, they do not address the possibility that a borrower who has missed less than three payments is also unlikely to be able to refinance. Thus, in Table A.18

below, we re-estimate our most saturated refinance regression specification (columns 6 and 9 in Table 3.3) and include a control for borrowers that are behind on their payments, "Behind on Payments" (columns (1) and (3)) and a control for borrowers that have missed a payment in the previous 12 months (columns (2) and (4)). The coefficients associated with these controls are negative, large, and statistically significant, which confirms that borrowers who are behind on their mortgage payments are much less likely to be eligible to refinance. Comparing the race coefficients in Table A.18 to columns (6) and (9) in Table 3.3, we see only small differences, however. Thus, controlling for minor mortgage delinquencies does not materially affect our estimates of racial refinance disparities. Including controls for missed payments does significantly reduce the coefficient associated with the change in credit score ("Risk Score Change") however. This suggests that controlling for the change in credit score partially accounts for borrowers who have missed payments.

Table A.17: Refinance LPM Results: Treating Delinquency as a Competing Risk

Dependent Variable: Prepay Refinance (d)	GSE Loans						FHA Loans		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black (d)	-0.673***	-0.325***	-0.263***	-0.225***	-0.223***	-0.134***	-0.510***	-0.275***	-0.200***
	(0.079)	(0.043)	(0.033)	(0.031)	(0.032)	(0.026)	(0.056)	(0.031)	(0.036)
Hispanic (d)	-0.609***	-0.374***	-0.393***	-0.377***	-0.374***	-0.246***	-0.395***	-0.339***	-0.397***
	(0.109)	(0.055)	(0.055)	(0.053)	(0.057)	(0.039)	(0.084)	(0.037)	(0.048)
Asian (d)	0.461***	0.279***	0.296***	0.286***	0.327***	0.233***	0.415***	-0.051	0.082
	(0.144)	(0.093)	(0.098)	(0.098)	(0.100)	(0.068)	(0.095)	(0.066)	(0.078)
Equifax Risk Score		0.293***						0.280***	
		(0.057)						(0.042)	
LTV Origination		-0.009***						-0.008***	
		(0.002)						(0.002)	
Loan Amount		0.606***						0.932***	
		(0.062)						(0.074)	
LTV Change		0.007*	-0.042***	-0.041***	-0.018***	-0.016***		-0.002	-0.027***
		(0.004)	(0.004)	(0.004)	(0.005)	(0.005)		(0.003)	(0.002)
Refinance (d)		-0.176***	-0.224***	-0.205***	-0.209***	-0.218***		-0.056	-0.093*
		(0.060)	(0.050)	(0.049)	(0.049)	(0.050)		(0.042)	(0.050)
Female (d)			-0.066***	-0.065***	-0.063***	-0.086***			-0.106***
			(0.013)	(0.012)	(0.012)	(0.014)			(0.018)
Call Option			0.249***	0.252***					0.170***
			(0.017)	(0.017)					(0.017)
Call Option V2					0.241***	0.242***			
					(0.018)	(0.018)			
SATO			-1.528***	-1.508***	-1.309***	-1.289***			-0.246*
			(0.120)	(0.117)	(0.106)	(0.109)			(0.124)
Risk Score Change				0.538***	0.123**	0.107**			0.666***
				(0.070)	(0.047)	(0.047)			(0.060)
Loan Age	X	X	X	X	X	X	X	X	X
Underwriting Vars		X	X	X	X	X		X	X
HMDA Vars			X	X	X	X			X
Vintage Year-Qtr FE	X	X	X	X	X	X	X	X	X
State FE		X	X	X	X			X	
Zip Code FE						X			X
# Observations	14,883,532	12,125,625	11,684,983	11,613,933	10,683,997	10,683,981	5,484,924	4,057,993	3,514,044
# Loans	792,823	652,106	629,224	629,154	600,792	600,792	291,587	221,036	192,645
R^2	0.008	0.013	0.020	0.020	0.021	0.024	0.006	0.014	0.019

Table A.18: Refinance LPM Results: Controlling for Missed Payments

Dependent Variable: Prepay Refinance (d)				
	GSE Loans		FHA Loans	
	(1)	(2)	(3)	(4)
Black (d)	-0.133***	-0.129***	-0.194***	-0.185***
	(0.026)	(0.026)	(0.030)	(0.030)
Hispanic (d)	-0.270***	-0.269***	-0.392***	-0.391***
	(0.040)	(0.039)	(0.046)	(0.046)
Asian (d)	0.230***	0.229***	0.079	0.076
	(0.069)	(0.069)	(0.073)	(0.073)
LTV Change	-0.019***	-0.019***	-0.028***	-0.028***
	(0.004)	(0.004)	(0.002)	(0.002)
Refinance (d)	-0.216***	-0.213***	-0.115***	-0.107**
	(0.049)	(0.049)	(0.042)	(0.042)
Female (d)	-0.092***	-0.092***	-0.100***	-0.100***
	(0.014)	(0.014)	(0.016)	(0.016)
Call Option			0.142***	0.142***
			(0.015)	(0.014)
Call Option V2	0.235***	0.235***		
	(0.017)	(0.017)		
SATO	-1.246***	-1.243***	-0.167	-0.156
	(0.102)	(0.100)	(0.114)	(0.113)
Risk Score Change	0.096**	0.077*	0.527***	0.495***
	(0.044)	(0.039)	(0.043)	(0.038)
Missed Last Payment (d)	-1.965***		-1.506***	
	(0.108)		(0.176)	
Missed Payment (d)		-1.612***		-1.391***
(Previous 12 months)		(0.131)		(0.178)
Loan Age	X	X	X	X
Underwriting Vars	X	X	X	X
HMDA Vars	X	X	X	X
Vintage Year-Qtr FE	X	X	X	X
Zip Code FE	X	X	X	X
# Observations	10,887,548	10,887,548	3,920,375	3,920,375
# Loans	600,792	600,792	192,645	192,645
R^2	0.024	0.024	0.018	0.018

A.14 Logit Models

In this section we present prepayment due to refinance and home sale results as well as default results from logit models. These models are estimated on a 7.5 percent random sample of our HMDA-McDash-Equifax matched data set. Table A.19 contains the refinance results, Tablel A.20 contains the home sale results, and Table A.21 displays the default results. Both tables show the estimated average marginal effects associated with the racial/ethnic indicator variables. The covariates and fixed effects in each column correspond exactly to their counterparts in Tables 3.3, 3.4, and 3.5 in the main text. The omitted specifications are those with Zip code fixed effects. It was not possible to estimate those specifications using the logit framework.

Table A.19: Logit Prepayment due to Refinance Hazard Estimates

Dependent Variable: Prepay Refinance (d)						
	GSE Loans				FHA Loans	
	(1)	(2)	(3)	(4)	(7)	(8)
Black (d)	-0.686***	-0.421***	-0.353***	-0.282***	-0.585***	-0.419***
	(0.033)	(0.030)	(0.029)	(0.030)	(0.037)	(0.024)
Hispanic (d)	-0.654***	-0.475***	-0.489***	-0.449***	-0.405***	-0.389***
	(0.057)	(0.024)	(0.030)	(0.031)	(0.059)	(0.028)
Asian (d)	0.466***	0.275***	0.259***	0.247***	0.455***	-0.030
	(0.132)	(0.070)	(0.070)	(0.070)	(0.088)	(0.043)
Loan Age	X	X	X	X	X	X
Underwriting Vars		X	X	X		X
HMDA Vars			X	X		X
Vintage Year-Qtr FE	X	X	X	X	X	X
State FE		X	X	X		X
# Observations	15,460,588	11,983,398	11,547,035	11,469,141	6,184,502	4,316,733

Notes: This table reports estimated marginal effects estimates from a logit model of equation (3.1)—the likelihood of voluntary mortgage prepayment due to refinancing on a set of race/ethnicity indicator variables. The estimation is performed at the quarterly frequency on a 5% random sample of loans from a matched HMDA-McDash-Equifax data set. The unit of observation is a loan-quarter. Underwriting variables include the borrower's risk score at origination, LTV at origination, loan amount, change in LTV since origination, indicators for condos and 2–4 multi-family properties, low-documentation loans, non-owner occupant properties, and refinance loans. HMDA variables include borrower age (2nd order polynomial), borrower income, and indicators for gender and co-applicants. All columns include a 3rd order polynomial for the number of quarters since origination (duration). Standard errors are clustered by county. (*** $p<0.01$, ** $p<0.05$, * $p< 0.1$)

Table A.20: Logit Prepayment due to Sale Hazard Estimates

Dependent Variable: Prepay Sale (d)						
	GSE Loans				FHA Loans	
	(1)	(2)	(3)	(4)	(7)	(8)
Black (d)	-0.505***	-0.440***	-0.415***	-0.414***	-0.633***	-0.587***
	(0.014)	(0.013)	(0.012)	(0.012)	(0.017)	(0.016)
Hispanic (d)	-0.417***	-0.347***	-0.340***	-0.340***	-0.501***	-0.524***
	(0.017)	(0.015)	(0.018)	(0.018)	(0.022)	(0.020)
Asian (d)	-0.189***	-0.190***	-0.208***	-0.209***	-0.234***	-0.339***
	(0.020)	(0.019)	(0.020)	(0.020)	(0.031)	(0.024)
Loan Age	X	X	X	X	X	X
Underwriting Vars		X	X	X		X
HMDA Vars			X	X		X
Vintage Year-Qtr FE	X	X	X	X	X	X
State FE		X	X	X		X
# Observations	15,460,588	11,983,398	11,547,035	11,469,141	6,184,502	4,316,733

Notes: This table reports estimated marginal effects estimates from a logit model of equation (3.1)—the likelihood of voluntary mortgage prepayment due to sale on a set of race/ethnicity indicator variables. The estimation is performed at the quarterly frequency on a 5% random sample of loans from a matched HMDA-McDash-Equifax data set. The unit of observation is a loan-quarter. Underwriting variables include the borrower's risk score at origination, LTV at origination, loan amount, change in LTV since origination, indicators for condos and 2–4 multi-family properties, low-documentation loans, non-owner occupant properties, and refinance loans. HMDA variables include borrower age (2nd order polynomial), borrower income, and indicators for gender and co-applicants. All columns include a 3rd order polynomial for the number of quarters since origination (duration). Standard errors are clustered by county. (*** $p<0.01$, ** $p<0.05$, * $p< 0.1$)

Table A.21: Logit Default Hazard Estimates

Dependent Variable: Default (d)	GSE Loans			FHA Loans	
	(1)	(2)	(3)	(6)	(7)
Black (d)	0.350***	0.149***	0.101***	0.719***	0.340***
	(0.023)	(0.013)	(0.012)	(0.031)	(0.025)
Hispanic (d)	0.362***	0.185***	0.132***	0.162***	0.158***
	(0.038)	(0.012)	(0.012)	(0.028)	(0.025)
Asian (d)	0.015	0.011	0.010	-0.163***	-0.098***
	(0.015)	(0.012)	(0.012)	(0.030)	(0.037)
Loan Age	X	X	X	X	X
Underwriting Vars		X	X		X
HMDA Vars			X		X
Vintage Year-Qtr FE	X	X	X	X	X
State FE		X	X		X
# Observations	9,929,254	7,705,281	7,424,419	3,653,447	2,558,071

Notes: This table reports estimated marginal effects estimates from a logit model of the likelihood of mortgage default on a set of race/ethnicity indicator variables. The estimation is performed at the quarterly frequency on a 5% random sample of loans from a matched HMDA-McDash-Equifax data set. The unit of observation is a loan-quarter. Underwriting variables include the borrower's risk score at origination, LTV at origination, loan amount, change in LTV since origination, indicators for condos and 2–4 multi-family properties, low-documentation loans, non-owner occupant properties, and refinance loans. HMDA variables include borrower age (2nd order polynomial), borrower income, and indicators for gender and co-applicants. All columns include a 3rd order polynomial for the number of quarters since origination (duration). Standard errors are clustered by county. (*** p<0.01, ** p<0.05, * p< 0.1)

A.15 Involuntary Prepayments

In Section 3.3.4 we showed that minority borrowers are more likely to default on their loans. The default definition that we use in that section is based on borrowers becoming seriously delinquent on their loans by missing at least three payments (that is, 90-plus days past due). We now consider an alternative definition of default that focuses on involuntary mortgage prepayment due to foreclosure and/or distressed sale (that is, short sales). Like our voluntary prepayment variables (refinance and home sale), this default definition identifies a terminal state, and is likely more correlated with the actual losses that lenders experience on distressed loans. As such, it is likely more relevant to mortgage pricing than a serious delinquency definition of default.

Table A.22 displays the estimation results. The table is identical in structure to Table 3.5, with the only difference being the dependent variable. The results are very different, however. In column (1) we see that minority borrowers are significantly more likely to lose their homes due to foreclosure, and the magnitudes are large.[19] However, as we add more controls and fixed effects, the differences disappear. In our most saturated model with Zip-code-by-year-quarter fixed effects, minority GSE borrowers are significantly *less* likely to lose their homes to foreclosure. We see a similar pattern in the FHA sample, as Black borrowers are more than 8 percentage points less likely to lose their homes to foreclosure compared with white borrowers (column (8)).

[19]The sample average for involuntary prepayment is approximately 0.11 percentage point.

Table A.22: Involuntary Prepayment/Foreclosure Results

Dependent Variable: Involuntary Prepayment/Foreclosure (d)	GSE Loans				FHA Loans		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Black (d)	0.095***	0.022	0.006	-0.012	-0.007	-0.072***	-0.099***
	(0.024)	(0.017)	(0.008)	(0.008)	(0.014)	(0.017)	(0.019)
Hispanic (d)	0.142***	0.042***	-0.018***	-0.010	-0.027*	-0.021*	-0.042***
	(0.036)	(0.015)	(0.006)	(0.006)	(0.014)	(0.011)	(0.012)
Asian (d)	0.020*	0.025***	0.016***	0.012**	-0.066***	-0.027**	-0.018
	(0.011)	(0.007)	(0.005)	(0.005)	(0.015)	(0.013)	(0.015)
Risk Score Origination		-0.165***				-0.212***	
		(0.024)				(0.018)	
LTV Origination		0.006***				0.009***	
		(0.001)				(0.001)	
LTV Change		0.022***	0.011***	0.012***		0.019***	0.017***
		(0.002)	(0.001)	(0.001)		(0.002)	(0.002)
Female (d)			-0.007***	-0.006**			-0.031***
			(0.002)	(0.002)			(0.008)
Loan Age	X	X	X	X	X	X	X
Underwriting Vars		X	X	X		X	X
HMDA Vars			X	X			X
Vintage Year-Qtr FE	X	X	X	X	X	X	X
State FE		X	X			X	
Zip Code FE				X			X
# Observations	15,460,588	12,572,069	10,960,278	10,960,264	6,184,502	4,563,659	3,970,642
# Loans	792,823	622,936	601,094	601,028	291,587	209,827	182,517
R^2	0.004	0.008	0.004	0.007	0.004	0.006	0.007

A.16 Agarwal, Driscoll, and Laibson Closed-Form Refinance Rule

In this section we proxy for the moneyness of the prepayment option using an alternative measure developed by Agarwal, Driscoll, and Laibson (2013a) (hereafter ADL). ADL derived a closed-form solution for the optimal time to refinance from a borrower's perspective. Specifically, the rule states that a borrower should refinance when the current mortgage interest rate falls below the original rate by at least:

$$\frac{1}{\psi}\left[\varphi + W(-exp(-\varphi))\right]$$

where W is the Lambert W-function and

$$\psi = \frac{\sqrt{2\rho+\lambda}}{\sigma}$$

$$\varphi = 1 + \psi(\rho+\lambda)\frac{\kappa/M}{(1-\tau)}$$

$$\lambda = \mu + \frac{i_0}{exp[i_0\Gamma]-1} + \pi$$

In these expressions ρ is the discount rate, μ is the expected probability of moving, σ is the standard deviation of the mortgage rate, $\frac{\kappa/M}{1-\tau}$ is the ratio of the tax-adjusted refinancing cost and the remaining mortgage value, Γ is the remaining maturity of the mortgage, i_0 is the original mortgage rate, π is the expected inflation rate, and τ is the marginal tax rate. We assume the following parameter values, where σ is estimated by taking the standard deviation of changes in the Freddie Mac Primary Market Mortgage Survey rate from April 1971 to August 2020:

$\rho =$	0.02
$\sigma =$	0.95
$\pi =$	0.02
$\mu_b =$	0.02
$\mu_w =$	0.04
$\frac{\kappa/M}{1-\tau} =$	$\frac{2000}{M} + 0.01$

We assume different mobility rates, μ_b, μ_w, for Black and white borrowers, respectively, which we annualize based on the quarterly hazards from Table 3.1.[20] We specify two variables based on the above threshold. First, we create an indicator variable, *ADL Dummy*, which takes a value of 1 if the difference between the borrower's current interest rate and the market rate (PMMS survey) is greater than the ADL threshold. Second, we create a continuous variable, *ADL*, which measures how much higher/lower the difference between the current rate and market rate is from the ADL threshold. Positive values of *ADL* imply that the refinance option is in the money given the borrower type specific moving hazards and refi costs, while negative values imply that it is not.

We then re-estimate equation (3.2) and substitute our ADL variables for *Call Option*, which is our proxy for the moneyness of the refinance option from Deng, Quigley, and Van Order (2000b). We focus on the specifications in columns (1) and (2) of Table 3.6. Column (1) includes only a control for the moneyness of the option, while column (2) includes interactions between the moneyness of the option and the race dummies. Table A.23 displays the results. In columns (1) and (2) we show

[20]For simplicity we assume the same mobility rate for Black and Hispanic households.

results for the *Call Option* variable applied to the sample of loans with non-missing ADL values. Columns (3) and (4) display results for the *ADL Dummy*, and columns (5) and (6) display results for the *ADL* continuous variable.

A few notable patterns emerge from Table A.23. First, both ADL variables are positive and statistically significant as expected, which suggests that borrowers are more likely to refinance when their option is in the money. However, columns (2), (4), and (6) show that the refinancing behavior of minority borrowers is much less sensitive to changes in the value of the option. In fact, these differences appear to be much larger when we use the ADL variables, as the interaction coefficients in columns (4) and (6) are of about the same magnitude, but with the opposite sign as the ADL coefficients by themselves. This implies that minority borrowers are actually insensitive to macroeconomic changes in rates that make their prepayment option more valuable.

Table A.23: Prepayment due to Refinance with Interaction Effects

Dependent Variable: Prepay Refinance (d)						
	GSE Loans					
	(1)	(2)	(3)	(4)	(5)	(6)
Black (d)	-0.119***	0.288***	-0.116***	0.225***	-0.110***	-0.113***
	(0.021)	(0.046)	(0.020)	(0.037)	(0.021)	(0.034)
Hispanic (d)	-0.196***	0.150***	-0.197***	0.126***	-0.189***	-0.217***
	(0.025)	(0.043)	(0.025)	(0.032)	(0.025)	(0.024)
Call Option	0.170***	0.175***				
	(0.012)	(0.012)				
Black * Call Option		-0.050***				
		(0.004)				
Hispanic * Call Option		-0.046***				
		(0.004)				
ADL Dummy			0.864***	0.945***		
			(0.073)	(0.080)		
Black * ADL Dummy				-0.744***		
				(0.059)		
Hispanic * ADL Dummy				-0.748***		
				(0.066)		
ADL					0.506***	0.562***
					(0.048)	(0.049)
Black * ADL						-0.421***
						(0.032)
Hispanic * ADL						-0.422***
						(0.034)
Loan Age	X	X	X	X	X	X
Underwriting Vars	X	X	X	X	X	X
HMDA Vars	X	X	X	X	X	X
Vintage Year-Qtr FE	X	X	X	X	X	X
Zip Code FE	X	X	X	X	X	X
# Observations	11,105,357	11,105,357	11,105,357	11,105,357	11,105,357	11,105,357
R^2	0.016	0.016	0.012	0.012	0.012	0.012

A.17 Evidence from Survey of Consumer Finances

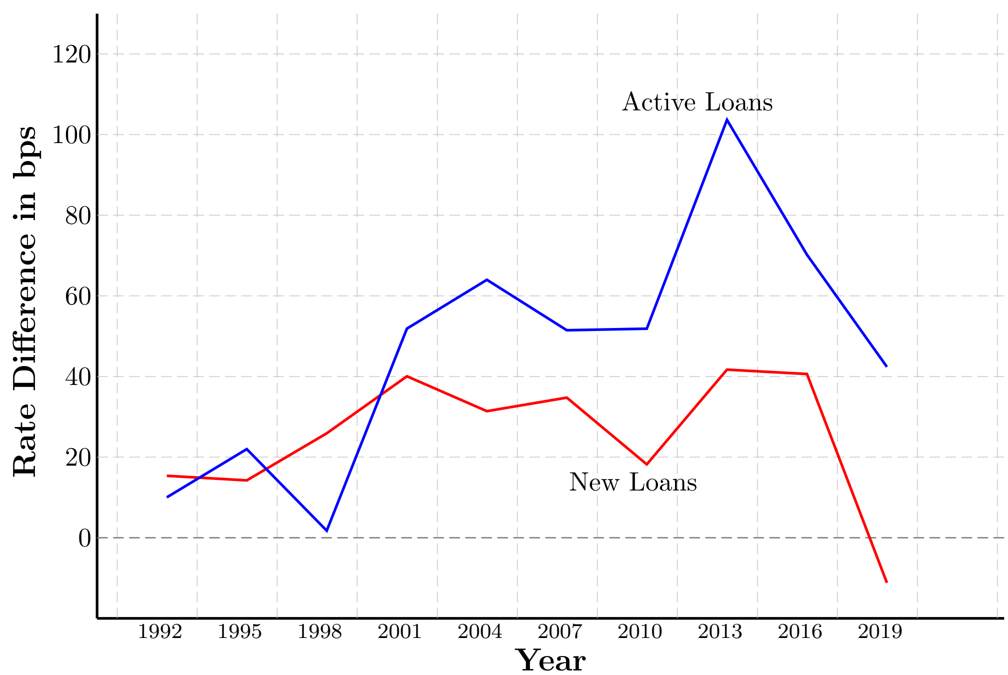
In this section we use data from the 1992–2019 Survey of Consumer Finances (SCF) to examine the rate gap between Black and white borrowers for active loans as well as new loans originated that year.

Data construction is as follows. For comparison with the fixed-rate, conforming, and FHA mortgages used in our main analysis, the observations are from respondents who report having a non-adjustable-rate mortgage (X820=5) that is either non-federally guaranteed (X724=1) or a FHA loan (X726=1) with a loan amount at origination (X804) of less than $450,000. The current interest rate is reported in X816, from which we remove outlier rates that are less than 2 percent or more than 4 percent over the average Freddie Mac PMMS rate during the year of origination, which is comparable to the restriction of rate to 3 percent to 8 percent in our main analysis.

The SCF definition of race underwent a slight revision in 1998 to include more categories. For the 1992–1995 SCF, we define respondent race based on Question X5909, "Are you Native American, Asian, Hispanic, black, white, or another race?", with an answer of 4 ("black or African-American") being our definition of a Black respondent and an answer of 5 ("white") being our definition of a White respondent. In the 1998–2019 SCF, we define respondent race based on the revised Question X6809, which asks, "Which of these categories do you feel best describe you: (white, black or African-American, Hispanic or Latino, Asian, American Indian or Alaska Native, Hawaiian Native or other Pacific Islander, or another race?)," with an answer of 2 ("black or African-American") being our definition of a Black respondent and answer of 1 ("white") being our definition of a white respondent.

We compute the mean rate for all active loans by respondent race using the provided survey weights by race to compute the active loan-rate gap. For the new loans rate gap, we take means by respondent race and by the year of origination (X802) rounded to the nearest SCF survey year. The mean rate differences between Black and white borrowers for active and new loans are shown in Figure A.16. While the estimates are much more noisy due to a smaller sample size (and potential survey error), we do find that the rate gap for active loans is higher than the rate gap for new loans, consistent with Figure 3.1 in the main text.

Figure A.16: Gap between interest rates for Black and white borrowers based on data from the SCF

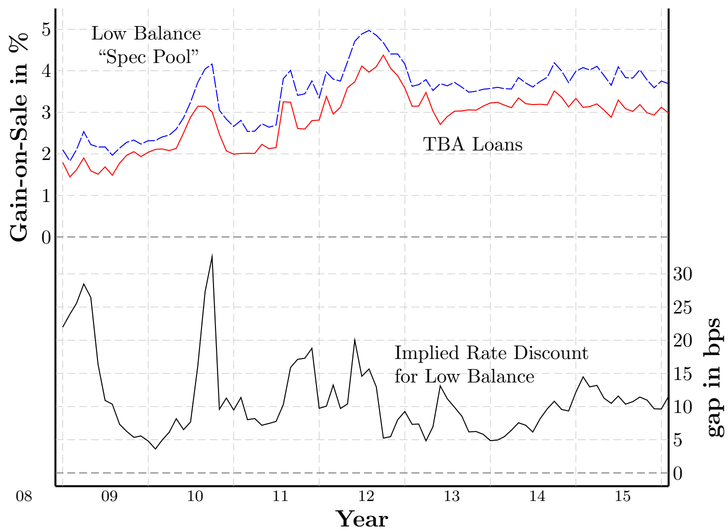


Notes: This graph reports estimates of the rate difference in bps between newly originated and active loans based on data from the Survey of Consumer Finances (SCF).

A.18 Analysis of Pricing Implications

In this section we present results related to the pricing implications in Section 3.4 of the main text.

Figure A.17: Mortgage pricing for low prepayment loans.



Notes: TBA loans are loans sold in "TBA" pools. Low Balance Spec Pool are "LLB" loans defined as loans with balances of less than $85K. Gain-on-sale is the gap between par and the interpolated price of an MBS paying a coupon equal to the FHLMC Primary Mortgage Market Survey 30-year FRM rate less the g-fee. Implied rate discount is the gap between the FHLMC PMMS 30-year FRM rate and the interest rate that yields the same gain-on-sale for an LLB mortgage.

Table A.24: Prepayment, Race, and Loan Amount

Dependent Variable: Prepay (d)		
	GSE Loans	
	(1)	(2)
Black (d)	-1.628***	
	(0.117)	
Hispanic (d)	-1.342***	
	(0.158)	
Asian (d)	0.440**	
	(0.188)	
Orig Amount ≤ 85k (d)		-1.697***
		(0.165)
85k < Orig Amount < 110k (d)		-1.225***
		(0.137)
110k < Orig Amount < 125k (d)		-1.050***
		(0.127)
125k < Orig Amount < 150k (d)		-0.854***
		(0.109)
150k < Orig Amount < 175k (d)		-0.681***
		(0.096)
Loan Age	X	X
Underwriting Vars		
HMDA Vars		
Vintage Year-Qtr FE	X	X
State FE		
Zip Code FE		
Zip Code-by-Year-Qtr FE		
# Observations	15,460,588	15,460,588
R^2	0.009	0.009

Table A.25: Payups regression with all covariates

	(1)	(2)
	≥$85k trades	≥$1mil trades
Tract_Black	0.913***	1.304***
	(0.289)	(0.208)
Loan size		
85-110k	-0.390***	-0.381***
	(0.0295)	(0.0280)
110-125k	-0.601***	-0.582***
	(0.0307)	(0.0289)
125-150k	-0.712***	-0.681***
	(0.0304)	(0.0281)
150-175k	-0.921***	-0.898***
	(0.0294)	(0.0276)
175-200k	-1.126***	-1.113***
	(0.0334)	(0.0312)
over 200k	-1.212***	-1.208***
	(0.0320)	(0.0303)
720≤Credit Score≤750	0.314***	0.334***
	(0.0436)	(0.0433)

Credit Score≥750	0.418***	0.421***
	(0.0440)	(0.0442)
80 ≤ LTV ≤ 90	-0.0871***	-0.0710***
	(0.0158)	(0.0146)
90 ≤ LTV ≤ 95	-0.387***	-0.329***
	(0.0481)	(0.0361)
LTV≥95	-0.360***	-0.235***
	(0.0760)	(0.0717)
Rate Spread	-0.202***	-0.222***
	(0.0769)	(0.0747)
Rate Spread Sq	-1.040***	-1.021***
	(0.121)	(0.120)
Rate Spread Cube	0.197***	0.177***
	(0.0601)	(0.0599)
Trade Size	5.559***	4.056***
	(0.878)	(1.302)
Trade Size Sq	-0.348***	-0.255***
	(0.0557)	(0.0808)
Trade Size Cube	0.00717***	0.00527***
	(0.00117)	(0.00166)

WeekXCoupon FE	Yes	Yes
Seller FE	Yes	Yes
Observations	16208	15139
R^2	0.724	0.748

Robust standard errors in parentheses

Standard errors clustered by CUSIPXweek

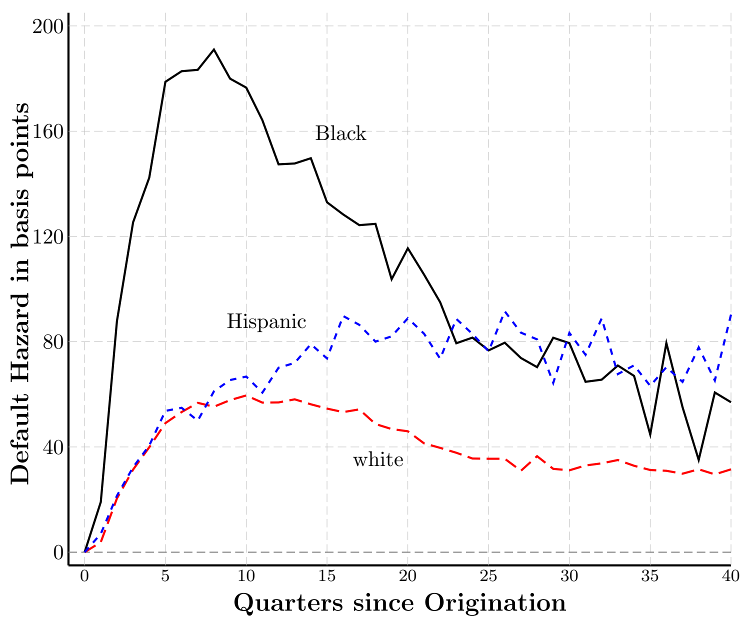
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.19 Additional Tables and Figures

Table A.26: Prepayment LPM Results

Dependent Variable: Prepay (d)	GSE Loans						FHA Loans			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Black (d)	-1.628***	-1.040***	-0.967***	-0.867***	-0.690***	-0.694***	-1.447***	-1.074***	-0.818***	-0.723***
	(0.117)	(0.060)	(0.050)	(0.047)	(0.040)	(0.042)	(0.066)	(0.050)	(0.056)	(0.053)
Hispanic (d)	-1.342***	-0.977***	-0.948***	-0.887***	-0.667***	-0.682***	-0.932***	-1.017***	-0.972***	-0.852***
	(0.158)	(0.089)	(0.087)	(0.078)	(0.049)	(0.048)	(0.099)	(0.048)	(0.054)	(0.049)
Asian (d)	0.440**	0.176	0.202	0.191	0.106	0.107	0.290**	-0.329***	-0.216**	-0.269***
	(0.188)	(0.125)	(0.131)	(0.131)	(0.098)	(0.097)	(0.128)	(0.071)	(0.085)	(0.073)
Risk Score Origination		0.482***						0.653***		
		(0.089)						(0.071)		
LTV Origination		-0.024***						-0.018***		
		(0.004)						(0.002)		
Loan Amount		0.765***						0.984***		
		(0.084)						(0.066)		
LTV Change		-0.025***	-0.089***	-0.084***	-0.084***	-0.075***		-0.051***	-0.077***	-0.038***
		(0.006)	(0.006)	(0.005)	(0.005)	(0.006)		(0.004)	(0.005)	(0.004)
Refinance (d)		-0.602***	-0.593***	-0.561***	-0.553***	-0.619***		-0.438***	-0.350***	-0.415***
		(0.090)	(0.079)	(0.075)	(0.075)	(0.073)		(0.048)	(0.042)	(0.047)
Female (d)			-0.079***	-0.072***	-0.102***	-0.107***			-0.075***	-0.110***
			(0.017)	(0.017)	(0.019)	(0.020)			(0.021)	(0.021)
Refi Money			0.404***	0.410***	0.411***	1.128***			0.220***	0.960***
			(0.026)	(0.026)	(0.027)	(0.083)			(0.025)	(0.102)
SATO			-1.965***	-1.907***	-1.846***	-7.215***			-0.216	-5.761***
			(0.171)	(0.162)	(0.163)	(0.586)			(0.168)	(0.625)
Risk Score Change				0.899***	0.874***	0.908***			1.198***	1.172***
				(0.130)	(0.129)	(0.116)			(0.100)	(0.098)
Loan Age	X	X	X	X	X	X	X	X	X	X
Underwriting Vars		X	X	X	X	X		X	X	X
HMDA Vars			X	X	X	X			X	X
Vintage Year-Qtr FE	X	X	X	X	X	X	X	X	X	X
State FE		X	X	X				X		
Zip Code FE					X				X	
Zip Code-by-Year-Qtr FE						X				X
# Observations	15,460,588	11,983,398	11,547,035	11,469,141	11,469,141	11,318,445	6,184,502	4,316,733	3,732,349	3,559,947
R^2	0.009	0.012	0.020	0.020	0.023	0.080	0.006	0.013	0.019	0.146

Figure A.18: Kaplan Meier unconditional default hazard rates



Notes: This figure displays the Kaplan-Meier hazard estimates of default broken down by racial/ethnic groups. The Kaplan-Meier estimate of the hazard function is: $\lambda_p(t_j) = \frac{d_{pj}}{n_j}$, where the number of loans that have reached time t_j without being terminated or censored is given by n_j, and the number of terminations due to default at t_j is given by d_{pj}. The underlying data come from the Black Knight McDash database.

References

Abadie, Alberto. 2020. "Statistical Nonsignificance in Empirical Economics." *American Economic Review: Insights* 2(2): 193–208.

Abel, Joshua, and Andreas Fuster. 2021. "How Do Mortgage Refinances Affect Debt, Default, and Spending? Evidence from HARP." *American Economic Journal: Macroeconomics* 13(2): 254–291.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83(6): 2411–2451.

Adelino, Manuel, Kristopher Gerardi, and Barney Hartman-Glaser. 2019. "Are lemons sold first? Dynamic signaling in the mortgage market." *Journal of Financial Economics* 132(1): 1 – 25. ISSN 0304-405X. doi:https://doi.org/10.1016/j.jfineco.2018.09.005.

Available at http://www.sciencedirect.com/science/article/pii/S0304405X18302575.

Agarwal, Sumit, Brent W. Ambrose, and Vincent W. Yao. 2020. "Lender Steering in Residential Mortgage Markets." *Real Estate Economics* 48(2): 446–475. doi:https://doi.org/10.1111/1540-6229.12203.

Agarwal, Sumit, Gene Amromin, Souphala Chomsisengphet, Tim Landvoigt, Tomasz Piskorski, Amit Seru, and Vincent Yao. 2020a. "Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinancing Program." *Working paper*.

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao. 2017a. "Systematic Mistakes in the Mortgage Market and Lack of Financial Sophistication." *Journal of Financial Economics* 123(1): 42–58. ISSN 0304-405X.

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao. 2017b. "Systematic mistakes in the mortgage market and lack of financial sophistication." *Journal of Financial Economics* 123(1): 42–58.

Agarwal, Sumit, Souphala Chomsisengphet, Hua Kiefer, Leonard C. Kiefer, and Paolina C. Medina. 2020b. *Working paper*.

Agarwal, Sumit, John C Driscoll, and David I Laibson. 2013a. "Optimal Mortgage Refinancing: A Closed-Form Solution." *Journal of Money, Credit and Banking* 45(4): 591–622.

Agarwal, Sumit, John C. Driscoll, and David I. Laibson. 2013b. "Optimal Mortgage Refinancing: A Closed-Form Solution." *Journal of Money, Credit and Banking* 45(4): 591–622.

Agarwal, Sumit, John Grigsby, Ali Hortaçsu, Gregor Matvos, Amit Seru, and Vincent Yao. 2020c. "Searching for Approval." Working Paper 27341. National Bureau of Economic Research.

Agarwal, Sumit, Richard J. Rosen, and Vincent Yao. 2016. "Why Do Borrowers Make Mortgage Refinancing Mistakes?" *Management Science* 62(12): 3494–3509.

Akerlof, George A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84(3): 488–500.

Allen, Jason, Robert Clark, and Jean-François Houde. 2014. "The Effect of Mergers in Search Markets: Evidence from the Canadian Mortgage Industry." *American Economic Review* 104(10): 3365–96. doi:10.1257/aer.104.10.3365.

Allen, Jason, Robert Clark, and Jean-François Houde. 2019. "Search Frictions and Market Power in Negotiated-Price Markets." *Journal of Political Economy* 127(4): 1550–1598. doi:10.1086/701684.

Ambokar, Sumedh, and Kian Samaee. 2019. "Inaction, Search Costs, and Market Power in the US Mortgage Market." *Working Paper*.

Ambrose, Brent W, James Conklin, and Luis A Lopez. 2020. "Does Borrower and Broker Race Affect the Cost of Mortgage Credit?" *Review of Financial Studies (Forthcoming)*.

Amromin, Gene, Jennifer Huang, Clemens Sialm, and Edward Zhong. 2018. "Complex Mortgages." *Review of Finance* 22(6): 1975–2007.

Andersen, Steffen, John Y. Campbell, Kasper Meisner Nielsen, and Tarun Ramadorai. 2018. "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market." *Working Paper*.

Andersen, Steffen, John Y. Campbell, Kasper Meisner Nielsen, and Tarun Ramadorai. 2020a. "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market." *American Economic Review* 110(10): 3184–3230. doi:10.1257/aer.20180865.

Andersen, Steffen, John Y. Campbell, Kasper Meisner Nielsen, and Tarun Ramadorai. 2020b. "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market." *American Economic Review* 110(10): 3184–3230. doi:10.1257/aer.20180865.

Available at https://www.aeaweb.org/articles?id=10.1257/aer.20180865.

Andrews, Donald W. K., and Gustavo Soares. 2010. "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection." *Econometrica* 78(1): 119–157.

Archer, Wayne R., and David C. Ling. 1993. "Pricing Mortgage-Backed Securities: Integrating Optimal Call and Empirical Models of Prepayment." *Real Estate Economics* 21(4): 373–404.

Arcidiacono, Peter, Patrick Bayer, Federico A Bugni, and Jonathan James. 2013. "Approximating High-dimensional Dynamic Models: Sieve Value Function Iteration." In *Structural Econometric Models (Advances in Econometrics, Vol. 31)*, 45–95. Emerald Group Publishing Limited.

Arrow, Kenneth J., and Gerard Debreu. 1954. "Existence of an Equilibrium for a Competitive Economy." *Econometrica* 22(3): 265–290.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace. 2019a. "Consumer-Lending Discrimination in the FinTech Era." *Working Paper*.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace. 2019b. "Consumer-lending discrimination in the FinTech era." *Working paper*.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace. 2021. "Consumer-lending discrimination in the FinTech Era." *Journal of Financial Economics*.

Barwick, Panle Jia, and Parag A. Pathak. 2015. "The Costs of Free Entry: an Empirical Study of Real Estate Agents in Greater Boston." *RAND Journal of Economics* 46(1): 103–145. doi: https://doi.org/10.1111/1756-2171.12082.

Bell, Alex. 2019. "Job Amenities & Earnings Inequality." *Working Paper*.

Benetton, Matteo, Alessandro Gavazza, and Paolo Surico. 2020. "Inaction, Search Costs, and Market Power in the US Mortgage Market." *Working Paper*.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra. 2018a. "Regional Heterogeneity and the Refinancing Channel of Monetary Policy." *Quarterly Journal of Economics* 134(1): 109–183.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra. 2018b. "Regional Heterogeneity and the Refinancing Channel of Monetary Policy*." *The Quarterly Journal of Economics* 134(1): 109–183. ISSN 0033-5533. doi:10.1093/qje/qjy021.

Available at https://doi.org/10.1093/qje/qjy021.

Berger, David W., Konstantin Milbradt, Fabrice Tourre, and Joseph Vavra. 2020. "Mortgage Prepayment and Path-Dependent Effects of Monetary Policy." *Working paper*.

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan. 1994. "Race, redlining, and residential mortgage loan performance." *The Journal of Real Estate Finance and Economics* 9: 263–294.

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan. 1998. "Discrimination, Competition, and Loan Performance in FHA Mortgage Lending." *The Review of Economics and Statistics* 80(2): 241–250.

Bhutta, Neil, and Glenn B Canner. 2013. "Mortgage Market Conditions and Borrower Outcomes: Evidence from the 2012 HMDA data and Matched HMDA-Credit Record Data." *Federal Reserve Bulletin* 99(4): 1–58.

Bhutta, Neil, Andreas Fuster, and Aurel Hizmo. 2019. "Paying Too Much? Price Dispersion in the US Mortgage Market." *Working Paper*.

Bhutta, Neil, Andreas Fuster, and Aurel Hizmo. 2020. "Paying too much? Price dispersion in the US mortgage market." *Working Paper*.

Bhutta, Neil, and Aurel Hizmo. 2020. "Do Minorities Pay More for Mortgages?" *The Review of Financial Studies* 34(2): 763–789.

Bhutta, Neil, and Aurel Hizmo. 2021. "Do Minorities Pay More for Mortgages?" *The Review of Financial Studies* 34(2): 763–789. ISSN 0893-9454. doi:10.1093/rfs/hhaa047.

Available at https://doi.org/10.1093/rfs/hhaa047.

Bhutta, Neil, and Daniel Ringo. 2021. "The effect of interest rates on home buying: Evidence from a shock to mortgage insurance premiums." *Journal of Monetary Economics* 118: 195–211.

Black, Harold A., Thomas P. Boehm, and Ramon P. DeGennaro. 2003. "Is there discrimination in mortgage pricing? The case of overages." *Journal of Banking & Finance* 27(6): 1139 – 1165. ISSN 0378-4266. doi:https://doi.org/10.1016/S0378-4266(02)00250-9.

Black, Harold A, and Robert L Schweitzer. 1985. "A Canonical Analysis of Mortgage Lending Terms: Testing for Lending Discrimination at a Commercial Bank." *Urban Studies* 22(1): 13–19. ISSN 0042-0980.

Bocian, Debbie Gruenstein, Keith S. Ernst, and Wei Li. 2008. "Race, ethnicity and subprime home loan pricing." *Journal of Economics and Business* 60(1): 110 – 124. ISSN 0148-6195. doi: https://doi.org/10.1016/j.jeconbus.2007.10.001. Financing Community Reinvestment and Development.

Boehm, Thomas P., Paul D. Thistle, and Alan Schlottmann. 2006. "Rates and race: An analysis of racial disparities in mortgage rates." *Housing Policy Debate* 17(1): 109–149. doi:10.1080/10511482.2006.9521563.

Boudoukh, Jacob, Robert F Whitelaw, Matthew Richardson, and Richard Stanton. 1997. "Pricing mortgage-backed securities in a multifactor interest rate environment: A multivariate density estimation approach." *The Review of Financial Studies* 10(2): 405–446.

Boyarchenko, Nina, Andreas Fuster, and David O. Lucca. 2019. "Understanding Mortgage Spreads." *Review of Financial Studies* 32(10): 3799–3850.

Brueckner, Jan K. 1994. "Borrower Mobility, Adverse Selection, and Mortgage Points." *Journal of Financial Intermediation* 3(4): 416–441.

Bubb, Ryan, and Alex Kaufman. 2014. "Securitization and moral hazard: Evidence from credit score cutoff rules." *Journal of Monetary Economics* 63: 1–18.

Bundorf, M. Kate, Jonathan Levin, and Neale Mahoney. 2012. "Pricing and Welfare in Health Plan Choice." *American Economic Review* 102(7): 3214–3248.

Calza, Alessandro, Tommaso Monacelli, and Livio Stracca. 2013. "HOUSING FINANCE AND MONETARY POLICY." *Journal of the European Economic Association* 11(s1): 101–122. doi: 10.1111/j.1542-4774.2012.01095.x.

Available at https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1542-4774.2012.01095.x.

Campbell, John Y. 2006a. "Household finance." *The journal of finance* 61(4): 1553–1604.

Campbell, John Y. 2006b. "Household Finance, Presidential Address to the American Finance Association." *Journal of Finance* LXI(4): 1553–1604.

Campbell, John Y., Nuno Clara, and João F. Cocco. 2021. "Structuring Mortgages for Macroeconomic Stability." *Journal of Finance* 76: 2525–2576.

Campbell, John Y., and Joao F. Cocco. 2003. "Household Risk Management and Optimal Mortgage Choice." *Quarterly Journal of Economics* 118(4): 1449–1494.

Campbell, John Y., and Joao F. Cocco. 2015. "A Model of Mortgage Default." *Journal of Finance* 70(4): 1495–1554.

Canner, Glenn B., Stuart A. Gabriel, and J. Michael Woolley. 1991. "Race, Default Risk and Mortgage Lending: A Study of the FHA and Conventional Loan Markets." *Southern Economic Journal* 58(1): 249–262. ISSN 00384038.

Available at http://www.jstor.org/stable/1060046.

Chang, Yan, and Abdullah Yavas. 2009. "Do borrowers make rational choices on points and refinancing?" *Real Estate Economics* 37(4): 635–658.

Chari, V. V., and Ravi Jagannathan. 1989. "Adverse Selection in a Model of Real Estate Lending." *Journal of Finance* 44(2): 499–508.

Chen, Hui, Michael Michaux, and Nikolai Roussanov. 2020. "Houses as ATMs: Mortgage Refinancing and Macroeconomic Uncertainty." *Journal of Finance* 75(1): 323–375.

Cheng, Ping, Zhenguo Lin, and Yingchun Liu. 2015. "Racial discrepancy in mortgage interest rates." *The Journal of Real Estate Finance and Economics* 51(1): 101–120.

Clapp, John M, Gerson M Goldberg, John P Harding, and Michael LaCour-Little. 2001. "Movers and shuckers: interdependent prepayment decisions." *Real estate economics* 29(3): 411–450.

Cloyne, James, Clodomiro Ferreira, and Paolo Surico. 2019. "Monetary Policy when Households have Debt: New Evidence on the Transmission Mechanism." *The Review of Economic Studies* 87(1): 102–129. ISSN 0034-6527. doi:10.1093/restud/rdy074.

Available at https://doi.org/10.1093/restud/rdy074.

Corbae, Dean, and Erwan Quintin. 2015. "Leverage and the Foreclosure Crisis." *Journal of Political Economy* 123(1): 1–65.

Courchane, Marsha, and David Nickerson. 1997. "Discrimination Resulting from Overage Practices." *Journal of Financial Services Research* 11(1): 133–151.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross. 1985. "A Theory of the Term Structure of Interest Rates." *Econometrica* 53(2): 385–407.

Cunningham, Chris, Kristopher Gerardi, and Lily Shen. 2021. "The Double Trigger for Mortgage Default: Evidence from the Fracking Boom." *Management Science* 67(6): 3943–3964.

Cutler, David M., and Sarah J. Reber. 1998. "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection." *Quarterly Journal of Economics* 113(2): 433–466.

Deng, Yongheng, and Stuart Gabriel. 2006. "Risk-based pricing and the enhancement of mortgage credit availability among underserved and higher credit-risk populations." *Journal of money, Credit and Banking* 1431–1460.

Deng, Yongheng, and John M Quigley. 2012. "Woodhead behavior and the pricing of residential mortgages." *NUS Institute of Real Estate Studies Working Paper Series IRES2012-025*.

Deng, Yongheng, John M. Quigley, and Robert Van Order. 2000a. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options." *Econometrica* 68(2): 275–307.

Deng, Yongheng, John M Quigley, and Robert Van Order. 2000b. "Mortgage terminations, heterogeneity and the exercise of mortgage options." *Econometrica* 68(2): 275–307.

Di Maggio, Marco, Amir Kermani, Benjamin J Keys, Tomasz Piskorski, Rodney Ramcharan, Amit Seru, and Vincent Yao. 2017. "Interest rate pass-through: Mortgage rates, household consumption, and voluntary deleveraging." *American Economic Review* 107(11): 3550–88. doi:https://doi.org/10.1257/aer.20141313.

Di Maggio, Marco, Amir Kermani, and Christopher J Palmer. 2020. "How Quantitative Easing works: Evidence on the Refinancing Channel." *Review of Economic Studies* 87(3): 1498–1528.

Diep, Peter, Andrea L Eisfeldt, and Scott Richardson. 2021. "The cross section of MBS returns." *The Journal of Finance* 76(5): 2093–2151.

Ding, Lei. 2017. "Borrower credit access and credit performance after loan modifications." *Empirical Economics* 52(3): 977–1005.

Eichenbaum, Martin, Sérgio Rebelo, and Arlene Wong. 2018. "State Dependent Effects of Monetary Policy: the Refinancing Channel." CEPR Discussion Papers 13223. C.E.P.R. Discussion Papers.

Einav, Liran, and Amy Finkelstein. 2011. "Selection in Insurance Markets: Theory and Empirics in Pictures." *Journal of Economic Perspectives* 25(1): 115–138. doi:10.1257/jep.25.1.115.

Einav, Liran, Amy Finkelstein, and Mark R. Cullen. 2010. "Estimating Welfare in Insurance Markets Using Variation in Prices." *Quarterly Journal of Economics* 125(3): 877–921.

Einav, Liran, Amy Finkelstein, Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen. 2013. "Selection on Moral Hazard in Health Insurance." *American Economic Review* 103(1): 178–219.

Ellis, Andrew, and Yusufcan Masatlioglu. 2021. "Choice with Endogenous Categorization." *The Review of Economic Studies*.

Fang, Zheng, and Andres Santos. 2018. "Inference on Directionally Differentiable Functions." *The Review of Economic Studies* 86(1): 377–412.

Firestone, Simon, Robert Van Order, and Peter Zorn. 2007. "The Performance of Low-Income and Minority Mortgages." *Real Estate Economics* 35(4): 479–504.

Fisher, Jack, Alessandro Gavazza, Lu Liu, Tarun Ramadorai, and Jagdish Tripathy. 2021. "Refinancing Cross-Subsidies in the UK Mortgage Market." *Working Paper*.

Fusari, Nicola, Wei Li, Haoyang Liu, and Zhaogang Song. 2020. "Asset pricing with cohort-based trading in mbs markets." *FRB of New York Staff Report* (931).

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther. 2018. "Predictably unequal? The effects of machine learning on credit markets." *Working paper*.

Fuster, Andreas, Aurel Hizmo, Lauren Lambie-Hanson, James Vickery, and Paul Willen. 2021. "How Resilient is Mortgage Credit Supply? Evidence from the COVID-19 Pandemic." *Working Paper*.

Fuster, Andreas, Stephanie H. Lo, and Paul S. Willen. 2017. "The Time-Varying Price of Financial Intermediation in the Mortgage Market." *Working paper*.

Fuster, Andreas, Stephanie H. Lo, and Paul S. Willen. 2019. "The Time-Varying Price of Financial Intermediation in the Mortgage Market." *Working Paper*.

Fuster, Andreas, and Paul Willen. 2010. "$1.25 trillion is still real money: Some facts about the effects of the federal reserve's mortgage market investments." *FRB of Boston Public Policy Discussion Paper* (10-4).

Gabaix, Xavier, Arvind Krishnamurthy, and Olivier Vigneron. 2007. "Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market." *The Journal of Finance* 62(2): 557–595. doi:10.1111/j.1540-6261.2007.01217.x.

Gabaix, Xavier, and David Laibson. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*.

Gal, Tomas, and Harvey J Greenberg. 2012. *Advances in sensitivity analysis and parametric programming*, vol. 6. Springer Science & Business Media.

Galichon, Alfred. 2016. *Optimal transport methods in economics*. Princeton University Press.

Gerardi, Kristopher, Paul Willen, and David Zhang. 2020. "Mortgage Prepayment, Race, and Monetary Policy." *Working paper*.

Gerardi, Kristopher, Paul Willen, and David Hao Zhang. 2021. "Prepayment, Race, and Monetary Policy." *Working Paper*.

Ghent, Andra C., Rubén Hernández-Murillo, and Michael T. Owyang. 2014. "Differences in subprime loan pricing across races and neighborhoods." *Regional Science and Urban Economics* 48: 199 – 215. ISSN 0166-0462. doi:https://doi.org/10.1016/j.regsciurbeco.2014.07.006.

Glaeser, Edward L, and Charles G Nathanson. 2017. "An Extrapolative Model of House Price Dynamics." *Journal of Financial Economics* 126(1): 147–170.

Goldberger, Arthur S. 1984. "Reverse Regression and Salary Discrimination." *The Journal of Human Resources* 19(3): 293–318.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104(4): 1091–1119.

Green, Richard K., and Michael LaCour-Little. 1999a. "Some Truths about Ostriches: Who Doesn't Prepay Their Mortgages and Why They Don't." *Journal of Housing Economics* 8(3): 233–248.

Green, Richard K, and Michael LaCour-Little. 1999b. "Some truths about ostriches: Who doesn't prepay their mortgages and why they don't." *Journal of Housing Economics* 8(3): 233–248.

Gul, Faruk, and Wolfgang Pesendorfer. 2001. "Temptation and Self-Control." *Econometrica* 69(6): 1403–1435.

Guren, Adam M, Arvind Krishnamurthy, and Timothy J McQuade. 2021. "Mortgage Design in an Equilibrium Model of the Housing Market." *Journal of Finance* 76(1): 113–168.

Hackmann, Martin B., Jonathan T. Kolstad, and Amanda E. Kowalski. 2015. "Adverse Selection and an Individual Mandate: When Theory Meets Practice." *American Economic Review* 105(3): 1030–66.

Handel, Ben, Igal Hendel, and Michael D. Whinston. 2015. "adverse selection and Inertia in Health Exchanges: Adverse Selection versus Reclassification Risk." *Econometrica* 83(4): 1261–1313.

Handel, Benjamin R. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review* 103(7): 2643–2682.

Handel, Benjamin R., Jonathan T. Kolstad, and Johannes Spinnewijn. 2019. "Information Frictions and Adverse Selection: Policy Interventions in Health Insurance Markets." *Review of Economics and Statistics* 101(2): 326–340.

Ho, Kate, and Robin S Lee. 2021. "Health Insurance Menu Design for Large Employers." *Working Paper*.

Hong, Han, and Jessie Li. 2020. "The numerical bootstrap." *Annals of Statistics* 48(1): 397–412.

Horrace, William C., and Ronald L. Oaxaca. 2006. "Results on the bias and inconsistency of ordinary least squares for the linear probability model." *Economics Letters* 90(3): 321 – 327. ISSN 0165-1765. doi:https://doi.org/10.1016/j.econlet.2005.08.024.

Available at http://www.sciencedirect.com/science/article/pii/S0165176505003150.

Hsieh, Yu-Wei, Xiaoxia Shi, and Matthew Shum. 2021. "Inference on estimators defined by mathematical programming." *Journal of Econometrics*.

Hurst, Erik, Benjamin J. Keys, Amit Seru, and Joseph Vavra. 2016. "Regional Redistribution through the US Mortgage Market." *American Economic Review* 106(10): 2982–3028.

Hwang, Hae-shin, W Robert Reed, and Carlton Hubbard. 1992. "Compensating wage differentials and unobserved productivity." *Journal of Political Economy* 100(4): 835–858.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil. 2013. "Securitization and Loan Performance: Ex Ante and Ex Post Relations in the Mortgage Market." *The Review of Financial Studies* 27(2): 454–483. ISSN 0893-9454. doi:10.1093/rfs/hht073.

Available at https://doi.org/10.1093/rfs/hht073.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil. 2014. "Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency." *The Review of Economics and Statistics* 96(1): 1–18. doi:10.1162/REST_a_00387.

Available at https://doi.org/10.1162/REST_a_00387.

Johnson, Eric J, Stephan Meier, and Olivier Toubia. 2018a. "What's the Catch? Suspicion of Bank Motives and Sluggish Refinancing." *Review of Financial Studies* 32(2): 467–495.

Johnson, Eric J, Stephan Meier, and Olivier Toubia. 2018b. "What's the Catch? Suspicion of Bank Motives and Sluggish Refinancing." *The Review of Financial Studies* 32(2): 467–495. ISSN 0893-9454. doi:10.1093/rfs/hhy061.

Available at https://doi.org/10.1093/rfs/hhy061.

Kau, James B, Lu Fang, and Henry J Munneke. 2019a. "An unintended consequence of mortgage financing regulation–a racial disparity." *The Journal of Real Estate Finance and Economics* 59(4): 549–588.

Kau, James B, Lu Fang, and Henry J Munneke. 2019b. "An unintended consequence of mortgage financing regulation–a racial disparity." *The Journal of Real Estate Finance and Economics* 59(4): 549–588.

Kau, James B, Donald C Keenan, and Henry J Munneke. 2012. "Racial discrimination and mortgage lending." *The Journal of Real Estate Finance and Economics* 45(2): 289–304.

Kaufman, Alex. 2014. "The Influence of Fannie and Freddie on Mortgage Loan Terms." *Real Estate Economics* 42(2): 472–496. doi:10.1111/1540-6229.12030.

Available at https://onlinelibrary.wiley.com/doi/abs/10.1111/1540-6229.12030.

Keane, Michael P., and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy* 105(3): 473–522.

Kelly, Austin. 1995. "Racial and ethnic disparities in mortgage prepayment." *Journal of Housing Economics* 4(4): 350–372.

Keys, Benjamin J., Devin G. Pope, and Jaren C. Pope. 2016a. "Failure to refinance." *Journal of Financial Economics* 122(3): 482–499.

Keys, Benjamin J, Devin G Pope, and Jaren C Pope. 2016b. "Failure to refinance." *Journal of Financial Economics* 122(3): 482–499.

Keys, Benjamin J., Amit Seru, and Vikrant Vig. 2012. "Lender Screening and the Role of Securitization: Evidence from Prime and Subprime Mortgage Markets." *The Review of Financial Studies* 25(7): 2071–2108. ISSN 0893-9454. doi:10.1093/rfs/hhs059.

Available at https://doi.org/10.1093/rfs/hhs059.

Kladivko, Kamil. 2021. *Maximum Likelihood Estimation of the Cox-Ingersoll-Ross Process: the Matlab Implementation*. MATLAB Central File Exchange.

Lambie-Hanson, Lauren, and Carolina Reid. 2018. "Stuck in Subprime? Examining the Barriers to Refinancing Mortgage Debt." *Housing Policy Debate* 28(5): 770–796.

LeRoy, Stephen F. 1996. "Mortgage Valuation Under Optimal Prepayment." *Review of Financial Studies* 9(3): 817–844.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies* 70(2): 317–341.

Maturana, Gonzalo, and Jordan Nickerson. 2019. "Teachers Teaching Teachers: The Role of Workplace Peer Effects in Financial Decisions." *The Review of Financial Studies* 32(10): 3920–3957.

Mayer, Chris, Tomasz Piskorski, and Alexei Tchistyi. 2013. "The Inefficiency of Refinancing: Why Prepayment Penalties are Good for Risky Borrowers." *Journal of Financial Economics* 107(3): 694–714.

McCloskey, Adam. 2017. "Bonferroni-based size-correction for nonstandard testing problems." *Journal of Econometrics* 200(1): 17 – 35.

McConnell, John J., and Manoj Singh. 1994. "Rational Prepayments and the Valuation of Collateralized Mortgage Obligations." *Journal of Finance* 49(3): 891–921.

Montiel Olea, José Luis, and Mikkel Plagborg-Møller. 2019. "Simultaneous confidence bands: Theory, implementation, and an application to SVARs." *Journal of Applied Econometrics* 34(1): 1–17.

Munnell, Alicia H., Geoffrey M. B. Tootell, Lynn E. Browne, and James McEneaney. 1996. "Mortgage Lending in Boston: Interpreting HMDA Data." *The American Economic Review* 86(1): 25–53. ISSN 00028282.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64(6): 1263–1297.

Piskorski, Tomasz, Amit Seru, and Vikrant Vig. 2010. "Securitization and distressed loan renegotiation: Evidence from the subprime mortgage crisis." *Journal of Financial Economics* 97(3): 369 – 397. ISSN 0304-405X. doi:https://doi.org/10.1016/j.jfineco.2010.04.003. The 2007-8 financial crisis: Lessons from corporate finance.

Available at http://www.sciencedirect.com/science/article/pii/S0304405X10000929.

Politis, Dimitris N., and Joseph P. Romano. 1994. "Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions." *Annals of Statistics* 22(4): 2031–2050.

Romano, Joseph P., Azeem M. Shaikh, and Michael Wolf. 2014. "A Practical Two-Step Method for Testing Moment Inequalities." *Econometrica* 82(5): 1979–2002.

Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90(4): 629–649.

Santambrogio, Filippo. 2015. "Optimal transport for applied mathematicians." *Birkäuser, NY* 55(58-63): 94.

Schwartz, Allie. 2006. "Household refinancing behavior in fixed rate mortgages." *unpublished paper, Harvard University*.

Schwartz, Eduardo S., and Walter N. Torous. 1989. "Prepayment and the Valuation of Mortgage-Backed Securities." *Journal of Finance* 44(2): 375–392.

Shapiro, Alexander. 1990. "On concepts of directional differentiability." *Journal of Optimization Theory and Applications* 66: 477–487.

Shapiro, Alexander. 1991. "Asymptotic analysis of stochastic programs." *Annals of Operations Research* 30: 169–186.

Skanderson, David, Ivan Darius, and Tammy Butler. 2014. "Containing Fair Lending Risk." *Mortgage Banking*.

Sommerfeld, Max, and Axel Munk. 2018. "Inference for empirical Wasserstein distances on finite spaces." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 80(1): 219–238.

Song, Zhaogang, and Haoxiang Zhu. 2018. "Mortgage Dollar Roll." *The Review of Financial Studies* 32(8): 2955–2996.

Stanton, Richard. 1995. "Rational Prepayment and the Valuation of Mortgage-Backed Securities." *Review of Financial Studies* 8(3): 677–708.

Stanton, Richard, and Nancy Wallace. 2003. "Mortgage Choice: What's the Point?" *Real Estate Economics* 26: 173–205.

Sunderam, Adi, and David Scharfstein. "Market Power in Mortgage Lending and the Transmission of Monetary Policy." *Working Paper* 2017.

Tameling, Carla, Max Sommerfeld, and Axel Munk. 2019. "Empirical optimal transport on countable metric spaces: Distributional limits and statistical applications." *The Annals of Applied Probability* 29(5): 2744–2781.

Tootell, Geoffrey M. B. 1996. "Redlining in Boston: Do Mortgage Lenders Discriminate Against Neighborhoods?*." *Quarterly Journal of Economics* 111(4): 1049–1079. ISSN 0033-5533. doi: 10.2307/2946707.

Vickery, James I, and Joshua Wright. 2013. "TBA Trading and Liquidity in the Agency MBS Market." *Economic Policy Review* 19(1).

Villani, Cédric. 2008. *Optimal transport: old and new*, vol. 338. Springer Science & Business Media.

Willen, Paul S., and David Hao Zhang. 2021. "Do Lenders Still Discriminate? A Robust Methodology for Detecting Differences in Menus." *Working Paper*.

Wilson, Charles. 1977. "A Model of Insurance Markets with Incomplete Information." *Journal of Economic Theory* 16(2): 167–207.

Wong, Arlene. 2019. "Refinancing and The Transmission of Monetary Policy to Consumption." *Working Paper*.

Woodward, Susan E. 2008. "A Study of Closing Costs for FHA Mortgages." *U.S. Department of Housing and Urban Development Report*.

Woodward, Susan E., and Robert E. Hall. 2012. "Diagnosing Consumer Confusion and Sub-Optimal Shopping Effort: Theory and Mortgage-Market Evidence." *American Economic Review* 102(7): 3249–3276.

Zhang, David Hao, and Paul S Willen. 2021. "Do Lenders Still Discriminate? A Robust Methodology for Detecting Differences in Menus." *Working Paper*.