Closing Costs, Refinancing, and Inefficiencies in the Mortgage Market

David Zhang*

October 12, 2023

Abstract

I use a structural model to quantify the cross-subsidization in the US mortgage market due to heterogeneous borrower refinancing tendencies. The presence of borrowers with high refinancing inertia reduces mortgage interest rates by 0.21–0.97 percentage points, particularly on lower upfront closing cost mortgages with more valuable refinancing options. As a result, actively refinancing borrowers refinance excessively relative to a no cross-subsidization benchmark, generating deadweight administrative costs. Adding more of the price of mortgage origination to the loan balance reduces transfers and increases surplus. Automatically refinancing mortgages also produce efficiency benefits but at the cost of a higher initial rate.

*I thank John Campbell, Ed Glaeser, Adi Sunderam, Ariel Pakes, and Robin Lee for continuous advice on this paper. I also thank João Cocco, Matthew Curtis, Mark Egan, Xiang Fang, Jiacheng Feng, Xavier Gabaix, Daniel Green, Daniel Greenwald (discussant), Shiyang Huang, Jun Ishii, Larry Katz, Amir Kermani, Alex Kopytov, David Laibson, Feng Liu, David Low (discussant), Doug McManus (discussant), Konstantin Milbradt, Fabrice Tourre, Pascal Noel, Henry Overman (discussant), Frank Pinter, David Scharfstein, Jeremy Stein, Dragon Tang, Audrey Tiew, Fabrice Tourre, Boris Vallee, Nancy Wallace, Jeffrey Wang, Paul Willen, Susan Woodward (discussant), Ron Yang, and seminar/conference participants at CityUHK, HKU, TAMU, SMU, BC, UofT, Purdue, Rice, Fed Board, Minnesota, CU Boulder, Penn State, 2022 Chicago Booth Household Finance Conference, 2022 NBER Summer Institute Real Estate, 2022 North American Summer Meeting of the Econometric Society, 2022 Asian Meeting of the Econometric Society, 2022 CFPB Research Conference, 2023 ASSA/AREUEA Meeting, 2023 UCLA/SF Fed Conference, 2023 Boulder Consumer Research Conference, and 2023 WFA meeting for valuable comments which greatly improved the paper. All errors are mine. This research was conducted while the author was a Visiting Fellow at the Federal Reserve Bank of Boston. The views expressed here are solely those of the author and not those of the Federal Reserve Bank of Boston nor the Federal Reserve System. David Zhang, dzhang@rice.edu, is at Rice University, Jones Graduate School of Business, 1900 Rice Boulevard, Houston, TX 77005.

1 Introduction

In the United States, many borrowers are slow to refinance or never refinance their mortgages when interest rates fall, while others are more quick to do so. This heterogeneity in refinancing tendencies has long been recognized as an important friction in household finance.[1] In this paper, I use a structural model to quantify the distributional and efficiency implications of the heterogeneous consumer refinancing tendencies.

A structural model of the heterogeneous borrower refinancing tendencies in this context is important for two reasons. First, to the extent that refinancing differences contribute to wealth inequality, it is important to quantify this inequality in dollar value terms. Second, a novel insight of my model is that refinancing heterogeneity under the predominant US mortgage contract design generates deadweight losses, due to excessive refinancing and inefficient contract selection in terms of upfront closing costs, which adds to the economic rationale for considering alternative contract designs.

While the study of the impact of heterogeneous refinancing tendencies in the US mortgage market is important, it is also difficult. First, non-underwriting variables, such as updated credit scores and mark-to-market loan to value (LTV) ratios, help determine borrowers' incentive to refinance and are unavailable in publicly available loan performance datasets from Fannie Mae and Freddie Mac. Second, a first order factor in models of optimal refinancing in the United States is the upfront closing costs, which pin down the interest rate threshold at which borrowers optimally refinance, but these upfront costs are endogenous when borrowers choose how much to pay for their mortgage origination upfront versus through the rate. Third, the lender's interest rate spreads across contract choices are determined by the investors' expectations of moving and refinancing, which is a complicated dynamic problem. Fourth, computing the implications of different refinancing strategies within each borrower

[1] See, e.g., Schwartz and Torous (1989), Archer and Ling (1993), McConnell and Singh (1994), Stanton (1995), Green and LaCour-Little (1999), Campbell (2006), Agarwal, Rosen, and Yao (2016), Keys, Pope, and Pope (2016), Johnson, Meier, and Toubia (2018), Andersen, Campbell, Nielsen, and Ramadorai (2020), Gerardi, Willen, and Zhang (2023), and Byrne, Devine, King, McCarthy, and Palmer (2023).

type also involves solving a complicated dynamic problem. Through a mix of novel data and computationally intensive modelling, my paper addresses these difficulties.

In particular, my paper identifies two main channels through which heterogeneous refinancing inertia affects on the US mortgage market. First, the existence of borrowers with high inertia implies that lenders can afford to charge lower interest rates upfront. This effect reflects a cross-subsidization from slow-to-refinance borrowers to quicker-to-refinance ones, which I quantify. Second, the non-uniformity of the interest rate reduction effect across upfront closing cost choices distorts borrower contract choices, which generates economic inefficiencies through excessive refinancing incentives.

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. Second, they can raise the interest rate on the mortgage, holding fixed its principal balance and then recovering their costs from the secondary market. Most lenders offer a menu of rate and upfront closing cost options to prospective borrowers through a choice of how many points to pay to or receive from the lender.[2] Borrowers therefore can opt for, through their choice of points, a mortgage with a lower rate and higher closing costs or one with a higher rate and lower closing costs. Using data linking points and borrower refinancing and prepayment behavior, I begin my empirical analysis with a set of motivating facts regarding their correlation, showing evidence of both selection and pooling of borrower types by point choices. Then, I show that most of the lenders' price of mortgage origination ends up being charged through the rate, as secondary marketing revenue, rather than upfront, which substantially increases the cross-subsidization and generates excessive refinancing incentives.

To quantify the size of the cross-subsidy by borrower refinancing tendencies and study its efficiency consequences, I develop a structural equilibrium model that captures borrower heterogeneity in refinancing and moving tendencies while endogenizing borrower choices

[2] In the industry, each mortgage point refers to 1% of the loan amount that borrowers pay upfront. Positive points in the form of discount points increase the upfront closing costs while reducing the interest rate, while negative points the form of lender credit reduce upfront closing costs while increasing the interest rate.

of upfront closing costs. I embed the time and state dependence of borrower refinancing behavior described by Andersen et al. (2020) into a heterogenous agent life-cycle model that gives welfare estimates interpretable in dollar-equivalent terms. A zero-profit condition with a Monte Carlo model of the pricing of mortgage-backed securities pins down the supply side.

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 their being able to refinance; (ii) moving or exogenous prepayment probabilities; (iii) discount factors; and (iv) liquid wealth and income. The time-varying ability to refinance and the refinancing hassle cost are separately identified from borrower delays in refinancing after their refinancing thresholds have been reached (Andersen et al., 2020) and could reflect both demand-side differences in preferences and any supply-side-driven differential costs to refinance coming from potential discrimination in the market. The variance in moving or exogenous prepayment expectations is identified by the correlation between sluggish refinancing and subsequent exogenous prepayment behavior. Discount factors are identified based on choices of upfront closing costs. Borrowers' liquid wealth and income are calibrated using data from the Survey of Consumer Finances (SCF) and the Home Mortgage Disclosure Act (HMDA).

Three main conclusions emerge. 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 can always refinance at a hassle cost of $200 and holding lender revenue fixed, a mortgage with a one percent upfront closing cost carries a 0.97 percentage points 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 0.21 percentage points. The reason for the larger cross-subsidization of lower upfront closing cost mortgages is that, from the perspective of the lender, slow-to-refinance refinancing borrowers overpay for closing costs when they pay them through the rate because they keep paying a higher rate longer, but continue to select these contracts likely for reasons such as a lack of liquidity or financial sophistication.

A key advantage of my approach is that I can quantify the consequences of this cross-subsidization in the population. I estimate the model using maximum likelihood on a novel dataset linking borrowers' upfront closing cost choices to their subsequent prepayments in a sample of conforming mortgages, and, as my second conclusion, I find that the cross-subsidization of mortgage closing costs generates large differences in payoffs between borrowers, up to \$8,171 of the loan amount in present value terms between a typical nonrefinancing borrower and an actively refinancing borrower with a new purchase origination in the 2013–2019 period in expectation. Black borrowers are particularly hurt in the pooling equilibrium, with an expected loss of \$1,983 per new purchase origination over the period.

As my third conclusion, I show that the fact that most of the price of mortgage origination ends up being added to the rate of the mortgage substantially increases refinancing cross-subsidies and generates welfare losses to the tune of \$8 billion per year. I investigate this effect under an alternative contract design where the portion of the borrowers' price of mortgage origination added to the rate are instead added to the mortgage balance, holding total borrower upfront payment fixed. There, I find an increase in average borrower utility of \$1,461 per borrower. Non-refinancing borrowers recover over three quarters of their cross-subsidies paid, and, while the most actively refinancing borrowers become worse off, they are still better off than they would be if they have not received any cross-subsidization. The benefit of this alternative contract design comes from eliminating the cross-subsidization of lower upfront closing cost mortgages and thereby making all borrowers pay for their own price of mortgage origination.

Using the model, I also study the case of automatically refinancing mortgages. This contract eliminates the cross-subsidization between borrowers with different refinancing speeds and, despite a slightly higher initial mortgage interest rate due to lower MBS values, leads to an increase in average borrower utility of \$1,362 per borrower. The drawback of a higher initial rate in my model is particularly strong for liquidity constrained borrowers, however, and automatically refinancing mortgages therefore has less desirable distributional properties than the first contract. Taken as a whole, my results suggest that the equity-efficiency trade-off is not binding in the US mortgage context: it is possible to reduce refinancing inequality while increasing total welfare.

My model generates cross-sectional in borrower refinancing behavior through heterogeneous borrower-specific refinancing costs that could come from either the demand or supply side and is consistent with all borrowers being rational. Nevertheless, if one instead views the slow-to-refinance borrowers as not understanding the true cost of a higher interest rate, my work can also be interpreted as an empirical model of a shrouded equilibrium like that of Gabaix and Laibson (2006). In this case, the alternative contract designs I consider can be viewed as unravelling the shrouded equilibrium by removing discretion in either contract choice or refinancing.

My paper is primarily related to the literature on borrower heterogeneity in mortgage refinancing. Many papers document large borrower heterogeneity, conditional on the interest rate savings available, including Archer and Ling (1993), McConnell and Singh (1994), Stanton (1995), Deng, Quigley, and Van Order (2000), Agarwal, Rosen, and Yao (2016), Keys, Pope, and Pope (2016), Johnson, Meier, and Toubia (2018), Andersen et al. (2020), Beraja, Fuster, Hurst, and Vavra (2018), Ambokar and Samaee (2019), Belgibayeva, Bono, Bracke, Cocco, and Majer (2020), Gerardi, Willen, and Zhang (2023), and Byrne et al. (2023). I use a structural life-cycle model to quantify the cross-subsidization across borrowers with different refinancing tendencies in market equilibrium and study its efficiency implications under endogenous upfront closing cost choices.

Methodologically, my paper also contributes to the literature on life-cycle models of mortgage refinancing. This includes Campbell and Cocco (2003), Corbae and Quintin (2015), Campbell and Cocco (2015), Eichenbaum, Rebelo, and Wong (2018), Belgibayeva et al. (2020), Campbell, Clara, and Cocco (2021), Guren, Krishnamurthy, and McQuade (2021), Boar, Gorea, and Midrigan (2021), Liu (2022), and MacGee and Yao (2023). Of these papers, none incorporate unobserved cross-sectional heterogeneity in refinancing costs. My model does incorporate the role of time-and-state dependent unobserved heterogeneity as emphasized in Andersen et al. (2020) and Byrne et al. (2023), and considers the distributional and efficiency consequences of this heterogeneity under the prevailing US mortgage contract design.

In complementary work, Fisher, Gavazza, Liu, Ramadorai, and Tripathy (2023) and Berger, Milbradt, Tourre, and Vavra (2023) which uses structural models to study refinancing heterogeneity and cross-subsidization in the UK and US mortgage market, respectively. Neither studies the inefficiencies generated by this cross-subsidization due to distortions in the upfront closing cost choices of quick to refinance borrowers, which is the main conceptual contribution of this paper. I show that these inefficiencies are economically important and presents a reason to consider alternative contract designs beyond redistribution. Furthermore, I compute results by race and ethnicity and show that their expected loss under the current US system relative to a no cross-subsidization benchmark is sizable even in an ex ante sense.

In particular, while Berger et al. (2023) focuses on a heterogeneous agent framework for counterfactual MBS prices, I derive novel results about cross-subsidization and its efficiency consequences using a richer life-cycle demand side with endogenous upfront closing costs. Heterogeneous agent counterfactual MBS prices has limited application in my setting because I am able to estimate the distribution of borrower types conditional on the observed interest rate process and the counterfactual of automatically refinancing mortgages via single-agent backward induction. The counterfactual of switching the portion of the price of mortgage origination added to the rate to added to the balance does involve heterogeneous agents, however, and I assume in that counterfactual that market interest rates shift by a level constant. The effect of such an assumption is likely small since mortgages are originated at par in the counterfactual so prepayment only has a small effect on valuations.[3] The benefit of my model, in addition to endogenizing the choices of upfront costs and quantifying the welfare

[3] As described in Section 7.1, the level shift in interest rates needed for lender break-even is only 1.9 basis points.

effects inefficient refinancing in the US mortgage market equilibrium, is to capture richer implications of counterfactual designs by incorporating borrower heterogeneity in liquidity across demographic groups.

In terms of institution, my paper relates to a growing literature on borrower choices of mortgage upfront closing costs, which are also called points. In pioneering work, 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 the selection effect explored in these theories seriously in a quantitative manner. I find evidence of both selection and pooling, which are incorporated in my estimates of the welfare implications of refinancing heterogeneity in the United States. Chari and Jagannathan (1989) studies the role of insurance to income shocks for the institution of mortgage points, which I also incorporate in my model. Empirical work on consumer behavior with mortgage points includes Woodward and Hall (2012) and Mota, Palim, and Woodward (2022), who document how points may lead to suboptimal shopping; Agarwal, Ben-David, and Yao (2017), who show that many borrowers pay too much in points, given their predicted refinancing propensities. Since both of my alternative contract designs eliminate the choice of upfront closing costs and reduce price competition between lenders to a single dimension, this literature suggests that these designs may have benefits from a shopping and financial mistakes perspective.

The rest of this paper is structured as follows. Section 2 presents the background about the upfront closing cost and interest rate trade-off. Section 3 describes the data used. Section 4 presents motivating facts. Section 5 presents my model and simulation results. Section 6 presents estimation results. Section 7 describes the counterfactual analyses. Section 8 concludes.

2 Background

US borrowers face a choice between a mortgage with a higher interest rate and a lower upfront closing cost or one with a lower interest rate and a higher upfront closing cost. Mortgage points paid to the lender, which is a component of the upfront closing costs, is the way through which borrowers can make this choice. The number of mortgage points paid can be positive or negative and affects the interest rate the borrower pays. Negative points, where the lender pays the borrower a credit for taking on a higher rate, are available until the borrower pays zero in upfront closing costs.[4] I illustrate this choice in Figure 1, which shows a series of options for rates and points options from a lender ratesheet. The first column of the table in Figure 1 shows the choices of interest rates available to a borrower, while the 15-day, 30-day, and 45-day columns show the corresponding number of points that borrowers would have to pay to receive the rate once the loan is originated within the given lock period. A rate is "locked" when a lender commits the given terms for a mortgage originated within the stated lock period of, for example, 15, 30, or 45 days. In particular, Figure 1 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. Appendix Figure A.1 shows an example of how borrowers were shown a series of rate and upfront closing cost choices from a price comparison website.

Lenders generate revenue from both upfront closing costs paid by the borrower (including points, application fees, and other upfront fees) and from selling the mortgage on the secondary market for an amount greater than the original loan balance. To be more precise, focusing on the setting where lenders are selling the mortgages they originate on the

[4] Negative points are also sometimes called "lender credit."

secondary market,[5] I decompose lenders' total origination revenue from making a loan as:

$$\underbrace{\text{lender origination revenue}}_{\text{price of mortgage origination}} = \underbrace{\text{upfront closing costs}}_{\text{paid upfront}} + \underbrace{\text{secondary marketing income}(c)}_{\text{added into rate}} \quad (1)$$

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. Mortgages with higher interest rates tend to be more valuable on the secondary market, and originating a mortgage with a high enough interest rate generates positive secondary market income. Therefore the lender's price of mortgage origination may either be paid upfront, through upfront closing costs, or through the interest rate c, from secondary market income. To illustrate what the secondary market income as a function of interest rates might look like on a given day, Appendix Figure A.2 plots the secondary market value of mortgages based on MBS TBA prices as a percentage of the loan amount at various interest rates on January 2, 2014. The TBA market is a highly liquid market where most MBS are traded and is described in more detail by Vickery and Wright (2013).

A stylized fact in the US mortgage industry is that most of the lender origination revenue in Equation (1) comes from secondary market income rather than upfront closing costs. The implication is then that new production mortgage backed securities trade significantly above par. I explore this fact in more detail in Appendix A.4.2, where I find that on average 68.6%–82.6% of the lender's origination revenue is earned through the rate of the mortgage rather than upfront fees. This fact implies that a significant fraction of the price of mortgage origination for quick-to-refinance borrowers may be being paid for through the rate by slow-to-refinance borrowers.

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

3 Data

For my loan-level analyses, I use a combination of three datasets. The first is the 2013–2019 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. Mortgage lenders use rate-locking platforms, such as Optimal Blue, to assist their loan originators and mortgage brokers in identifying options for rate and upfront closing costs for their clients. It contains information about interest rates, points paid or received by the borrower, and time of the lock. Second, I use the January 2013–May 2022 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 dataset. It contains information on loan performance and a measure of the borrower's time-varying credit score in the form of the Equifax Risk Score. The CRISM data also allows me to classify prepayments as moves or refinances.[6] Third, I use the 2013–2019 Home Mortgage Disclosure Act (HMDA) data to capture borrower demographics.

For my main empirical analysis, I construct a novel match of these datasets, leading to the 2013–2019 Optimal Blue-HMDA-CRISM match, with performance until May 2022. I present summary statistics of this match in Table 1. I focus on 30-year, conforming, fixed-rate mortgages due to their status as the predominant form of mortgage contract during this period.[7] Details of the matching as well as summary statistics can be found in the Appendix A.2.

Finally, I obtain actual data on the rate and upfront closing cost menus from LoanSifter.[8] Summary statistics and more detailed descriptions of the LoanSifter data are shown in

[6]I follow the procedure of Lambie-Hanson and Reid (2018) and Gerardi, Willen, and Zhang (2023) to identify moving by classifying a prepayment as a move if the borrower's address changed within a six-month window surrounding the prepayment date.

[7]Among mortgages with balances below the conforming loan limit, complex mortgage contracts used to be more common before the financial crisis, but their market share had plummeted by the start of my sample period (Amromin, Huang, Sialm, and Zhong, 2018).

[8]These two datasets have also been used by Fuster, Lo, and Willen (2022) to study the time-varying price of mortgage intermediation.

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 market income relationship as implied by MBS TBA prices from Morgan Markets in Appendix A.3.

4 Motivating facts

In this section, I present some stylized facts that motivate my model. First, I show that borrowers have heterogeneous refinancing tendencies in Section 4.1. Second, I explore evidence on the selection of borrowers with different prepayment tendencies into upfront closing cost choices in Section 4.2. Third, I look at the price of mortgage origination that is added to rate in Section 4.3.

4.1 Heterogeneous refinancing tendencies

It is well known that some borrowers are slow to refinance, while others are quicker to do so when interest rates fall.[9] This is also true in my Optimal Blue-HMDA-CRISM sample. I study this examining 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 since origination, is greater than 1.2%, which is larger than the optimal refinancing threshold in typical calibrations of both the Agarwal, Driscoll, and Laibson (2013) model and my model as presented in Section 5.

Specifically, the Kaplan-Meier estimates are calculated as follows. Let the number of terminations, due to prepayment at time t, be p_t and the number of loans remaining at time t be n_t, where t is monthly. Then the Kaplan-Meier hazard function is $\hat{\lambda}_p(t) = \frac{p_t}{n_t}$. The Kaplan-Meier survival function is then the cumulative effect of the Kaplan-Meier hazard function, or $\hat{S}_p(t) = \prod_{t'<t} \left(\frac{p_{t'}}{n_{t'}}\right)$.

[9] See, e.g., Archer and Ling (1993), McConnell and Singh (1994), Stanton (1995), Agarwal, Rosen, and Yao (2016), Keys, Pope, and Pope (2016), Johnson, Meier, and Toubia (2018), Andersen et al. (2020), Belgibayeva et al. (2020), Gerardi, Willen, and Zhang (2023), and Byrne et al. (2023).

Figure 2 presents 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 a mark-to-market 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 than half have not prepaid after 10 months of a relatively high refinancing incentive.

4.2 Selection in choices of upfront closing costs

Second, I examine borrower choices of upfront closing costs in my Optimal Blue-HMDA-CRISM data, paying particular attention to selection by borrower prepayment propensities. 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 with different propensities. The choice of upfront closing costs would then perfectly screen borrowers by type, as described in the models of Brueckner (1994), LeRoy (1996), and Stanton and Wallace (2003). While I find some evidence of such selection in the data, I also find evidence of substantial within-choice heterogeneity in moving and refinancing behavior not incorporated in these models which leaves room for cross-subsidization.

In this section, I measure borrower upfront closing costs in terms of points. While the focus on points abstracts from other components of upfront closing costs, such as application fees, to the extent these fees are constant within lender, geography, period, and loan type, my lender-by-county-by-year fixed effects within the sample of 30-year, fixed rate mortgages with controls for loan balance alleviates the effects of this potential measurement error.

To examine selection, I look at how borrower choices of points relate to their moving and refinancing behavior. To do so, with quarterly panel data, I run the the linear probability model on an indicator variable for moving or refinancing:[10]

$$\mathbb{1}_{i,t}(\text{move/refi}) = \sum_{j=1}^{N} \beta_j \mathbb{1}(\psi_i = j) + \gamma Z_i + \xi_{l_{i,t} \times c_{i,t} \times t} + \epsilon_{i,t} \tag{2}$$

where $\mathbb{1}_{i,t}(\text{move/refi})$ is an indicator variable that equals 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 (2000), 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. The call option value of refinancing is defined as:

$$Call\ Option_{i,k} = \frac{V_{i,m} - V_{i,r}}{V_{i,m}} \tag{3}$$

where

$$V_{i,m} = \sum_{s=1}^{TM_i - k_i} \frac{P_i}{(1 + m_{it})^s} \tag{4}$$

$$V_{i,r} = \sum_{s=1}^{TM_i - k_i} \frac{P_i}{(1 + c_i)^s} \tag{5}$$

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 $Call\ Option_{i,k}$ variable represents the potential interest rate savings from refinancing, which is positively correlated with refinancing behavior and is used as a control variable. Finally, $\xi_{l_{i,t} \times c_{i,t} \times t}$ represents lender-by-county-by-year fixed

[10]As noted by Jörn-Steffen Pischke, the linear probability model will be a good approximation to the conditional expectation function if marginal effects are of interest: http://www.mostlyharmlesseconometrics.com/2012/07/probit-better-than-lpm/.

effects, and $\epsilon_{i,t}$ is the error term.

Figure 3 present the results. In particular, Figure 3a 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 decreases with the number of points paid, which is consistent with a selection story. Figure 3b shows the same pattern but for refinancing.

Table 2 shows the regression coefficients that underlie these results, with column (1) showing the results with an indicator variable for whether the borrower moved as the dependent variable and column (2) showing the results with an indicator variable for whether the borrower refinanced as the dependent variable. 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 the control variables, the *Call Option* and log of the loan amount are positively correlated with moving and refinancing, whereas the spread at origination *SATO* is negatively correlated with refinancing.

I next examine the within-choice heterogeneity in borrower prepayment behavior. Figure 4 plots the distribution of borrowers' choices of points by their eventual refinancing or five year prepayment behavior. Here, 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 1.2% at some point within five years after origination. 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.

To ensure that the result of Figure 4 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 Appendix Table A.3. Indeed, while I find a statistically significant positive correlation between non-refinancing borrowers and their payment of points and a statistically significant negative correlation between borrowers who prepay within five years and their choices of points, the magnitude of the difference in points paid is small at no more than 13 basis points. This analysis suggests that most of the heterogeneity between borrower prepayment behavior remains conditional on choices of upfront closing costs.

To further demonstrate the pooling of borrower prepayment types by point choices, I present regression estimates of how borrower choices of points correlate with their demographics with choices of points and various prepayment outcomes as the dependent variable. The regressions are of the form:

$$Y_{i,t} = \beta X_i + \gamma Z_i + \xi_{l_{i,t} \times c_{i,t} \times t} + \epsilon_{i,t} \tag{6}$$

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 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 $Y_{i,t}$ 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 1.2% refinancing rate incentive (i.e. a non-refinancing borrower).

Results are shown in Table 3, with the dependent variable in column (1) being the continuous measure of points paid as an outcome variable, column (2) being an indicator variable for the borrowers' prepayment behavior, and column (3) being an indicator variable for a non-refinancing borrower. I find that while borrowers with larger loan amounts pay more points, and the correlation is low in terms of other borrower characteristics. In particular, borrower characteristics that strongly predict prepayment have low correlation with the amount of points paid. For example, Black borrowers and first-time home buyers are significantly less likely to prepay their mortgages 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.[11] Similarly, Hispanic borrowers are less likely to prepay and refinance but only pay 0.045 more points at the point estimate. These results suggest substantial within-choice variation in prepayment behavior.

4.3 Price of mortgage origination added to the rate

I estimate the extent to which mortgage closing costs are added to the rate based on Equation (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 (2022),[12] where the revenue that lenders generate from the secondary market y_{it} as a fraction of the mortgage balance M_{it} is given as:

$$y_{it} = \frac{p_{it}^{TBA+payup}(c_{it} - gfees_t) - M_{it}}{M_{it}} \tag{7}$$

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 by loan balance from Morgan Markets to the MBS TBA price.

Figure 5 plots the distribution of this secondary marketing income that this secondary

[11] 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 by Bhutta and Hizmo (2020).

[12] The methodology of Fuster, Lo, and Willen (2022) 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.

market income. It is fairly sizable, at 3.46% of the loan amount on average. Furthermore, Appendix A.4.3 shows substantial cross-subsidies in this price of mortgage origination which generates excessive refinancing incentives.

5 Model

The motivating facts in Section 4 show that the existence of significant refinancing inertia in the US mortgage market as well as selection of mortgage contracts by borrower prepayment types. Because borrower contract selection is an important determinant of subsequent refinancing behavior, an equilibrium model that captures the heterogeneity in the selection of borrowers into point choices by their prepayment expectation and the within-choice heterogeneity in borrower refinancing behavior is necessary to study the welfare consequences of borrower financing heterogeneity. My model accomplishes both of these tasks, and I use it to both study the distributional implications of refinancing cross-subsidies in the US as well as the effect of counterfactual contract designs.

On the demand side, following the state-of-the-art from Andersen et al. (2020) and Byrne et al. (2023), 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 exogenous prepayment (e.g. moving) probabilities and discount factors. These decisions are then embedded in a workhorse life-cycle model of mortgage choice from Campbell and Cocco (2015) and Campbell, Clara, and Cocco (2021). A competitive supply side with Monte Carlo mortgage-backed securities pricing pins down mortgage interest rates at various levels of upfront closing costs and closes the model. Novel aspects of the model include embedding rich unobserved heterogeneity in borrower prepayment tendencies in a standard life-cycle context along with supply-side estimates of the mortgage market equilibrium.

Calibration of the model shows evidence of large cross-subsidization of lower upfront closing cost mortgages from slow-to-refinance borrowers. In addition, the fully estimated model allows me to measure the distributional and efficiency implications of alternative mortgage contract designs in terms of adding more of the price of mortgage orignation to the balance of the loan.

5.1 Setup

5.1.1 Demand side

On the demand side, following Campbell and Cocco (2015), households maximize non-housing consumption with time-separable utility with a bequest motive for terminal wealth taking housing choice and initial mortgage size as exogenous:

$$\max \mathbb{E}_1 \sum_{t=1}^{T} \beta_i^{t-1} \frac{(C_{it})^{1-\gamma}}{1-\gamma} + \beta_i^T b \frac{W_{i,T+1}^{1-\gamma}}{1-\gamma}, \tag{8}$$

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

Households receive real log labor income L_{it} and a risk-free nominal return on liquid savings r_t in each period. They also face financial constraints in their liquid savings, such that $S_{it} \geq 0$. In addition, they face inflation π^i, and a tax on labor income τ. Their mortgage interest as well as mortgage points paid are assumed to be deductible.

The value function $V_i(c_{i,t-1}, S_{i,t-1}, \bar{c}_t, r_{t-1}, H_{it}, H_{ig,t}, H_{ig,t-2}, L_{it}, L_{ig,t}, M_{i1}, p_i^a, \kappa_i, p_i^m, t)$ is a function of the state variables including the last period interest rate on the mortgage $c_{i,t-1}$, last period savings $S_{i,t-1}$, the current market interest rate $\bar{c}_t$, the last period risk-free rate r_{t-1}, house price H_{it}, house price growth $H_{ig,t}$, lagged house price growth $H_{ig,t-2}$, labor income L_{it}, labor income growth $L_{ig,t}$, initial mortgage balance M_{i1}, borrower attention probability p_i^a, borrower refinancing hassle cost κ_i, borrower moving probability p_i^m, and quarter t. Of these variables, $c_{i,t-1}, S_{i,t-1}$ evolve endogenously in that they are influenced by the decision to refinance and borrower's consumption decision, while the other states evolve exogenously. Borrower action in each period include the choice of savings $S_{i,t}$ and the the choice of refinancing which determines $c_{i,t}$.

When first getting a mortgage of size M_{i1}, borrowers face with a menu of options for points paid or received, which depends on the rate on the mortgage c. They choose their consumption and savings along with a mortgage interest rate c, which is associated with upfront closing cost $\psi_{it}(c)$, to maximize their expected utility in the first period:

$$\mathbb{E}_1 U_{i1} = \max_{c, S_{i1}} \frac{C_{i1}(c, S_{i1})^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_1 \tilde{V}_{i2}(c, S_{i1}), \tag{9}$$

where S_{i1} is the liquid savings of borrower i at period 1, $C_{i1}(S_{i1}, c)$ is the first period consumption, and the expected continuation value with the mortgage interest rate and next period savings as state variables is $\mathbb{E}_1 \tilde{V}_{i2}(c, S_{i1})$. For brevity, I write the continuation value function as a function of the choice variables only, with the notation $\tilde{V}_{it}(c, S_{i,t-1}) = V_i(c, S_{i,t-1}, \bar{c}_t, r_{t-1}, H_{it}, H_{ig,t}, H_{ig,t-2}, L_{it}, L_{ig,t}, M_{i1}, p_i^a, \kappa_i, p_i^m, t)$.

Conditional on the borrower's choice of their mortgage interest rate c and savings S_{i1}, the implied equation for their first period consumption $C_{i1}(c, S_{i1})$ is:

$$C_{i1}(c, S_{i1}) = (1-\tau)\exp(L_{i1}) + (r_0 - \pi_1) S_{i0} \tag{10}$$

$$- (\bar{\kappa}(M_{i1}) + \psi_{it}(c, M_{i1}) M_{i1}) + \tau \max\{\psi_{it}(c, M_{i1}) M_{i1}, 0\} - (S_{i1} - S_{i0}), \tag{11}$$

where $\exp(L_{i1})$ is the exponential of log labor income, $(r_0 - \pi_1) S_{i0}$ is the return on liquid savings where S_{i0} excludes any liquid assets allocated for the down payment, $\bar{\kappa}(M_{i1}) + \psi_{it}(c) M_{i1} \geq 0$ is the initial cost of mortgage origination consisting of a fixed financial component $\bar{\kappa}(M_{i1})$ and a variable component that depends on the number of points paid $\psi_{it}(c, M_{i1})$, $\tau \max\{\psi_{it}(c, M_{i1}) M_{i1}, 0\}$ is the tax deduction from paying points for a tax rate τ, and $\Delta S_{i1} = S_{i1} - S_{i0}$ is the change in liquid savings.

Borrowers have an option to refinance in each period. To refinance, they need to pay a cost $\tilde{\kappa}_{it}$. I model the borrowers' heterogeneous refinancing cost $\tilde{\kappa}_{it}$ by incorporating a time-dependent component in which borrowers can only refinance with probability p_i^a as well as a state-dependent component κ_i:

$$\tilde{\kappa}_{it} = \begin{cases} \bar{\kappa}(M_{it}) + \kappa_i, \text{ with probability } p_i^a \\ \infty, \text{ with probability } 1 - p_i^a \end{cases} \tag{12}$$

where p_i^a is the probability that a borrower can refinance in a particular period, representing a time varying ability to refinance or the probability that the borrower is paying attention. In addition, $\bar{\kappa}(M_{it})$ is the financial cost of refinancing a mortgage of size M_{it} at zero points, and κ_i is an individual specific hassle cost of refinancing which may reflect either psychic costs experienced by the borrower or supply side frictions. As described by Andersen et al. (2020), the inclusion of time- and state-varying refinancing costs is necessary to fit the data where borrowers do not immediately refinance when facing their cutoff. The borrower's time and state-varying refinancing cost $\tilde{\kappa}_{it}$ may be generated by a combination of demand and supply side factors, where the former factors may include inattention and inertia and the latter may include lender discrimination and the costs of accessing financial institutions.

Mortgage payments follow the standard 30-year amortization schedule. In particular, the real quarterly mortgage payment under inflation π^i is $P_t^M(c_{it}) = \frac{1}{(1+\pi^i)^t} M_{i1} \frac{c_{it}/4(1+c_{it}/4)^n}{(1+c_{it}/4)^n - 1}$. Similarly, the nominal quarterly mortgage interest payment is computed by finding the nominal remaining balance and then multiplying by $c_{it}/4$. Note that the mortgage amortization for an initial balance M_{i1} is computed based on the current rate rather than the full history of rates, which preserves state space and increases the computational tractability of the model. I add a correction for the difference in amortization as an additional upfront payment to be made by the borrower during refinancing so as to be more numerically correct, but the error resulting from this issue is likely to be small for minor differences in rates.

In periods when the borrower does not refinance, households with interest rate c_{it} make a real mortgage payment $P_t^M(c_{it})$, earn interest r_{1t} on savings minus inflation π^i, and receive a tax deduction on their mortgage interest payment $\tau P_t^{IM}(c_{it})$ where $P_t^{IM}(c_{it})$ is the interest portion of the mortgage payment, and so in non-refinancing periods their non-durable consumption C_{it} in real terms can be written as:

$$C_{it}^{nr}(S_{it}) = (1-\tau)\exp(L_{it}) + (r_{1,t-1} - \pi^i)S_{it-1} \tag{13}$$

$$- P_t^M(c_{it}) + \tau P_t^{IM}(c_{it}) - (S_{it} - S_{i,t-1}), \tag{14}$$

where $\Delta S_{it} = S_{it} - S_{i,t-1}$ is the change in the borrower's savings.

In periods when the borrower refinances, that person's consumption as a function of their chosen rate c can be written as:

$$\begin{aligned} C_{it}^r(c, S_{it}) =& (1-\tau)\exp(L_{it}) + (r_{1,t-1} - \pi^i)S_{it-1} - P^M(c_{it}) - (\tilde{\kappa}_{it} + \psi_{it}(c)M_{it}) \\ &+ \tau P^{IM}(c_{it}) + \tau \max\{\psi_{it}(c, M_{it})M_{it}, 0\} - (S_{it} - S_{i,t-1}) \end{aligned} \tag{15}$$

Therefore, conditional on not moving, the borrower's utility $\mathbb{E}_t U_{it}^{nm}$ can be written as the maximum of what can be obtained by refinancing and not refinancing:

$$\mathbb{E}_t U_{it}^{nm}(c_{i,t-1}, S_{i,t-1}) = \max \begin{cases} \max_{S_{it}} \frac{(C_{it}^{nr}(S_{it}))^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_t \tilde{V}_{i,t+1}(c_{it}, S_{it}) \\ \max_{S_{it},c} \frac{(C_{it}^{r}(c,S_{it}))^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_t \tilde{V}_{i,t+1}(c, S_{it}) \end{cases} \tag{16}$$

where the first line of Equation (16) 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 paid $\psi_{it}(c)$. The choice that gives the maximum value in this problem represents the borrower's refinancing policy.

Given the same refinancing cost $\tilde{\kappa}_{it}$, borrowers with more income and liquid savings will refinance at a weakly lower interest rate incentive because they can more incur their refinancing cost without increasing their marginal utility of consumption. Indeed the liquid savings constraint of $S_{it} \geq 0$ will prevent some borrowers from refinancing altogether if they do not have sufficient income and liquid savings to pay the refinancing cost. Therefore my model's estimated refinancing costs automatically incorporates the fact that borrowers with lower incomes and discount rates may be less willing to refinance.

In addition to the maximization problem in Equation (16), I require that, to refinance, borrowers must have a mark-to-market loan-to-value (LTV) ratio of at most 95%, which is required by Freddie Mac[13] and captures the constraints to refinancing in periods of house price decline as described in Hurst, Keys, Seru, and Vavra (2016). This requirement is rarely binding in my empirical estimation, due to the January 2013 to May 2022 period in which I assess mortgage performance. Defaults are also rare in my sample period, and I do not include mortgage default in the model to conserve state space because it is unlikely that the borrowers' upfront closing costs and refinancing decisions in my sample period are significantly influenced by their default expectations.

Furthermore, I assume that borrowers expect to get the market rate-upfront closing cost trade-off when they refinance, which increases the computational tractability of the model. The market rate-upfront closing cost menu is modelled as the average rate and upfront closing cost trade-off, as implied by MBS TBA prices for deviations from the PMMS survey rate. A more detailed description of the construction of the trade-offs is described in Section 5.1.2. When fitting the model to loan-level performance data, I use an individual-specific refinancing rate based on an interest rate that depends on the borrower's updated Equifax Risk Score, a mark-to-market loan-to-value (LTV) ratio, property state, and the borrower's loan amount. To the extent that refinancing expectations are primarily based on expectations of macroeconomic rate movements, the implication of the restriction on borrower expectations for the borrower's optimal refinancing policy is likely small.

In terms of exogenous state transitions, I assume that the risk-free rate r_t follows the

[13] 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.

model of Cox, Ingersoll, and Ross (1985), which has a natural zero lower bound. I take inflation $\pi^i = 1.68\%$ as a constant equal to the average in my sample. Real (log)labor income L_{it}, house price H_{it}, and changes in the market mortgage interest rate given by the Freddie Mac Primary Mortgage Market Survey (PMMS) $\bar{c}_t$ are modelled as a vector autoregression (VAR) with the risk-free rate r_{1t} as an exogenous covariate, one lag for labor income growth, and two lags for house price growth, the details of which are described in Appendix A.5.1.

I model exogenous prepayment, which includes moving and cash-out refinancing, as an exogenous refinance to the new mortgage with an interest rate $\bar{c}_t$ that is associated with $2,000 plus 1 point in upfront closing costs as in Agarwal, Driscoll, and Laibson (2013) occurring with probability p_i^m for borrower i. More specifically, the borrower's utility upon exogenous prepayment is:

$$\mathbb{E}_t U_{it}^m(c_{i,t-1}, S_{i,t-1}) = \max_{S_{it}} \frac{(C_{it}^r(S_{it}, \bar{c}_t))^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_t \tilde{V}_{i,t+1}(S_{it}, \bar{c}_t) \tag{17}$$

Gerardi, Willen, and Zhang (2023) finds that cash-out refinancing volume is not strongly correlated with interest rate movements and may be more driven by the household's idiosyncratic demand for liquidity. Since I do not study prepayment penalties as a counterfactual, to the extent that utility from moving and cash-out refinancing are independent of borrower rate and term refinancing types, the assumption that they are exogenous should also not significantly affect my results on rate and term refinancing heterogeneity. Combined, the value function of the borrower can be written as:

$$\mathbb{E}_t V_{it} = (1 - p_i^m)\mathbb{E}_t U_{it}^{nm}(c_{i,t-1}, S_{i,t-1}) + p_i^m \mathbb{E}_t U_{it}^m(c_{i,t-1}, S_{i,t-1}). \tag{18}$$

where the value function $\mathbb{E}_t V_{it} = V_i(c_{i,t-1}, S_{i,t-1}, \bar{c}_t, r_{t-1}, H_{it}, H_{ig,t}, H_{ig,t-2}, L_{it}, L_{ig,t}, M_{i1}, p_i^a, \kappa_i, p_i^m, t)$ is a function of the state variables as described earlier, and the borrower choices are in each period are the savings $S_{i,t}$ and the refinancing decision which determines $c_{i,t}$. Table 4 lists the calibrated parameters of the model that is imposed throughout the paper.

This model cannot be solved analytically. Computationally, to handle the large number of states I solve the model in parallel on a large high performance compute cluster for a grid for the time invariant variables $\{M_{i1}, p_i^a, \kappa_i, p_i^m\}$, with each element of the grid corresponding to a value function of only the remaining time variant variables. To solve the model, I use the nested fixed-point method with code adapted from Druedahl and Jørgensen (2017).

5.1.2 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 mortgage-backed securities (MBS) value of mortgages. That is, in equilibrium, the relationship between the upfront points paid as a fraction of the loan amount $\psi_{it}(c, M_{it})$ for borrower i at time t and the mortgage interest rate c for a new purchase mortgage is pinned down by a zero profit condition:

$$\pi_{it}^l = \bar{\kappa}(M_{it}) + \psi_{it}(c, M_{it})M_{it} + \phi_t(c)M_{it} - \bar{m}_t^l - m_{it}^l = 0 \tag{19}$$

where π_{it}^l is lender profit from an originating loan to borrower i at time t, $\bar{\kappa}(M_{it})$ is the upfront cost of mortgage origination at zero points, $\phi_t(c)$ is the MBS premium of the mortgage as a percentage 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_{it}^l is the borrower and loan amount specific marginal cost incurred by the lender for originating the loan with $E(m_{it}^l) = 0$. Assuming that the marginal cost of loan origination $\bar{m}_t^l + m_i^l$ does not vary by the borrower's choice of points, we have by re-arranging:

$$\psi_{it}(c, M_{it}) = \frac{\bar{m}_t^l + m_{it}^l - \bar{\kappa}(M_{it})}{M_{it}} - \phi_t(c). \tag{20}$$

Therefore, all else equal, a mortgage with a higher interest rate c and MBS value $\phi_t(c)$ would require fewer upfront points ψ_{it}. In particular, my model implies that the MBS value of mortgages with higher interest rates will be passed through to borrowers in terms of lower upfront closing costs. This pass-through implication is approximately true in empirically, as I show in Figure A.3.

Computing the zero profit condition in Equation (20) requires an estimate of MBS prices $\phi_t(c)$ both in the current world and in counterfactuals. Actual MBS prices incorporate heterogeneous borrower refinancing behavior in a pooling equilbirium, but MBS prices would differ in separating equilibrium counterfactuals without cross-subsidization between borrower refinancing types. Estimation of counterfactual MBS prices is therefore key to establishing the interest rate effect of heterogeneous borrower refinancing.

I use a one-parameter, option-adjusted spread (OAS) model to estimate the MBS value of mortgages $\phi_t(c)$, based on an standard expected NPV method, where the cashflows from MBS are assumed to be discounted based on a stochastic risk-free rate r_t 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 by Gabaix, Krishnamurthy, and Vigneron (2007), Song and Zhu (2018), and Boyarchenko, Fuster, and Lucca (2019), and Diep, Eisfeldt, and Richardson (2021).[14] Under this setup, the MBS value of mortgages net of their current balance may be written as:

$$\phi_t(c) = \frac{1}{M_{it}} E_t \sum_{t'=t}^{t+T} \delta_{t'} q_{t'} [(1 - p_{t'}) P^M(c) + \hat{p}_{t'} B^M_{t'}] - 1 \tag{21}$$

where $p_{t'}$ is the prepayment probability of the borrower at time t', $q_{t'} = \prod_{j=t}^{t'-1}(1 - p_j)$ is the remaining proportion of borrowers who have not prepaid, $B^M_{t'}$ is the remaining principal the lender gets when a borrower prepays, the lender gets remaining principal $B^M_{t'}$, and $P^M(c)$ is

[14] Another method of valuing MBS is via multivariate density estimation, following Boudoukh, Whitelaw, Richardson, and Stanton (1997), but that does not allow me to get counterfactual prices under alternative prepayment behavior or with alternative mortgage contract designs.

the regular mortgage payment. The discount factor is based on the cumulative risk-free rate in period j, r_{jf}, plus an estimated OAS term that compensates for liquidity and prepayment risk:

$$\delta_{t'} = \frac{1}{\prod_{j=t}^{t'}(1 + r_{jf} + OAS)}. \tag{22}$$

Based on Equations 21 and (22), an estimate of the OAS combined with borrower refinancing behavior allows me to arrive at counterfactual MBS prices. To estimate the OAS, I use actual MBS prices combined with an empirical prepayment hazard function $\hat{p}_{t'}$ and its implied empirical cumulative remaining balance $\hat{q}_{t'} = \prod_{j=t}^{t'-1}(1 - \hat{p}_j)$. Furthermore, I take the 10-year yield implied by my Cox, Ingersoll, and Ross (1985) as the r_{jf} on the supply side, as the 10-year yield is a commonly used discount rate for pricing MBS. Other details of the OAS estimation is shown in Appendix A.5.2.

Since in a segmented markets model of Gabaix, Krishnamurthy, and Vigneron (2007) the OAS partially reflects the price of prepayment risk, it could be higher if the quick to refinance borrowers were isolated by themselves. Assuming that it does not change in counterfactual analyses would therefore lead to conservative estimates of the cross-subsidization between borrower refinancing tendencies. The reason is that a higher OAS for quick to refinance borrowers would imply an even greater cross-subsidy, as they would be subsidized not only by the cash flows of the slow to refinance borrowers but also a reduced discount rate due to their presence..

As noted in Section 5.1.1, I make an assumption that, when computing their optimal refinancing policy, borrowers expect to receive the market rate when refinancing, which implies $m^l_{it} = 0$ for $t > 0$ in Equation (20). Abstracting from the borrower's expectation of the evolution of unobserved borrower specific markups is consistent with the literature (e.g., Agarwal, Driscoll, and Laibson (2013) and Berger, Milbradt, Tourre, and Vavra (2021)) and saves computational time. Borrowers still originate their purchase mortgage at an individual-specific markup M_{i1}^l, which is pinned down by their initial rate gap from the market rate, and when fitting the model to data I do use a borrower-specific refinancing rate based on their mark-to-market loan to value ratio and time-varying credit scores.

5.2 Equilibrium

The demand and supply side jointly determine a mortgage market equilibrium. In particular, in this equilibrium:

1. Borrower value functions are described by Equation (18),
2. Rate and upfront closing cost choices are determined by Equation (20),
3. MBS prices are determined by Equation (21).

Note that by varying $\bar{m}_t^l$ and m_{it}^l with $E(m_{it}^l) = 0$, I am able to fit any time-series variation in mortgage rates as well as the dispersion of mortgage rates at origination. Therefore, for estimation in Section 6 I am able to take the movement in mortgage interest rates as given by a vector auto-regression (VAR) as detailed in Appendix Section A.5.1 and recover the distribution of borrower types. Counterfactual analyses then require me to compute counterfactual MBS prices given by Equation (21) and their implied rate and upfront closing costs given by Equation (20) holding fixed lender costs $\bar{m}_t^l$ and m_{it}.

5.3 Calibration evidence of cross-subsidization

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. All of the analysis in this section is conducted for a calibrated borrower with parameters described in Table 5, where p^m is the mean of the estimates from Section 6 and $p_i^a = 1, \kappa_i = 200$ are chosen to represent the behavior of an optimally refinancing borrower, who can always refinance with a hassle cost of $200.

Figure 6 illustrates this pricing impact of cross-subsidization. The market rate and upfront closing cost trade-off, as implied by the model, is shown by the blue solid line, and the empirical rate and upfront closing cost trade-off, as implied by MBS TBA prices, is presented by the black dotted line.[15] The close match between the solid and dotted lines suggests that the supply side of the model, which is the previously described OAS model of MBS valuation, matches the average empirical rate and upfront closing cost trade-off well.

I then compare the counterfactual rate and upfront closing cost trade-off if there were no cross-subsidization. This counterfactual was computed by jointly iterating on the borrower and lenders' problems while holding the lenders' revenue constant. The joint iteration is conducted as follows. In the last period T, the borrower cannot refinance. In the second-to-last period $T-1$, the lender creates a rate-and-upfront closing costs schedule $\psi_{i,T-1}(c, M_{it})$ based on the borrower specific MBS value $\phi_{i,T-1}(c, M_{it})$ that is dependent on period T borrower prepayment, and then the borrower makes his or her $T-1$ period prepayment decision. This process was repeated until the model is solved via backward induction.

The red dotted line in Figure 6 shows that the interest rate trade-off would be significantly 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 lower than in the no cross-subsidization case, due to the presence of slow-to-refinance borrowers. In terms of numbers, I find that a mortgage with a one percent upfront closing cost would carry a 0.97 percentage points higher interest rate in the no cross-subsidization case, relative to the existing market equilibrium, whereas the difference is smaller but still substantial at 0.21 percentage points for a mortgage with a four percent upfront closing cost.[16] This pricing result is not specific to my model: Appendix A.6 shows that it is also true under the Agarwal, Driscoll, and Laibson (2013) model of borrower

[15] As explored in Appendix A.3, the rate and upfront closing cost trade-off in lender rate sheets in turn matches those implied by MBS TBA prices on average cross-sectionally and over time.

[16] This comparison excludes zero upfront closing cost mortgages, whose pricing may be affected by search frictions described by Woodward and Hall (2012). My calibration suggests that zero upfront closing cost mortgages would carry significantly higher interest rates for the quick-to-refinance borrower in the no cross-subsidization counterfactual, with an average interest rate above the upper limits of my grid at 7%.

behavior.

Two reasons explain why lower upfront closing cost mortgages are cross-subsidized more than higher upfront closing cost mortgages. First, the former mortgages carry a higher interest rate and therefore a higher option value of refinancing. Slow-to-refinance borrowers end up paying more for these mortgages to the extent they do not take advantage of the option value of refinancing. Second, the selection of slow-to-refinance borrowers to higher upfront closing cost mortgages has to be weak enough so that the first effect dominates. This is because slow-to-refinance borrowers can choose to pay more in upfront closing costs, and, if they do so, they would avoid the excess cross-subsidization of lower upfront closing cost mortgages. My calibration result that lower upfront closing cost mortgages receive more cross-subsidization from slow-to-refinance borrowers is consistent with my motivating fact in Figure 4, which shows that, while there is some selection of contracts by borrower prepayment tendencies, the magnitude of the selection is small when compared to the variance in borrower prepayment behavior.

Figure 7 presents the welfare implications of this calibration for the borrower, lender, and society. Compared to a no cross-subsidization counterfactual, the calibrated quick-to-refinance borrower benefits from cross-subsidization and would have to be paid 2.4% of the loan amount in liquid assets in the no cross-subsidization counterfactual to be indifferent from the current world. The lender, on the other land, loses 5.5% of the loan amount in profit from the quick-to-refinance borrower in the current world, which they compensate for through higher interest rates on all borrowers. Overall cross-subsidization leads to a social loss of 3.0% of the loan amount. This social loss comes from the excessive refinancing that this quick-to-refinance borrower undertakes at the expense of the slow-to-refinance borrowers.

6 Estimation

To estimate the distribution of borrower heterogeneity in the population, I allow $p_i^a, \kappa_i, \beta_i, p_i^m, M_i$ to vary by individual and parametrize their distributions, where p_i^a is the probability that an individual is available to refinance in a particular period, κ_i is the individual's refinancing hassle cost when refinancing, 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 $\gamma = 2$ and a bequest motive of $b = 400$ in accordance with Campbell and Cocco (2015). When evaluating the policy function at home purchase and refinance, I calibrate the model's liquid assets to the distribution of liquid savings from SCF by race and ethnicity, as described in Appendix Section A.5.3.

Since I construct counterfactual mortgage interest rates based on a MBS pricing model estimated on TBA prices, to maintain comparability to the TBA market, I further restrict my sample to 30-year purchase mortgages with a balance above \$150k, FICO above 680, and LTV below 85%, which are mortgages that are likely to be sold through the TBA market (Fusari, Li, Liu, and Song, 2020).

When estimating the distribution of borrower heterogeneity, I include control variables for whether the mortgage was originated by a non-bank as well as FICO, LTV, and loan amount bins. This allows me to incorporate the fact that, for example, mortgages from non-bank lenders are faster prepaying on average which may be already priced in by the lenders. My cross-subsidization results controls for the effect of these observables by taking loan-level rate and upfront closing costs as given and looking at counterfactual prices based on the estimated distribution of borrower heterogeneity conditional on the observables.

I first present the identification argument in Section 6.1, the maximum likelihood estimation procedure in Section 6.2, results in Section 6.3, some calibration based on my estimates in Section 5.3, and finally the implications of my estimates for transfers and welfare in Section 6.4.

6.1 Identification

Of the unknown parameters, the distribution of M_i is observed. I discuss the identification for the distribution of $p_i^a, \kappa_i, \beta_i, p_i^m$ 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. 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 by Andersen et al. (2020).

Of the other parameters, prepayments at low interest rate incentives identify the mean of exogenous prepayments p_i^m, while the correlation between sluggish refinancing and subsequent exogenous prepayments identify its variance. Conditional on refinancing and exogenous prepayment 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. To the extent that slow to refinance borrowers may also face behavioral frictions in their choices of points, their β_i may be underestimated, which would lead to a conservative estimate of their NPV loss from failure to refinance. Since most of the excessive refinancing comes from the quick to refinance borrowers responding to their incentives, the model's welfare implications are likely not significantly affected by the mechanism used to rationalize the slow to refinance borrowers' choices.

6.2 Parametrization

I estimate the distribution of the borrower types using mortgage performance data. Specifically, I use a log-normal distribution to model p_i^a, p_i^m and κ_i as well as a logit-normal distribution[17] to model β_i. I allow β_i to be correlated to p_i^a, p_i^m and κ_i through coefficients $\rho_{\beta,p^a}, \rho_{\beta,\kappa}, \rho_{\beta,p^m}$. The precise parametrization is as follows:

$$p_i^a \sim LogNormal(\mu_{p^a}(X_i), \sigma_{p^a}) \tag{23}$$

$$\kappa_i \sim LogNormal(\mu_\kappa(X_i), \sigma_\kappa) \tag{24}$$

$$p_i^m \sim LogNormal(\mu_{p^m}(X_i), \sigma_{p^m}) \tag{25}$$

$$\beta_i \sim \textit{Logit-Normal}\left(\mu_\beta(X_i) + \rho_{\beta,p^a} p_i^a + \rho_{\beta,p^m} p_i^m + \rho_{\beta,\kappa}\kappa_i, \sigma_\beta\right) \tag{26}$$

where the means of the distributions $\mu_{p^a}(X_i), \mu_\kappa(X_i), \mu_{p^m}(X_i), \mu_\beta(X_i)$ are allowed to depend on a set of observed covariates X_i, including the race of the borrower, FICO bins, LTV bins, loan amount bins, and an indicator for whether the borrower chose a bank or nonbank lender.

The model is estimated via maximum likelihood over both borrower prepayment decisions and choices of upfront closing costs. In terms of prepayment decisions, the likelihood function for a prepayment decision y_{it} for borrower i at time t, given a set of model parameters $k_i = \{p_i^a, \kappa_i, \beta_i, p_i^m, M_i\}$ is:

$$l_{j,t}(k_i) = (1 - p_{jt}(k_i))^{1-y_{jt}} p_{jt}(k_i)^{y_{jt}}. \tag{27}$$

where $p_{jt}(k_i) = p^m(k_i) + (1 - p^m(k_i))(\mathbb{1}_{refi,jt}(k_i; z_{jt}))$ is the model predicted prepayment including the exogenous component $p_j^m(k_i)$ and the refinancing policy $\mathbb{1}_{refi,it}(k_i; z_{jt})$, which depends on both parameters k_i and the estimated refinancing interest rate of the borrower z_{jt} from the borrower's updated Equifax risk score, mark-to-market LTV, the property state, the remaining balance of the loan, and the period t.

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

[17] 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.

closing costs ψ_{j0} that is equal to the optimal choice implied by the model ψ_0^*, in whole numbers from -2 to 2, is:

$$l'_j(\theta) = \prod_{q=-2}^{2} \Pr(\psi_0^* = q|\theta, X_i)^{\mathbb{1}(\psi_{j0}=q)}. \tag{28}$$

where $\Pr(\psi_0^* = q|\theta, X_i)$ is the probability of a borrower paying points q, given parameters θ and observables X_i.

To estimate the model, I simulate individuals with a grid for $k_i = \{p_i^a, \kappa_i, \beta_i, p_i^m, M_i\}$ based on a set of parameters θ, with $k_i \sim \mathcal{F}(\theta)$ where $\mathcal{F}(\theta)$ is the distribution of types from Equations (23) to (26). I then get their point choices (in whole numbers from -2 to 2) and time-varying prepayment (i.e., refinancing and moving) decisions for each loan-time observation $p_{jt}(k_i)$ and search for the set of parameters that maximizes the combined likelihood of the data following the simulated maximum likelihood formulation:

$$\mathcal{L}(\theta) \propto \sum_j \log\left(l'_j(\theta)\right) + \sum_j \sum_i \log\left(\prod_{t=1}^{T_j} l_{j,t}(k_i)\right) |M_j = M_i, k_i \sim \mathcal{F}(\theta), \tag{29}$$

where $nsim = 4,320$ is the number of simulations used to compute the likelihood function. In particular, I solve the model in parallel on 2.90 GHz cores across the $nsim = 4,320$ grid of values for $M_i, \beta_i, p_i^m, p_i^a, \kappa_i$, with approximately 10 hours per agent, and then estimate the distribution of these parameters using the simulated maximum likelihood, as in Equation 29.

6.3 Distribution of borrower types

In this section I present my estimates for the distribution of borrower types in the population. The parameters and their standard errors are shown in Table 6, and I plot their distributions in the rest of this section.

Figure 8 presents the estimates on the distribution of refinancing types in the population. In the left panel of Figure 8a, the results show that many borrowers have a low probability of being able to refinance in a particular quarter, though with significant variance. Mean able-to-refinance probability is 16.4% quarterly or 48.8% annualized. This is consistent with my stylized fact in Section 4.1 showing that around half of all borrowers fail to refinance following 10 months of a relatively high refinancing incentive. In the right panel in Figure 8b, the results show that the implied hassle cost of refinancing for most borrowers is below $2,000, though higher for Black borrowers. Taken together, the results suggest that most of the inaction in refinancing is due to a time-varying inability to refinance rather than hassle costs. The importance of the time-varying ability to refinance is consistent with the randomized controlled trial results of Byrne et al. (2023), who found borrower attention to be a significant friction in refinancing. Interestingly, Hispanic borrowers have similar refinancing ability and hassle costs as white borrowers, despite being less likely to refinance, and that is explained by their lower liquid assets, as found in the SCF (Section A.5.3).

Figure 9 presents my estimates for borrower moving probabilities and discount factors. Figure 9a presents my estimates of the distribution of exogenous prepayment probabilities by borrower. As Figure 9a shows, the annualized exogenous prepayment probabilities are centered around 7.2% per year, which implies a median half-life of 9.3 years. The mean exogenous prepayment probability is 7.4% per year. Figure 9b plots the distribution of discount factors, which is above 0.9 for most borrowers. The median discount factor is 0.925. The fact that borrowers have lower discount factors than MBS investors implies that they are less willing to pay points, even if they have lower-than-average refinancing propensities, which generates the cross-subsidization of lower upfront closing cost mortgages.

6.4 Implications for cross-subsidization across the population

In this section, I examine the implied cross-subsidization across the population implied by my empirical estimates of the distribution of borrower types. To do so, I estimate the implied lender profit per loan from each borrower type, which by construction averages to zero overall, and examine how it varies by borrower refinancing type as well as race.

The results are shown in Table 7. In particular, row 1 of Table 7 shows that the non-refinancing borrower (borrowers with $p^a = 0$) in the population on average loses \$3,634 to cross-subsidization, whereas row 2 of Table 7 shows quick to refinance borrowers that is always able to refinance with \$200 in hassle costs $p^a = 1, \kappa = 200$ receives \$4,537 in cross-subsidies. The average Black borrower loses \$1,983, the average Hispanic borrower loses \$387, and the other borrowers on average gain \$77. Figure 10 presents these results visually, with Figure 10a plotting the average cross-subsidies received by borrower refinancing ability in the population and Figure 10b plotting those cross-subsidies by race.

7 Counterfactual contract designs

I conduct two additional counterfactual analyses. First, I consider an alternative mortgage contract design where the portion of the price of mortgage origination added to the rate has to instead be added to the balance of the loan. Notably, it does not change the total level of household borrowing nor their upfront payment since it is only a shift in the form of borrowing, from an a spread on the rate to the balance of the loan. This design eliminates the cross-subsidization of the price of mortgage origination. Second, I consider the case of automatically refinancing mortgages, which is a mortgage whose interest rate resets downward automatically to a lower rate when the market rates falls by more than 0.5%.[18] In both of these cases, I re-estimate equilibrium interest rates and MBS prices.

7.1 Adding more of the price of mortgage origination to the balance

First, I consider the utility changes of borrowers when they switch the portion of their price of mortgage origination added to the rate of the mortgage to the balance of the mortgage, with LTV limits relaxed correspondingly. This does not change the amount of total borrowing

[18]This form of contract has been discussed by Campbell (2006).

nor the upfront payment required of the borrower. In particular, I take the amount of the price of mortgage origination added to the rate y_{it} as calculated in Equation (7) and instead add it to the balance of the loan while setting the mortgage rate to the par.

Mathematically, borrower mortgage balance at origination becomes $M'_{it} = M_i(1 + y_{it})$, where y_{it} is the lender's price of origination added to the rate, and borrower mortgage payment becomes:

$$P_{it}(M'_{it}) = M'_{it} \frac{c_{it}/4(1 + c_{it}/4)^n}{(1 + c_{it}/4)^n - 1}. \tag{30}$$

where c_{it} is set equal to the interest rate where the lender's secondary marketing income is equal to zero. In periods where borrowers can 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 $M'_{it'}$ to $M'_{it'}(1 + y_{it'})$. Hence the mortgage size M'_{it} is an endogenous state variable that affects the size of the mortgage payment $P_{it}(M'_{it})$.

Without changing the borrower's upfront payments, LTV restrictions are assumed to be relaxed to enable the portion of the price of mortgage origination added to the rate to be added to the balance of the loan instead. One potential objection to this contract design is that the resulting higher LTV may increase default rates. However, this concern is mitigated by the fact that this design leads to lower monthly payments compared to adding that the price of mortgage origination to the rate. To the extent that liquidity and payment size is a more important determinant of default than LTV (Ganong and Noel, 2020), this contract design may reduce default rates in addition to being more equitable in terms of eliminating the cross-subsidization of the price of mortgage origination. Furthermore, the FHA has already implemented a more limited version of this contract, where they allow the 1.75% upfront mortgage insurance premium to be added to the balance of the loan without regard to LTV restrictions.

The welfare effects of this counterfactual contract design are shown in column (1) of

Table 8. Row (1) shows that, in terms of total welfare, I find that on average borrower welfare increases by $1,461 per new purchase origination. This implies a substantial welfare improvement, on the order of $7.3 billion per year assuming roughly 5 million new purchase mortgage originations per year as in the 2020–2021 HMDA data. Rows (2) and (3) of Table 8 show that non-refinancing borrowers with $p_i^a = 0$ gain $2,819 per new purchase origination under this contract design. Quick-to-refinance borrowers with $p_i^a = 1, \kappa_i = 200$ lose $1,310 per new purchase origination. Rows (4) to (6) show that all racial groups and, in particular, Black and Hispanic borrowers gain disproportionately from this contract design, at $2,642 and $2,478, respectively. Figure 11a plots the change in borrower utility by borrower refinancing ability, and Figure 11b presents the results by race.

To the extent that moving all borrowers to this contract design leads to a different the composition of borrowers than those who select into higher upfront closing cost mortgages, the equilibrium mortgage interest rate may differ from what is implied by the existing market rate and upfront closing cost trade-off. This difference is expected to be small since in this counterfactual the mortgage backed securities are issued at par, and so their value do not fluctuate as much with the likelihood of borrower prepayment. Nevertheless, I address this possibility using a recursive procedure. First, I obtain the implied borrower refinancing behavior and lender values based on my estimated distribution of borrower types. Then I increase market interest rates as well as the interest rates of new purchase originations by a constant additive factor so that the lender zero profit condition holds on average, and then iterating on the borrower's policy functions and lender zero profit conditions until convergence. This procedure leads to an increase in rate of 1.9 basis points, as shown in row 7, column 1 of Table 8, which confirms that this effect is small.

Taken together, my results in Table 8 suggest that a simple change in the form of financing the price of mortgage origination, from switching the portion that is added to the rate to added to the balance of the loan, substantially reduces the cross-subsidization between borrower types while increasing average welfare.

7.2 Making mortgages automatically refinancing

Second, I consider a counterfactual where mortgages are automatically refinancing, with a 0.5% buffer. In particular, the mortgage interest rate automatically drops from c to $c + \Delta$ whenever the change in the 10-year yield since origination Δ falls by more than half a percentage point ($\Delta < -0.5$). Making the automatically refinancing mortgage dependent on the 10-year yield instead of the mortgage rate avoids the recursion in mortgage rates, and the 0.5 percentage point buffer makes the interest rate trade-off of automatically refinancing mortgages lower than it otherwise would be. I eliminate the choice of manual refinancing in this model and assume that all mortgages are originated with zero upfront closing costs so that all closing costs are added to the rate of the mortgage. Computationally, I add the 10-year yield at origination r_{ft}^{o} as a state variable, such that the borrower's mortgage interest rate at time t, c_t, is:

$$c_t = \begin{cases} c_{t-1} \text{ if } r_{ft} - r_{ft}^{o} \geq -0.5 \\ c_{t-1} + r_{ft} - r_{ft}^{o} \text{ if } r_{ft} - r_{ft}^{o} < -0.5 \end{cases} \tag{31}$$

where c_{t-1} is the borrower's interest rate at time $t-1$, and r_{ft} is the 10 year yield.

An important feature of automatically refinancing mortgages is the higher interest rates at origination due to a higher option value embedded in the contract. I calculate counterfactual mortgage interest rates under this alternative contract design by backward induction. In particular, under this counterfactual, borrower refinancing behavior is no longer heterogenous, and the only relevant unobserved heterogeneity across borrowers is their moving types p_i^m. I solve the model for the mortgage interest rate increase at new purchase origination for lenders to break even taking as given my estimated distribution of borrower p_i^m, and find that interest rates would rise by 0.196 percentage points on average under this contract design.

The welfare implications of this counterfactual contract design are shown in column (2) of Table 8. In terms of overall welfare, I find that switching to automatically refinancing mortgages features an increase of \$1,362 per new purchase origination compared to the current world. Nonrefinancing borrowers gain \$1,498 per new purchase origination, whereas quick-to-refinance borrowers with $p_i^a = 1, \kappa = 200$ lose \$2,016. Black, Hispanic, and other borrowers also gain. Figure 12a plots the change in borrower utility by borrower refinancing ability, and Figure 12b presents the results by race.

Conceptually, there are two main channels through which automatically refinancing mortgages can affect average welfare. First, they generate resource savings by eliminating the administrative and hassle costs of refinancing. Second, they front-load the mortgage payments through an increase in mortgage rates, which may be less desirable to financially constrained borrowers. The second channel likely explains why Black and Hispanic borrowers as well as borrowers with lower p_i^a gain less from automatically refinancing mortgages relative to adding closing costs to the balance. The average welfare improvement from automatically refinancing mortgages is also lower. Therefore, my first alternative contract design of switching the price of mortgage origination from being added to the rate to being added to the balance of the loan may be more preferable for policymakers looking to eliminate the key inefficiencies identified in my paper.

8 Discussion

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 lower upfront cost mortgages are more cross-subsidized by slow-to-refinance borrowers. I show that, through this effect, slow-to-refinance borrowers end up incentivizing excessive origination, which generates deadweight administrative costs and significant welfare losses.

In addition to heterogeneous borrower prepayment tendencies, heterogeneous borrower default tendencies is also a channel for cross-subsidization. Since borrower refinancing and default tendencies are unlikely to be perfectly correlated with one another, however, one does not justify the other. Furthermore, the inefficient refinancing identified in this paper is a rationale for considering alternative contract designs on top of any pure distributional concerns.

References

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao. 2017. "Systematic Mistakes in the Mortgage Market and Lack of Financial Sophistication." *Journal of Financial Economics* 123(1): 42–58. ISSN 0304-405X.

Agarwal, Sumit, John C. Driscoll, and David I. Laibson. 2013. "Optimal Mortgage Refinancing: A Closed-Form Solution." *Journal of Money, Credit and Banking* 45(4): 591–622.

Agarwal, Sumit, Richard J. Rosen, and Vincent Yao. 2016. "Why Do Borrowers Make Mortgage Refinancing Mistakes?" *Management Science* 62(12): 3494–3509.

Ambokar, Sumedh, and Kian Samaee. 2019. "Inaction, Search Costs, and Market Power in the US Mortgage Market." *Working Paper*.

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. 2020. "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market." *American Economic Review* 110(10): 3184–3230. doi:10.1257/aer.20180865.

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.

Belgibayeva, Adiya, Teresa Bono, Philippe Bracke, João Cocco, and Tommaso Majer. 2020. "When Discounted Rates End: The Cost of Taking Action in the Mortgage Market." *FCA Occasional Paper No. 54*.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra. 2018. "Regional Heterogeneity and the Refinancing Channel of Monetary Policy." *Quarterly Journal of Economics* 134(1): 109–183.

Berger, David, Konstantin Milbradt, Fabrice Tourre, and Joe Vavra. 2023. "Refinancing Frictions, Mortgage Pricing and Redistribution." *Working Paper*.

Berger, David, Konstantin Milbradt, Fabrice Tourre, and Joseph Vavra. 2021. "Mortgage Prepayment and Path-Dependent Effects of Monetary Policy." *American Economic Review* 111(9): 2829–78. doi:10.1257/aer.20181857.

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, and Aurel Hizmo. 2020. "Do Minorities Pay More for Mortgages?" *The Review of Financial Studies* 34(2): 763–789.

Boar, Corina, Denis Gorea, and Virgiliu Midrigan. 2021. "Liquidity Constraints in the U.S. Housing Market." *The Review of Economic Studies* 89(3): 1120–1154. ISSN 0034-6527.

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.

Byrne, Shane, Kenneth Devine, Michael King, Yvonne McCarthy, and Christopher Palmer. 2023. "The Last Mile of Monetary Policy: Consumer Inattention, Disclosures, and the Refinancing Channel." *Working Paper*.

Campbell, John Y. 2006. "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.

Chari, V. V., and Ravi Jagannathan. 1989. "Adverse Selection in a Model of Real Estate Lending." *Journal of Finance* 44(2): 499–508.

Corbae, Dean, and Erwan Quintin. 2015. "Leverage and the Foreclosure Crisis." *Journal of Political Economy* 123(1): 1–65.

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.

Deng, Yongheng, John M. Quigley, and Robert Van Order. 2000. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options." *Econometrica* 68(2): 275–307.

Diep, Peter, Andrea L Eisfeldt, and Scott Richardson. 2021. "The cross section of MBS returns." *The Journal of Finance* 76(5): 2093–2151.

Druedahl, Jeppe, and Thomas Høgholm Jørgensen. 2017. "A general endogenous grid method for multi-dimensional models with non-convexities and constraints." *Journal of Economic Dynamics and Control* 74: 87–107.

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.

Fisher, Jack, Alessandro Gavazza, Lu Liu, Tarun Ramadorai, and Jagdish Tripathy. 2023. "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, Stephanie H. Lo, and Paul S. Willen. 2022. "The Time-Varying Price of Financial Intermediation in the Mortgage Market." *Working paper*.

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*.

Ganong, Peter, and Pascal Noel. 2020. "Liquidity versus Wealth in Household Debt Obligations: Evidence from Housing Policy in the Great Recession." *American Economic Review* 110(10): 3100–3138. doi:10.1257/aer.20181243.

Gerardi, Kristopher, Paul S. Willen, and David Hao Zhang. 2023. "Mortgage prepayment, race, and monetary policy." *Journal of Financial Economics* 147(3): 498–524.

Glaeser, Edward L, and Charles G Nathanson. 2017. "An Extrapolative Model of House Price Dynamics." *Journal of Financial Economics* 126(1): 147–170.

Green, Richard K., and Michael LaCour-Little. 1999. "Some Truths about Ostriches: Who Doesn't Prepay Their Mortgages and Why They Don't." *Journal of Housing Economics* 8(3): 233–248.

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.

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.

Johnson, Eric J, Stephan Meier, and Olivier Toubia. 2018. "What's the Catch? Suspicion of Bank Motives and Sluggish Refinancing." *Review of Financial Studies* 32(2): 467–495.

Keys, Benjamin J., Devin G. Pope, and Jaren C. Pope. 2016. "Failure to refinance." *Journal of Financial Economics* 122(3): 482–499.

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.

Liu, Lu. 2022. "The Demand for Long-Term Mortgage Contracts and the Role of Collateral." *Working paper.*

MacGee, James, and Yuxi Yao. 2023. "Accounting for the Rise in Mortgage Debt: Lower Inflation and Mortgage Innovations." *Working Paper.*

McConnell, John J., and Manoj Singh. 1994. "Rational Prepayments and the Valuation of Collateralized Mortgage Obligations." *Journal of Finance* 49(3): 891–921.

Mota, Nuno, Mark Palim, and Susan Woodward. 2022. "Mortgages are still confusing... and it matters – How borrower attributes and mortgage shopping behavior impact costs." *Working paper.*

Schwartz, Eduardo S., and Walter N. Torous. 1989. "Prepayment and the Valuation of Mortgage-Backed Securities." *Journal of Finance* 44(2): 375–392.

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.

Vickery, James I, and Joshua Wright. 2013. "TBA Trading and Liquidity in the Agency MBS Market." *Economic Policy Review* 19(1).

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.

Tables and Figures

Table 1: Summary statistics for the Optimal Blue-HMDA-CRISM sample

Panel A: Fixed Characteristics	All		Black		Hispanic	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Loan amount ($'000s)	252688.6	114694.2	231226.7	111223.1	237658.2	107874.4
Credit score	748.5	46.0	724.8	53.3	732.6	46.8
LTV (%)	79.8	15.0	84.5	14.6	81.6	14.4
DTI (%)	34.4	9.4	36.7	8.9	37.7	8.4
Interest rate	4.365	0.499	4.565	0.530	4.546	0.530
Points paid	0.006	0.932	-0.119	0.970	-0.093	0.948
First-time home buyer (d)	0.201	0.400	0.275	0.446	0.277	0.448
Single Female (d)	0.249	0.432	0.449	0.497	0.273	0.446
Single Male (d)	0.322	0.467	0.360	0.480	0.437	0.496
Credit Card Revolver (d)	0.110	0.313	0.132	0.339	0.094	0.292
# Observations	338,338		10,211		25,217	
Panel B: Time-Varying Characteristics						
Mark-to-market LTV (%)	67.9	16.7	71.7	15.5	69.2	16.3
Equifax Risk Score	770.3	70.7	743.5	86.6	752.0	75.5
Mark-to-market LTV >95 (d)	0.0068	0.0822	0.0135	0.1156	0.0104	0.1013
Rate gap	0.1618	0.7365	0.2653	0.7887	0.2826	0.7746
Moved (d)	0.0042	0.0646	0.0029	0.0534	0.0030	0.0550
Refied (d)	0.0093	0.0962	0.0071	0.0839	0.0081	0.0895
# Observations	13,192,408		342,089		863,323	

Notes: This table reports summary statistics from the Optimal Blue-HMDA-CRISM merged sample from January 2013 to December 2019, with performance until May 2022. 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.

Table 2: Choices of points and refinancing/prepayment behavior

	(1) Moved		(2) 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

Note: The data used in this table is the Optimal Blue-HMDA-CRISM data from January 2013 to May 2022, for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2013–2019. This table regression results from estimation Equation (2). Column (1)'s dependent variable is an indicator variable for whether the borrower has moved in a given month multiplied by 100. Column (2)'s dependent variable is an indicator variable for whether the bororwer has refinanced in a given month multiplied by 100. The control variables include the *Call Option* variable of Deng, Quigley, and Van Order (2000) as described in the text, spread of the mortgage interest rate to the Freddie Mac rate at origination (SATO), log of the loan amount, as well as five categories of credit score, four categories of LTV, and a linear control for DTI. CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data.

Table 3: Borrower choices of points and their prepayment behavior by characteristics

	(1)		(2)		(3)	
	Points		5-year prepayment		Non-Refi Borrower	
Black	0.0337	(0.81)	-0.120***	(-5.48)	0.152***	(7.20)
Hispanic	0.0445*	(1.91)	-0.0802***	(-8.43)	0.0872***	(10.64)
Single male	0.000181	(0.01)	-0.00327	(-0.45)	0.0156*	(1.80)
Single female	-0.0287*	(-1.71)	-0.0133*	(-1.84)	0.0181**	(2.18)
First-time home buyer	0.000776	(0.04)	-0.0340**	(-2.46)	0.0371***	(3.03)
Credit card revolver	-0.0287	(-1.60)	0.0225*	(1.81)	0.000864	(0.05)
1st quartile of education	0.00803	(0.40)	-0.00771	(-0.53)	0.0120	(1.02)
2nd quartile of education	-0.0170	(-0.86)	-0.00622	(-0.76)	-0.00266	(-0.29)
3rd quartile of education	0.00614	(0.56)	-0.00520	(-0.57)	0.00141	(0.18)
Log(loan amount)	0.0382**	(2.34)	0.111***	(9.97)	-0.164***	(-16.22)
Credit score controls	Yes		Yes		Yes	
LTV controls	Yes		Yes		Yes	
DTI control	Yes		Yes		Yes	
Constant	-0.414**	(-2.22)	-0.876***	(-6.53)	2.358***	(18.36)
Observations	25245		25245		25245	
LenderXCountyXYear FEs	Yes		Yes		Yes	

Robust t statistics clustered by lender and county in parentheses.

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

Note: The data used in this table is the Optimal Blue-HMDA-CRISM data from January 2013 to May 2022, for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2013–2019. This table regression results from estimation Equation (6). Column (1)'s dependent variable is the number of points paid, with outliers below -4 and above 4 being excluded from the analysis. Column (2)'s dependent variable is whether the borrower prepayed within five years of the mortgage being originated, conditional on the mortgage being originated before April 2016. Column (3)'s dependent variable is whether the borrower did not refinance or otherwise prepay within five years despite having faced a Freddie Mac Survey Rate decrease of at least 1.2%. CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data.

Table 4: Table of Calibrated Parameters

Parameter	Value
γ	2
τ	0.22
b	400
$\bar{\kappa}(M_{it})$	$\$2,000 + 0.01M_{it}$
π^i	0.0168
T	120

Note: Table 4 shows the list of calibrated parameters that is used throughout the model. Of these parameters γ is the coefficient of risk aversion, τ is the tax rate, b is the bequest motive, $\bar{\kappa}(M_{it})$ is the cost of refinancing at zero points, π^i is the expected inflation rate, and T is the number of periods measured in quarters.

Table 5: Additional parameters for an illustrative calibration of cross-subsidization from the perspective of a quick to refinance borrower

Parameter	Value
β_i	0.98
M_i	\$223,784
p_i^m	0.074
κ_i	200
p_i^a	1
Initial liquid assets	\$50,000
Initial risk-free rate	1.0%
Initial mortgage rate	3.25%
Initial income	\$75,000
Initial house price	\$300,000

Note: Table 5 shows the parameters of the model used for the calibration in Section 5.3. β_i refers to the discount factor, γ the coefficient of risk aversion, M_i the mortgage size, p_i^m the moving probability, κ_i the fixed component of refinancing costs, and p_i^a the time-varying ability of a borrower to refinance.

Table 6: Estimated model parameters and their standard errors

Parameter	Value	Standard Error
μ_{p^a}	-1.590***	(0.055)
σ_{p^a}	0.665***	(0.058)
μ_κ	6.716***	(0.126)
σ_κ	0.566***	(0.079)
μ_{p^m}	-2.673***	(0.085)
σ_{p^m}	0.242***	(0.047)
μ^β	2.500***	(0.197)
σ_β	0.154	(0.096)
ρ_{β,p^a}	-0.072	(0.107)
$\rho_{\beta,\kappa}$	0.012	(0.055)
ρ_{β,p^m}	0.264***	(0.101)
$\mu^b_{p^a}$	-0.145*	(0.076)
μ^b_κ	0.703***	(0.106)
$\mu^b_{p^m}$	-0.141	(0.115)
$\mu^h_{p^a}$	0.024	(0.067)
$\mu^h_{p^m}$	0.110*	(0.065)
μ^h_κ	-0.056	(0.086)
FICO controls	Yes	
LTV controls	Yes	
Loan amount controls	Yes	
Non-bank control	Yes	

Note: Table 6 shows the parameters of the model estimated from maximum likelihood as in Equation (29). μ_{p^a} and σ_{p^a} refers to the mean and standard deviation of the Log-Normal distribution of the probability that a borrower is able to refinance. μ_κ and σ_κ refers to location and scale parameter of the Log-Normal distribution of the borrower's refinancing hassle costs. μ_{p^m} and σ_{p^m} refers to the mean and standard deviation of the Log-Normal distribution of the probability that the borrower moves. μ_β and σ_β refers to the mean and standard deviation of the Logit-Normal distribution of the borrower's discount factors. $\rho_{\beta,p^a}, \rho_{\beta,\kappa}, \rho_{\beta,p^m}$ denotes the correlation between the borrower's discount factors and their refinancing/moving. Standard errors are derived from the inverse Hessian. * p<0.1, ** p<0.05, *** p<0.01

Table 7: Implied cross-subsidization received based on model estimates

Borrower Type	Implied cross-subsidization received ($)
$p_i^a = 0$	–3634
$p_i^a = 1, \kappa_i = 200$	+4537
Black	–1983
Hispanic	–387
Others	+77

Note: Table 7 shows the implied cross-subsidization received by each borrower type based on the heterogeneous agent life-cycle model. These implied cross-subsidies are measured in dollar terms from the lenders' perspective. They are computed in an ex-ante terms rather than taking the path of interest rates as given. The results show that, in expectation, the difference between non-refinancing borrowers with $p_i^a = 0$ and actively refinancing borrowers with $p_i^a = 1, \kappa_i = 200$ is \$8,171. Black borrowers pay \$1,983 in cross-subsidies ex ante, while Hispanic borrowers pay \$387.

Table 8: Welfare change with alternative contract designs, change in borrower utility ($) relative to existing equilibrium

	(1) Switching from adding to rate to adding to to balance	(2) Automatically refinancing
Overall	+1461	+1362
$p_i^a = 0$	+2819	+1498
$p_i^a = 1, \kappa_i = 200$	–1310	–2016
Black	+2642	+1863
Hispanic	+2478	+834
Others	+1263	+1386
Change in initial rate	+0.019	+0.196

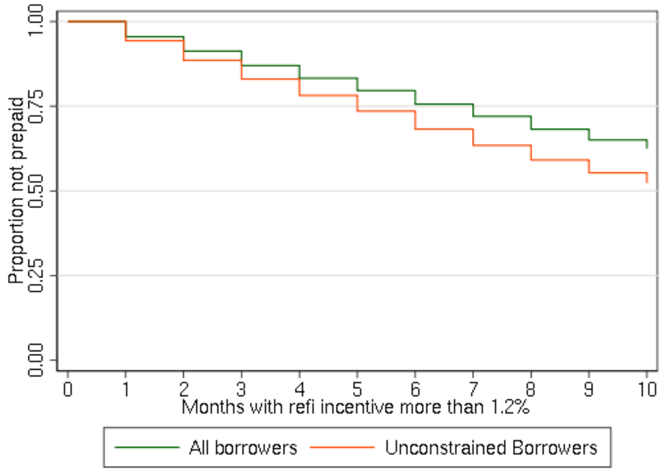
Note: Table 8 shows welfare comparison of the existing market equilibrium to the counterfactual of (1) adding closing costs to the balance of the loan and (2) automatically refinancing mortgages for 2013–2019 originations. Welfare is quantified in terms of the change in borrower liquid savings to make them indifferent between the existing market equilibrium and the alternative contract designs.

Figure 1: Rate and points 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)

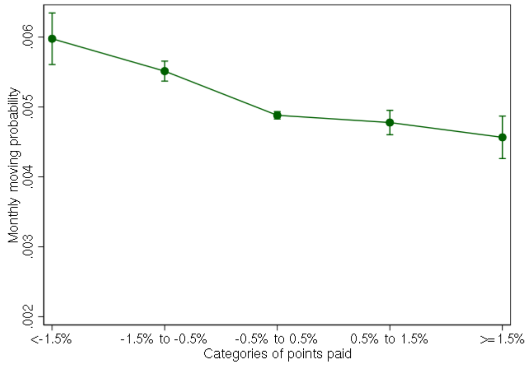
Note: Figure 1 shows a set of rate and points available to borrowers from an example wholesale lender ratesheet. The first column indicates the rate, while the next three columns shows the amount of points, in the form of points/percentages of the loan amount, the borrower would have to pay to lock the rate for 15, 30, or 45 days, respectively. Negative points are also possible in order to cover the other costs associated with the real estate transaction the borrowers might have to pay. Figure A.1 shows an example of how a price comparison website displayed the series of rate and upfront closing cost choices.

Figure 2: Kaplan-Meier survival hazards with months of interest rate incentive being greater than 1.2%

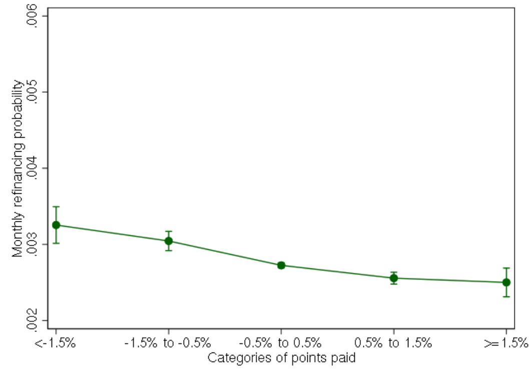


Note: The data used in this figure is the Optimal Blue-HMDA-CRISM data for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2013–2019, with performance until May 2022. The green line Figure 2 presents the Kaplan-Meier survival estimates of prepayment for mortgages with a refinancing incentive, here defined as a Freddie Mac survey rate decrease, of greater than or equal to 1.2%. The red line in Figure 2 shows the result of the same analysis among borrowers with an Equifax Risk Score that is above 700 and an estimated loan-to-value ratio of below 80% throughout the sample, which is a group of borrowers who are unlikely to face supply-side constraints in refinancing. CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data.

Figure 3: Moving/refinancing probability by points paid



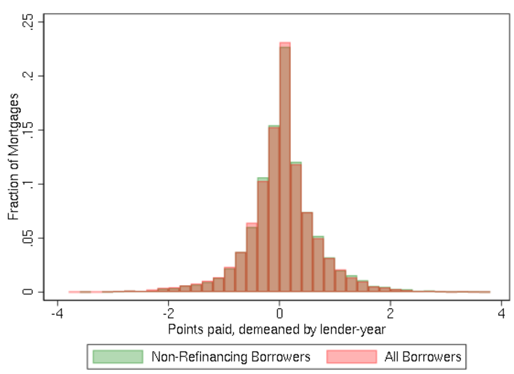
(a) Moving probability by points



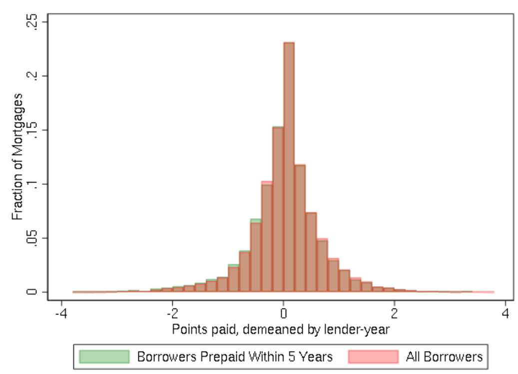
(b) Refinancing probability by points

Note: The data used in this figure is the Optimal Blue-CRISM data for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2013–2019, with performance until May 2022. CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data. Figure 3a presents the predicted probabilities in regressions of moving on control variables, while Figure 3b presents the predicted proabilities in regressions of refinancing on control variables. The regression estimates that these results were based on are presented in Table 2.

Figure 4: Points paid by borrower ex-post prepayment behavior



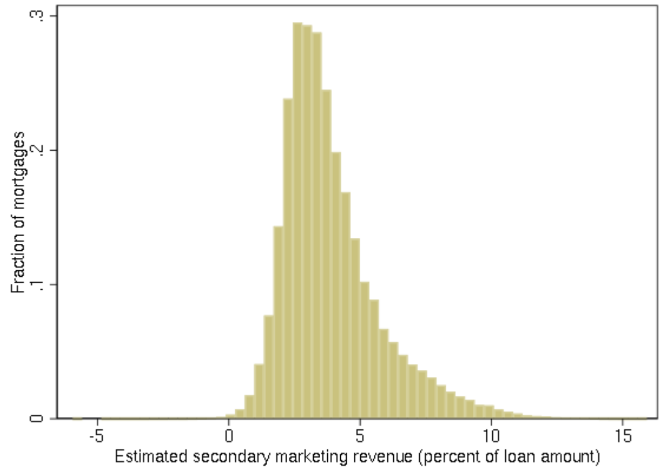
(a) Refinancing behavior



(b) Prepayment behavior

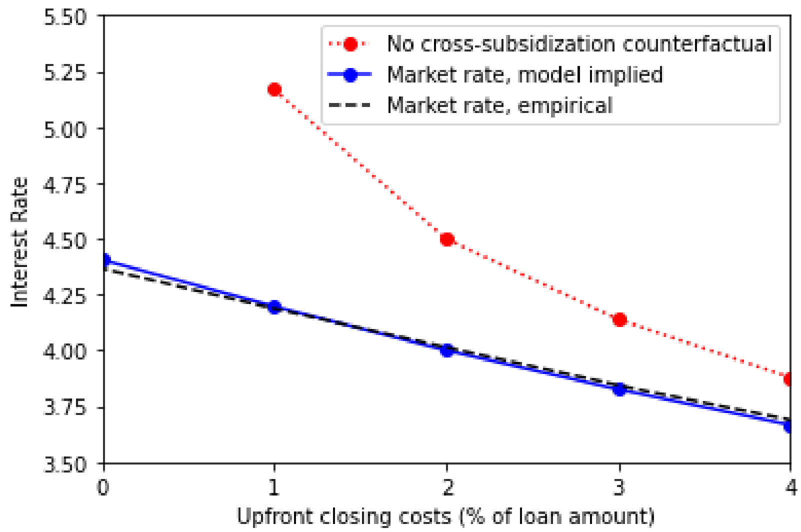
Note: The data used in this figure is the Optimal Blue-HMDA-CRISM data from January 2013 to May 2022, for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2013–2019. Figure 4a presents a histogram of borrower choices of points demeaned by lender by county by year groups, comparing between non-refinancing borrowers (defined as borrowers who did not refinance or otherwise prepay within five years despite facing a Freddie Mac Survey Rate decrease of at least 1.2%) and all borrowers who faced a Freddie Mac Survey Rate decrease of at least 1.2%. Figure 4b conducts the same analysis comparing borrowers who prepaid within five years versus all mortgages that have been originated for at least five years. CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data.

Figure 5: Price of mortgage origination that is added to the rate of the mortgage



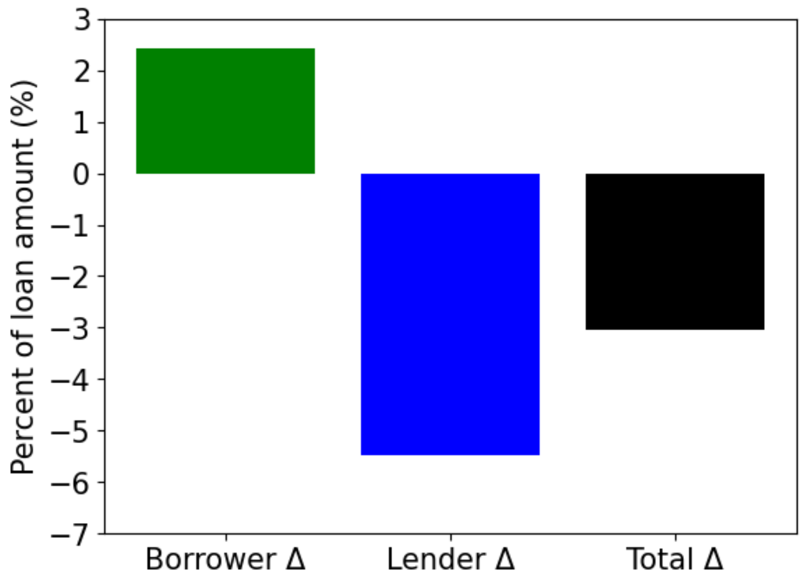
Note: Figure 5 plots histograms of estimated lender revenue that comes from secondary marketing revenue. The data used in this figure is the Optimal Blue-HMDA-CRISM data from January 2013 to May 2022, for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2013–2019. The data contains information on rates and net origination charges paid upfront and was linked to MBS TBA data following Fuster, Lo, and Willen (2022) to estimate secondary marketing revenue.

Figure 6: Market interest rate vs no cross-subsidization counterfactual interest rate for the calibrated quick to refinance borrower



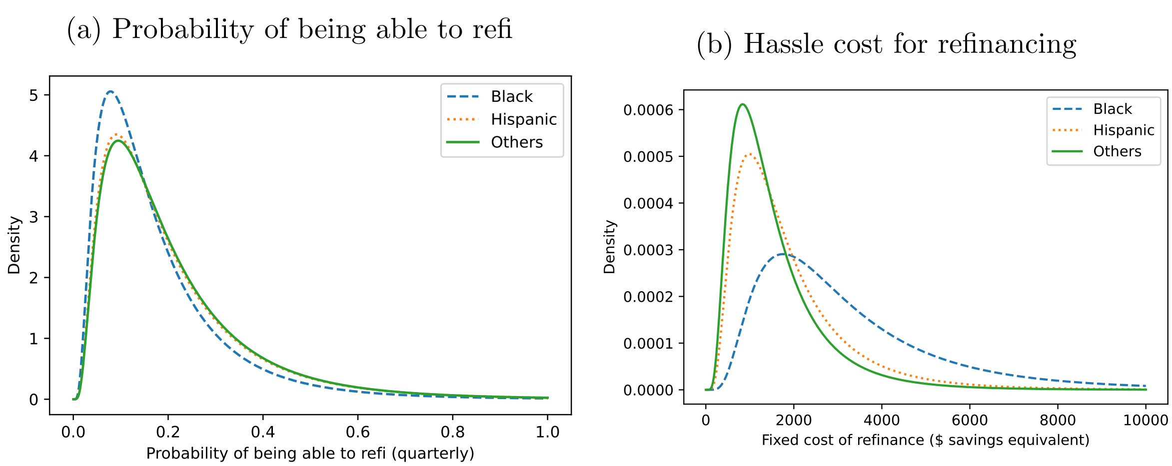
Note: Figure 6 presents the equilibrium rate and upfront closing costs trade-off from the model and compares it to the empirical rate and upfront closing costs trade-off that I estimate from the data. The "Market rate, model implied" solid line refers to the average equilibrium rate and closing cost trade-off given our model for the period of 2013–2019. The "Market rate, empirical" dotted line is the average implied rate and upfront closing cost options from MBS TBA prices combined with the PMMS survey rate for the period of 2013–2019. I show that the rate and upfront closing cost options implied by MBS TBA prices is consistent with the options actually presented to borrowers in rate sheets from a regression with rate sheet fixed effects in Appendix A.3. Finally, the "No cross-subsidization counterfactual" dashed line refers to the model-implied equilibrium rate and closing cost trade-off in a world where the lender is pricing their mortgages for the calibrated quick to refinance borrower with perfect information on their type. This counterfactual was computed by jointly iterating on the borrower and lender's value functions.

Figure 7: Welfare relative to no cross-subsidization counterfactual, calibrated quick to refinance borrower



Note: Figure 7 plots (i) the upfront cash the calibrated quick to refinance borrower would have to receive to remain indifferent in the no cross-subsidization counterfactual, (ii) the upfront cash the lender's difference in profit from the quick to refinance borrower's loan between the current world and the no cross-subsidization counterfactual, and (iii) the sum of (i) and (ii). The results suggests that under the current system, the quick to refinance borrower gains 2.4% of loan amount in dollar terms, whereas lender loses 5.5% of the loan amount in profit, with a total social loss of 3.0% of the loan amount.

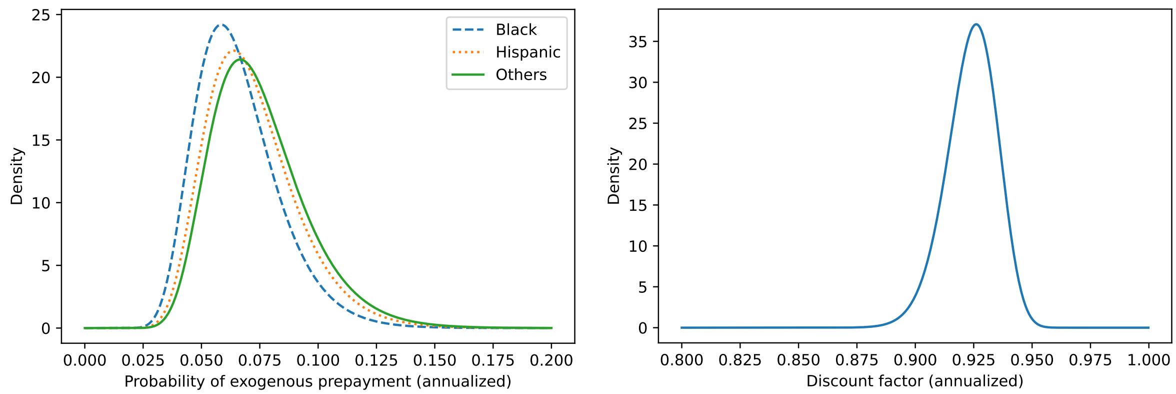
Figure 8: Distribution of borrower refinancing tendencies



Note: Figure 8a plots the estimated density for the probability of being able to refinance coming from the marginal of the multivariate Log-Normal distribution of Equation (23). Figure 8b plots the estimated density for the hassle cost of refinancing from the Log-Normal distribution of Equation (24). The density of borrower types is computed by taking a weighted sum of the probability density functions for all categories of covariates X_i.

Figure 9: Moving probability and discount factor

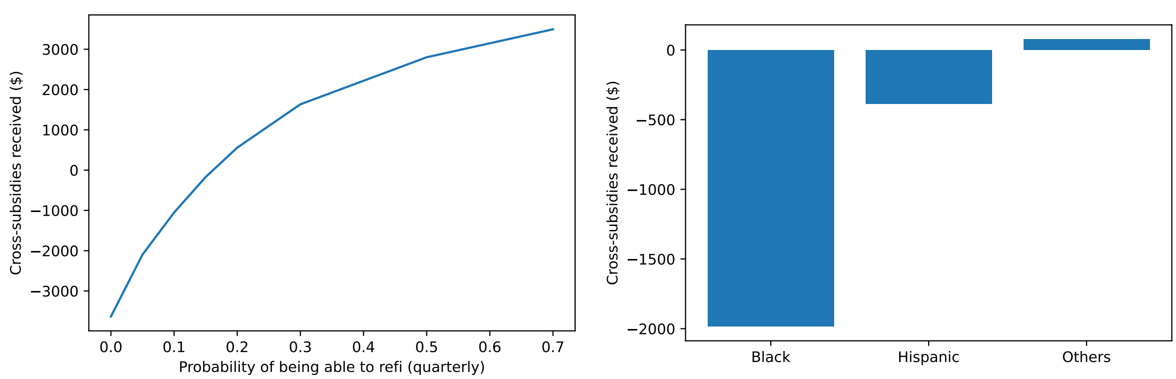
(a) Moving probability (b) Discount factor



Note: Figure 9a plots the estimated density of moving probabilities across borrower types from the Log-Normal distribution of Equation (25). Figure 9b plots the estimated density for the discount factor coming from the marginal of the Logit-Normal distribution of Equation (26). The density of borrower types is computed by taking a weighted sum of the probability density functions for all categories of covariates X_i.

Figure 10: Magnitude of cross-subsidies within the population

(a) By borrower refinancing ability (b) By borrower race/ethnicity

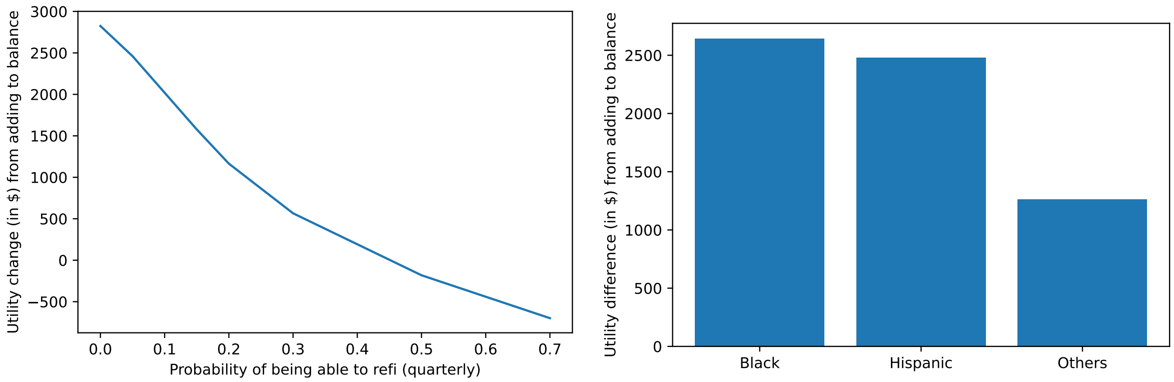


Note: Figure 10 plots the average cross-subsidy received by borrowers of a given type. Figure 10a shows the results by borrower refinancing ability, and Figure 10b shows the results by borrower race.

Figure 11: Counterfactual utility change under adding more of the price of mortgage origination to balance

(a) By borrower refinancing ability

(b) By borrower race/ethnicity

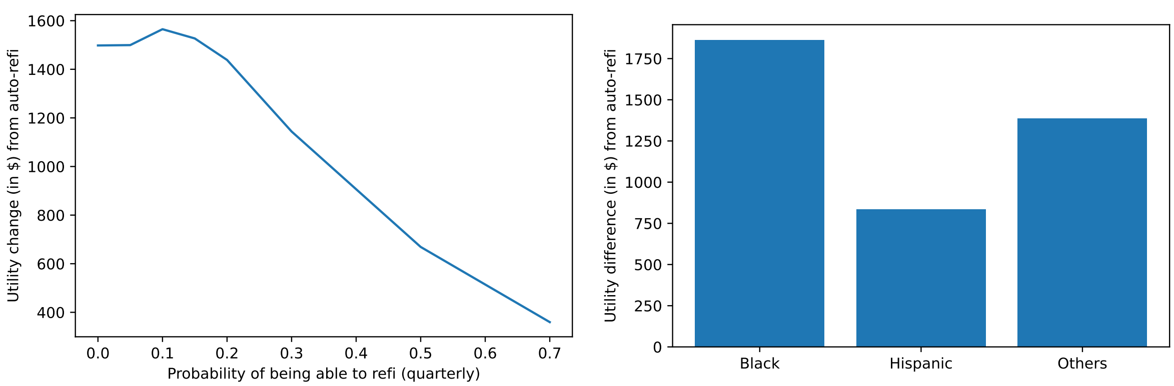


Note: Figure 11 plots the average upfront dollar savings that would make borrowers indifferent between the existing system and what they would otherwise obtain in the counterfactual of adding the portion of the price of mortgage origination that is added to the rate to the balance instead, as described in Section 7.1. Figure 11a shows the results by borrower refinancing ability, and Figure 11b shows the results by borrower race.

Figure 12: Counterfactual utility change under automatically refinancing mortgages

(a) By borrower refinancing ability

(b) By borrower race/ethnicity



Note: Figure 12 plots the average upfront dollar savings that would make borrowers indifferent between the existing system and what they would otherwise obtain in the automatically refinancing counterfactual as described in Section 7.2. Figure 12a shows the results by borrower refinancing ability, and Figure 12b shows the results by borrower race.

Internet Appendix

This appendix supplements the empirical analysis of this paper. Below is a list of the sections contained in this appendix.

Table of Contents

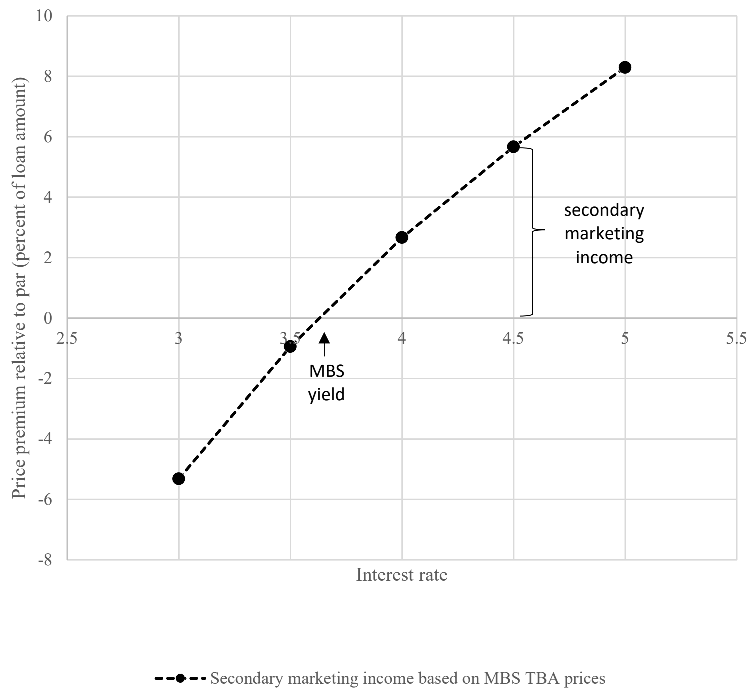
A.1 Additional Background About Rate and Upfront Closing Costs

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

Note: Figure A.1 shows a screenshot obtained by the author from Bankrate.com for a $250,000 refinancing mortgage on September 18, 2021. It shows how a borrower may choose to pay 0 points for a 2.740% interest rate mortgage, 0.626 points for a 2.615% interest rate mortgage, or 1.5 points for a 2.490% interest rate mortgage.

Figure A.2: Secondary marketing income as a function of interest rates



Note: Figure A.2 plots the FNMA MBS TBA prices on January 2, 2014 expressed as a percentage point premium/discount over the loan amount on the y-axis for a variety of coupon rates on the x-axis. Secondary marketing income is the extent to which the secondary market value of the mortgage is above its principal balance.

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 ID 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 ID 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 conversation with representatives from Optimal Blue.

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

	All		Black		Hispanic	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Loan amount ($'000s)	256695.6	117785.7	242574.6	117351.2	243938.2	112333.0
Origination cost ($)	1516.0	1807.2	1657.6	2062.3	1849.0	1969.4
Total loan cost ($)	3902.6	2362.4	4222.6	2713.2	4487.1	2547.2
Credit Score	747.9	44.5	728.7	47.3	732.7	45.5
LTV (%)	80.4	15.0	84.9	13.5	82.6	14.7
DTI (%)	34.973	9.681	37.363	8.849	38.239	8.600
Interest rate	4.544	0.579	4.692	0.612	4.674	0.603
Points paid	0.307	0.461	0.395	0.489	0.387	0.487
First-time home buyer (d)	0.252	0.434	0.436	0.496	0.268	0.443
Single Female (d)	0.330	0.470	0.356	0.479	0.441	0.497
# Observations	1,041,807		42,793		92,598	

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 May 2022. 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 2,563,550 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

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

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 Canner (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 is shown in Table 1.

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 (2022). 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.2.

Table A.2: Summary statistics for the LoanSifter data

Year	No. of Lenders	Rate at -2 points	Rate at 0 points	Rate at 2 points	N lender-days obs
2009	93	5.42	5.01	4.65	3923
2010	93	5.10	4.70	4.44	16025
2011	83	4.82	4.46	4.25	16589
2012	86	4.07	3.67	3.41	18105
2013	126	4.42	4.07	3.80	19993
2014	103	4.52	4.21	3.97	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 Pass-through of TBA prices to lender rate sheets

I examine how the secondary marketing income-interest rate trade-off matches the retail interest rate and upfront closing costs trade-off on average in the cross-section, with results in Figure A.3. I use data LoanSifter matched with MBS TBA pricing data from 2009Q3 to 2014. Following the methodology of Fuster, Lo, and Willen (2022), I focus on borrowers with a $300k conforming mortgage, 700 LoanSifter credit score, 80% LTV, and 30% DTI. I estimate (i) the secondary marketing revenue generated by lenders in as implied by MBS TBA prices, and compare that with (ii) the sum of the secondary marketing revenue and the upfront closing costs they charge in the form of points. The secondary marketing revenue generated by lenders in as implied by MBS TBA prices is estimated using Equation (7), with the *Payup* set to zero due to the \$300k loan amount.

Then, with the interest rate spread to the Freddie Mac Primary Mortgage Market Survey (PMMS) rate[3] rounded to the nearest 1/8th $\tilde{c}$, I run a linear regressions of the form:

$$\phi_{ijt} = \sum_{l=1}^{N} \gamma_l \mathbb{1}(c = c_l) + \xi_{jt} + \epsilon_{ijt}, \tag{32}$$

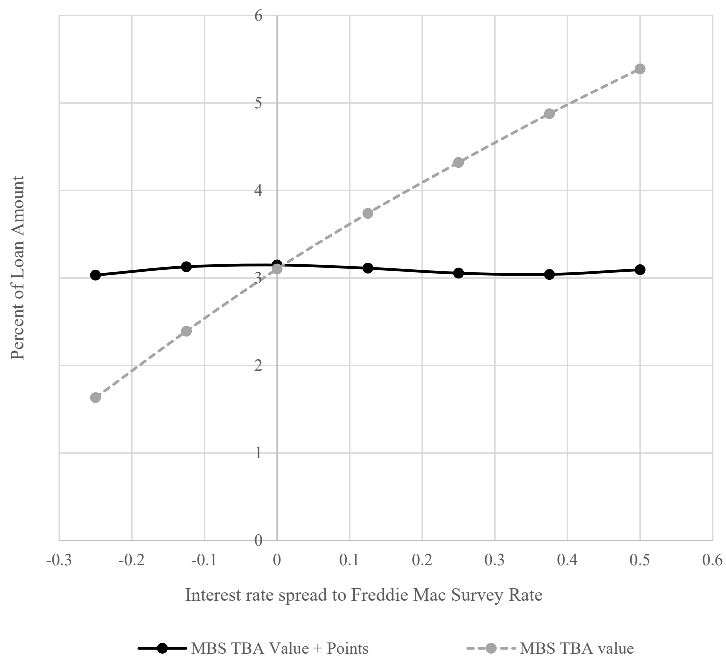
where c_l are the categorical variables of interest rate spread rounded to the nearest 1/8th, ξ_{jt} are lender-day fixed effects, and ϵ_{ijt} is the error term. ϕ_{ijt} is either the secondary marketing revenue generated the lender or sum of the revenue generated by lenders in the secondary market and the upfront closing costs in the form of points, both expressed as a percentage of the loan amount.

The predicted values of Equation (32) are plotted in Figure A.3, 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.[4] 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.

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

[4] The same patterns also exist in the time series, as I illustrate in Appendix Figure A.4. In Figure A.4, 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.4.

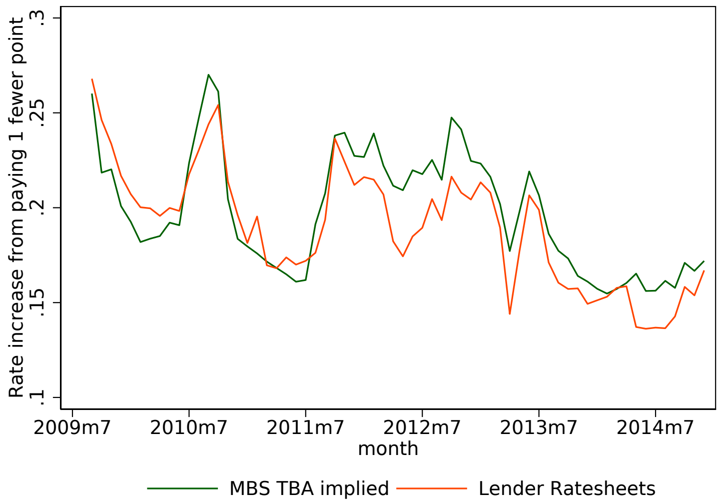
Figure A.3: Secondary marketing income and total lender revenues



Note: Figure A.3 presents estimates from a linear regression of (i) the estimated secondary marketing revenue as implied by MBS TBA prices and (ii) the sum of estimated secondary marketing revenue as implied by MBS TBA prices and upfront closing costs in the form of points on categorical variables of eighths of rate spreads with lender-day fixed-effects based on Equation (32). The grey dotted line plots the predicted values from the regression with estimated secondary marketing revenue as the regressor. The black solid line plots the predicted values from the same regression on the sum of estimated secondary marketing revenue and upfront closing costs in the form of points.

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.4. 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 from lender rate sheets 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.4.

Figure A.4: The interest rate increase from paying 1 less point in upfront closing cost over time, lender ratesheets (red) versus MBS TBA implied (green)



Note: Figure A.4 presents estimates from taking monthly means of (i) the required increase in rate to make the mortgage value increase by 1% of the loan amount in the MBS TBA data (ii) the increase in rate going from 0 points in lender rate sheet to 1 point in lender rate sheet in terms of upfront closing costs paid. The data used is Morgan Markets for MBS TBA prices and LoanSifter for rate sheets. MBS TBA values are linearly interpolated in between eighths of interest rates and LoanSifter rates are linearly interpolated to arrive at the rate at 0 and 1 point in upfront closing costs.

Figure A.4 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 upfront closing cost mortgages in the retail market as compared to the secondary market near the end of the sample.

A.4 Additional motivating facts

In this section, I present some additional stylized facts that illustrate the existence of cross-subsidization of mortgage closing costs and its sizable distributional implications. First, I show in Section A.4.2 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 heterogenous borrower prepayment tendencies implies 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 post, in Section A.4.3. Third, I assess magnitude of this difference by demographic groups in Section A.4.4.

A.4.1 Regression of choices of points as predicted by ex-post prepayment behavior

Table A.3: Choices of points as it relates to refinancing/prepayment behavior

	(1)		(2)	
	Points		Points	
Non-refi borrower	0.0659***	(5.32)		
5-year prepayment			-0.0841***	(-6.30)
Log(loan amount)	0.0511***	(2.71)	0.0497***	(2.73)
Credit score controls	Yes		Yes	
LTV controls	Yes		Yes	
DTI control	Yes		Yes	
Constant	-0.600***	(-2.83)	-0.519**	(-2.59)
Observations	25245		25245	
LenderXCountyXYear FEs	Yes		Yes	

Robust t statistics clustered by lender and county in parentheses.

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

Note: The data used in this figure is the Optimal Blue-HMDA-CRISM data from January 2013 to May 2022, for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2013–2019. The sample for (1) and (2) is further restricted to the set of borrowers whose mortgages originated before April 2016 and where the Freddie Mac Survey Rate decreased at least 1.2% since origination. Table A.3 presents OLS estimates of borrower choices of points on (1) an indicator variable for non-refinancing borrowers, defined as borrowers who did not refinance or otherwise prepay within five years despite facing a Freddie Mac Survey Rate decrease of at least 1.2%, and (2) borrowers who prepaid within five years.

A.4.2 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 charges and total loan 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 expected to generate from the secondary market at the time the rate is set.

I first estimate the extent to which mortgage closing costs are added to the rate based on Equation (1). This statistic has a mean of 3.49%. I then compute the lender revenue from origination as the sum of y_{it} and the net loan origination charges from 2018-2019 HMDA data. The fraction of y_{it} to the sum of y_{it} and the net loan origination charges represents the share of lender revenue from secondary marketing income.[5]

The results of my analysis are shown in Figure A.5. The left panel in Figure A.5a 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. Relative 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.[6]

The right panel of Figure A.5b shows that only a small fraction of lender revenue is paid as

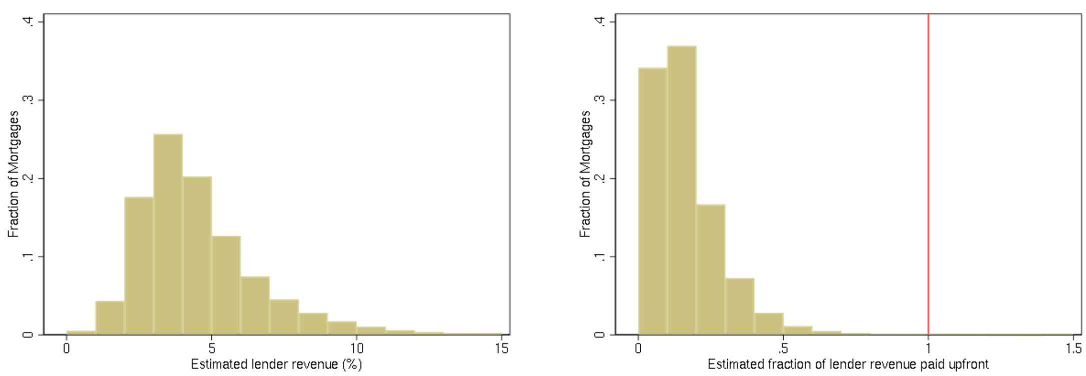
[5] A drawback of this approach is that it assumes that mortgage servicing has zero marginal cost, or that mortgage servicing rights have value equal to their expected stream of cash flows. Without explicit data on the value of mortgage servicing rights relative to their cash flow value, I compute a 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.6, which still shows that the majority of mortgages still have most of their price of origination paid for through the rate.

[6] https://www.mba.org/2019-press-releases/april/independent-mortgage-bankers-production-volume-and-profits-down-in-2018. The MBA 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.

upfront net origination charges, with an average of 17.4%. That is, even though most of the lender costs of origination are incurred upfront, 82.6% 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. Figures A.7 and A.8 repeats this exercise for purchase and refinance mortgages separately, and finds patterns that are broadly consistent.

Finally, Figures A.9, A.10, and A.11 repeats this exercise but with total loan costs in place of net origination charges. Total loan costs include appraisal and title search fees that may not go to the lender, though some lenders may generate revenue from part of these fees. But, even with this broader measure of upfront lender revenue only 31.4% of the price of mortgage origination is paid upfront on average.

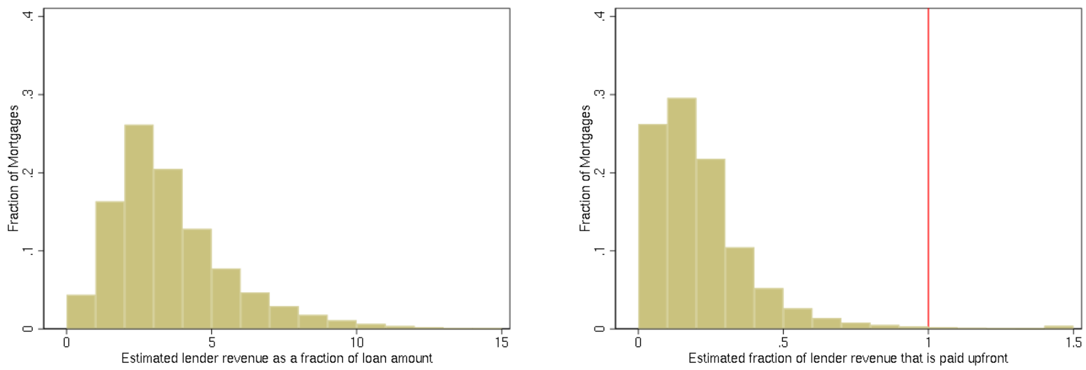
Figure A.5: Lender revenue and percentage paid as upfront net origination charges



(a) Estimated lender revenue (b) Fraction of lender revenue paid upfront

Note: The data used in this figure is the 2018–2019 Optimal Blue-HMDA data for 30-year, fixed-rate, conforming, primary residence mortgages. The data contains information on rates and net origination charges paid upfront and was linked to MBS TBA data following Fuster, Lo, and Willen (2022) to estimate secondary marketing revenue. Figure A.5a plots histograms of estimated lender revenue which consists of the sum of net origination charges and secondary marketing revenue. Figure A.5b then plots histograms of the fraction of lender revenue that is consists of net origination charges. The net origination charges used in this figure equals to the HMDA origination charges minus lender credit, and is set to zero for mortgages with less than zero net origination charges.

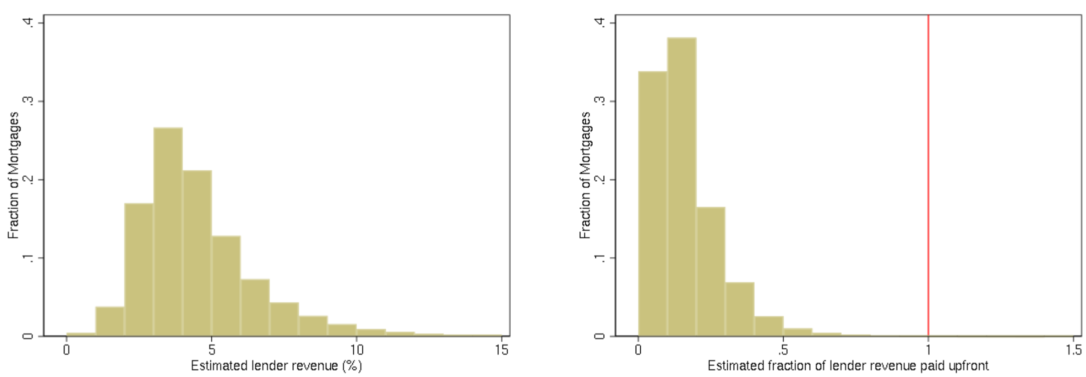
Figure A.6: Lender revenue and percent paid as upfront net origination charges, net of mortgage servicing revenue



(a) Estimated lender revenue (b) Fraction of lender revenue paid upfront

Note: The data used in this figure is the Optimal Blue data for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2018-2019 matched to the 2018–2019 HMDA data. The data contains information on rates and net origination charges paid upfront and was linked to MBS TBA data following Fuster, Lo, and Willen (2022) to estimate secondary marketing revenue. A further 25 basis points was subtracted from the coupon rate for mortgage servicing. Figure A.6a plots histograms of estimated lender revenue which consists of the sum of net origination charges and secondary marketing revenue. Figure A.14b then plots histograms of the fraction of lender revenue that is consists of net origination charges.

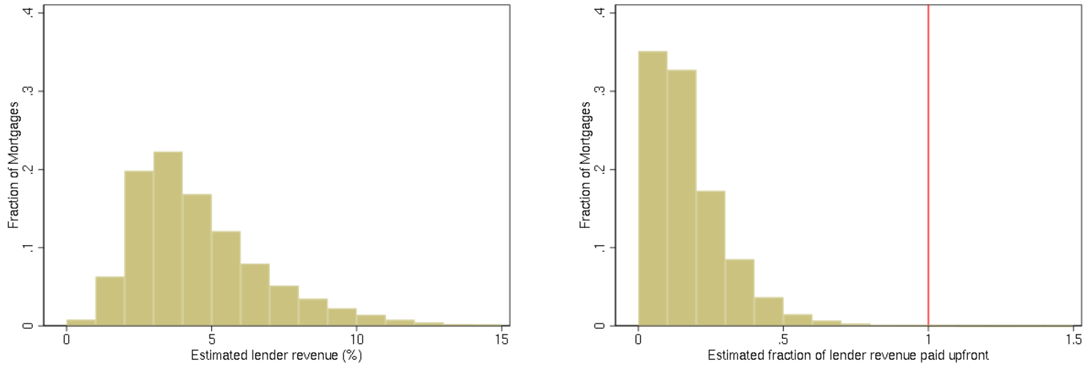
Figure A.7: Lender revenue and percentage paid as upfront net origination charges, purchase



(a) Estimated lender revenue (b) Fraction of lender revenue paid upfront

Note: The data used in this figure is the 2018–2019 Optimal Blue-HMDA data for 30-year, fixed-rate, conforming, primary residence purchase mortgages. The data contains information on rates and net origination charges paid upfront and was linked to MBS TBA data following Fuster, Lo, and Willen (2022) to estimate secondary marketing revenue. Figure A.7a plots histograms of estimated lender revenue which consists of the sum of net origination charges and secondary marketing revenue. Figure A.10b then plots histograms of the fraction of lender revenue that is consists of net origination charges. The net origination charges used in this figure equals to the HMDA origination charges minus lender credit, and is set to zero for mortgages with less than zero net origination charges.

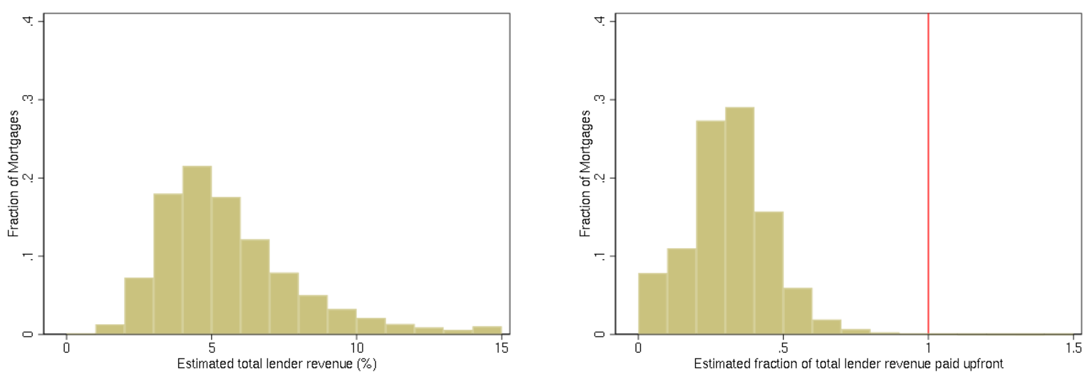
Figure A.8: Lender revenue and percentage paid as upfront net origination charges, refinance



(a) Estimated lender revenue (b) Fraction of lender revenue paid upfront

Note: The data used in this figure is the 2018–2019 Optimal Blue-HMDA data for 30-year, fixed-rate, conforming, primary residence refinance mortgages. The data contains information on rates and net origination charges paid upfront and was linked to MBS TBA data following Fuster, Lo, and Willen (2022) to estimate secondary marketing revenue. Figure A.8a plots histograms of estimated lender revenue which consists of the sum of net origination charges and secondary marketing revenue. Figure A.11b then plots histograms of the fraction of lender revenue that is consists of net origination charges. The net origination charges used in this figure equals to the HMDA origination charges minus lender credit, and is set to zero for mortgages with less than zero net origination charges.

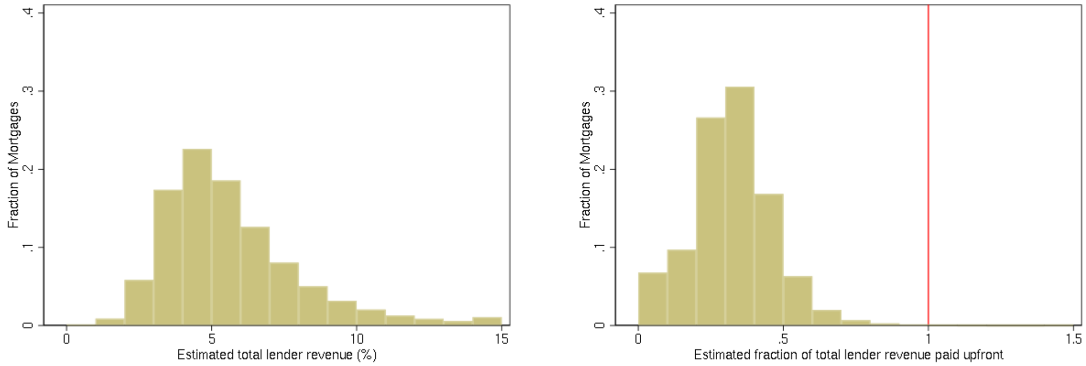
Figure A.9: Total lender revenue and percentage paid as total loan costs



(a) Estimated lender revenue (b) Fraction of lender revenue paid upfront

Note: The data used in this figure is the 2018–2019 Optimal Blue-HMDA data for 30-year, fixed-rate, conforming, primary residence mortgages. The data contains information on rates and total loan costs paid and was linked to MBS TBA data following Fuster, Lo, and Willen (2022) to estimated secondary marketing revenue. Figure A.9a plots histograms of estimated lender revenue which consists of the sum of total loan costs plus secondary marketing revenue. Figure A.9b then plots histograms of the fraction of lender revenue that is paid upfront. The total loan costs used in this figure equals to the HMDA total loan costs minus lender credit, and is set to zero for mortgages with less than zero total loan costs.

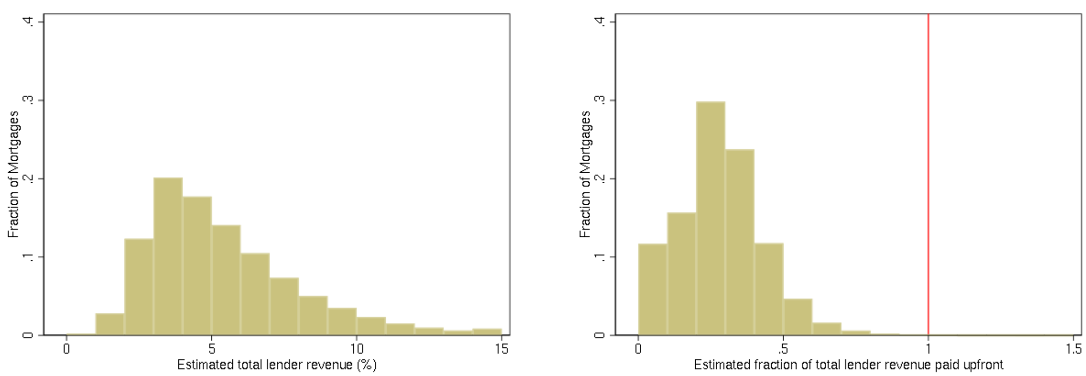
Figure A.10: Total lender revenue and percentage paid as total loan costs, purchase



(a) Estimated lender revenue (b) Fraction of lender revenue paid upfront

Note: The data used in this figure is the 2018–2019 Optimal Blue-HMDA data for 30-year, fixed-rate, conforming, primary residence purchase mortgages. The data contains information on rates and total loan costs paid and was linked to MBS TBA data following Fuster, Lo, and Willen (2022) to estimated secondary marketing revenue. Figure A.10a plots histograms of estimated lender revenue which consists of the sum of total loan costs plus secondary marketing revenue. Figure A.10b then plots histograms of the fraction of lender revenue that is paid upfront. The total loan costs used in this figure equals to the HMDA total loan costs minus lender credit, and is set to zero for mortgages with less than zero total loan costs.

Figure A.11: Total lender revenue and percentage paid as total loan costs, refinance



(a) Estimated lender revenue (b) Fraction of lender revenue paid upfront

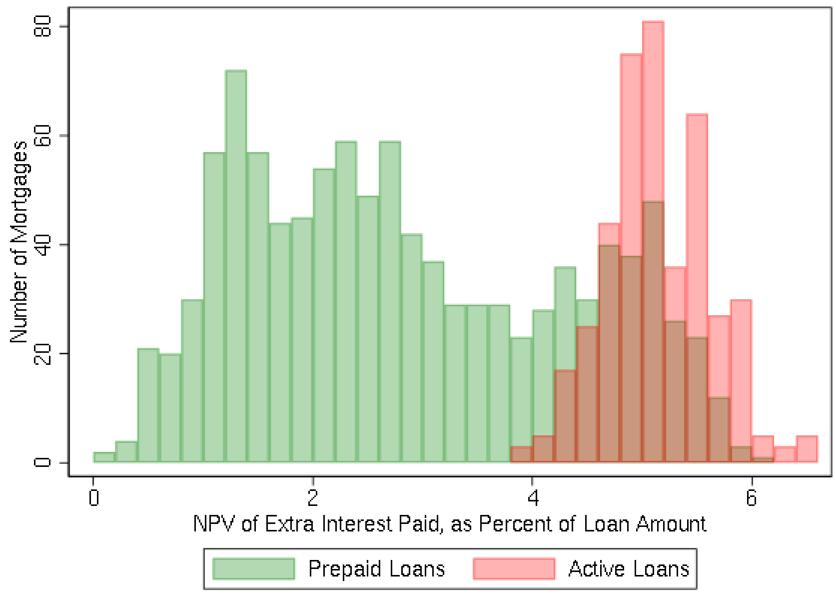
Note: The data used in this figure is the 2018–2019 Optimal Blue-HMDA data for 30-year, fixed-rate, conforming, primary residence refinance mortgages. The data contains information on rates and total loan costs paid and was linked to MBS TBA data following Fuster, Lo, and Willen (2022) to estimated secondary marketing revenue. Figure A.11a plots histograms of estimated lender revenue which consists of the sum of total loan costs plus secondary marketing revenue. Figure A.11b then plots histograms of the fraction of lender revenue that is paid upfront. The total loan costs used in this figure equals to the HMDA total loan costs minus lender credit, and is set to zero for mortgages with less than zero total loan costs.

A.4.3 Cross-subsidization of the price of mortgage origination when they are added to the rate

The interaction of the heterogeneity in refinancing tendencies and a component of the price of mortgage origination being added to the rate implies a cross-subsidization of the price of mortgage origination to the extent they are added to the rate. To illustrate this in my data, Figure A.12 looks at borrowers with similar amounts of the price of mortgage origination added to the rate, between 4.75-5.25% of the loan amount, 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.[7] Due to differences in prepayment behavior, I find large differences in how much borrowers end up paying for the same 4.75-5.25% of the loan amount in the price of mortgage origination added to the rate, ranging from close to 0% to more than 6%.

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

Figure A.12: NPV of extra interest paid, 2013 mortgages with 4.75–5.25% of the loan amount in the price of mortgage origination added to the rate



Note: The data used in this figure is the Optimal Blue-HMDA-CRISM data originated between January 2013 and December 2013, with performance to May 2022, for 30-year, fixed-rate, conforming, primary residence mortgages. The sample was further limited to mortgages with a secondary marketing revenue of 4.75–5.25% of the loan amount, as estimated based on MBS TBA prices following Fuster, Lo, and Willen (2022). The extra interest paid is relative to a mortgage with 0% secondary marketing revenue (i.e., originated at par) and is estimated as the difference between the mortgage interest rate at origination net of the fee for government guarantee (gfees) and the MBS TBA yields at the time of lock. The NPV of the extra monthly payment resulting from this difference in extra interest paid is then computed assuming a discount rate equal to the 10-year Treasury rate at the time of the rate lock and plotted in the histogram for loans that have prepaid (in green) and for loans that are still active (in red). CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data.

The reason for the variance in outcomes in Figure A.12 is that, when the price of mortgage origination is added to the rate of the mortgage, lenders can only recover their price of mortgage origination 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 price of mortgage origination in expectation when they are added to the rate by receiving cross-subsidization from other borrowers. Appendix Section A.4.4

examines the variation in this cross-subsidization by borrower demographics.

A.4.4 The predictability of cross-subsidization by demographics

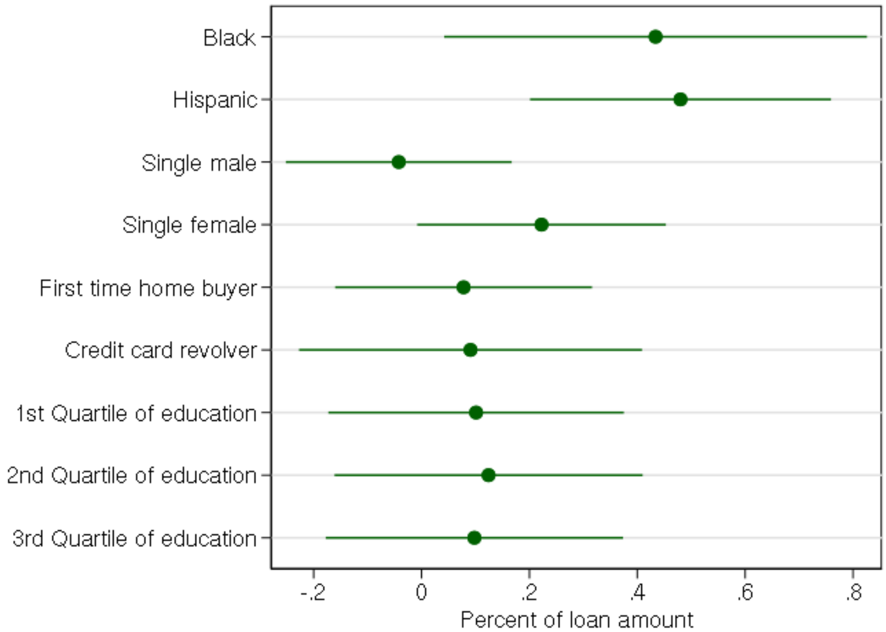
In this section I examine the extent of this ex-post cross-subsidization by demographics. To do so, I run the regression on loan level data in my Optimal Blue-HMDA-CRISM sample:

$$NPV_{i,t} = \beta X_i + \gamma Z_i + \xi_{\phi_{i,t} \times t} + \epsilon_{i,t} \tag{33}$$

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 loan amount; $\xi_{\phi_{i,t} \times t}$ is the amount of closing costs added to the rate by time fixed effects.

The results of this analysis are shown in Figure A.13 and Table A.4. 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 A.13: NPV of extra interest paid by demographic and borrower characteristics



Note: The data used in this figure is the Optimal Blue-HMDA-CRISM data from January 2013 to December 2013, for 30-year, fixed-rate, conforming, primary-residence mortgages originated in 2013. The graph plots regression coefficients from Column (2) of Table A.4. In particular, it shows that Black, Hispanic and single-applicant female borrowers pay more for their closing costs added to the rate than other borrowers. Other characteristics, such as single-applicant male borrowers, first-time home buyers, credit card revolvers (defined as someone with a more than 60% credit utilization and $10,000 in debt at the time of getting a mortgage), and quartiles by education are not statistically different from zero at the 5% level. CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data.

Table A.4: 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

Note: The data used in this table is the Optimal Blue-HMDA-CRISM data from January 2013 to December 2013, for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2013. This table contains regression results from estimating Equation (7). The dependent variable is the NPV of extra interest paid from the closing costs that are added to the rate. I include ϕ by month fixed effects, where ϕ refers to the amount of closing costs added to the rate rounded to the nearest percent of the loan amount. CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data.

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:

$$dr_{1t} = a(b - r_{1t})dt + \sigma\sqrt{r_{1t}}dW_t. \tag{34}$$

I estimate the evolution of exogenous states in the model via maximum likelihood[8] using the monthly three-month Treasury bill data from January 1987 to January 2021.[9] The results for the risk-free rate are as follows:

Table A.5: 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

Note: This table contains estimates from fitting the Cox, Ingersoll, and Ross (1985) model on the three-month Treasury bill data from January 1987 to January 2021. Estimation proceeds via the maximum likelihood, and standard errors are obtained from the inverse Hessian.

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 quarterly 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

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

[9]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.

house price coefficient to reduce dimensionality.[10] 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:

$$\boldsymbol{s_t} = \boldsymbol{\mu} + r_{1t}\boldsymbol{\beta_{r_{1t}}} + \Phi_1 \boldsymbol{s'_{t-1}} + \Phi_2 \Delta H_{t-1} + \boldsymbol{e_t}, \tag{35}$$

where $\boldsymbol{e_t} \sim N(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}})$ and $\boldsymbol{\mu}, \boldsymbol{\beta_{r_{1t}}}, \Phi_2$ 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,[11] H_t is obtained from the Case-Shiller National House Price Index,[12] and L_t is obtained from the US Personal Income[13] divided by the US population.[14] Furthermore, H_t and L_t are converted to real terms using the Consumer Price Index for All Urban Consumers.[15] The results of the VAR estimation are as follows:

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

[11]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.

[12]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.

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

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

[15]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.

Table A.6: VAR estimates of state transitions

Parameter	$\boldsymbol{\mu}$	$\boldsymbol{\beta_{r_{1t}}}$	$\boldsymbol{\Phi_1}$			$\boldsymbol{\Phi_2}$	$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}$		
$\bar{c}_t$	.418 (.135)	.075 (.027)	.892 (.031)	0	0	0	.110		
$100 * \Delta H_t$	.447 (.228)	-.079 (.072)	0	.513 (.088)	0	-.022 (.088)	.080	2.664	
$100 * \Delta L_t$	.277 (.320)	.086 (.104)	0	0	-.472 (.079)	0	-.019	.026	5.572

Note: This table contains estimates from fitting a constrained VAR described in Equation (35). Data on mean mortgage rates $\bar{c}_t$ is obtained from the Primary Mortgage Market Survey (PMMS), data on house prices H_t are taken from the Case-Shiller National Home Price index, and data on personal income Y_t are taken as the ratio of US aggregate personal income divided by the US population. House prices and income are divided by the CPI for urban consumers and then transformed into growth rates.

Furthermore, I assume that borrowers pay \$2000 + 1% of the loan amount to obtain the PMMS survey rate following Agarwal, Driscoll, and Laibson (2013), and that the upfront closing cost payment is a deterministic nonlinear function of the deviation from the Freddie Mac PMMS rate $c-\bar{c}_t$, as denoted by the nonlinear function $\bar{o}(c-\bar{c}_t)$. As Appendix Figure A.4 shows, the rate and upfront trade-off from lender rate sheets does vary over time and follows MBS TBA prices closely between the July 2009 to December 2014 period when I have rate sheet data. However, calibrating the model to the average trade-off as implied by MBS TBA prices over the period allows me to approximate the average borrower welfare while significantly reducing the computational burden.

With this assumption, the interest rate and upfront closing cost menu for refinances is given in Equation (37):

$$\psi_{it}(c, M_{it}) = \$2000 + 0.01M_{it} + \bar{o}(c - \bar{c}_t(\bar{m}_t^l), t > 0 \tag{36}$$

where M_{it} is the remaining balance as given by the amortization formula starting from M_{i1}, and $\bar{o}$ is the average rate and upfront closing cost trade-off conditional on the market interest rate $\bar{c}_t(\bar{m}_t^l)$, which in turn incorporates time-varying lender costs $\bar{m}_t^l$. For new purchase originations, the rate and upfront closing cost menu faced by borrowers with a mortgage amount M_{i1} is the same as in Equation (37) plus an additional M_{i1} term that incorporates a lender markup:

$$\psi_{it}(c, M_{it}) = \$2000 + 0.01M_{i1} + \bar{o}(c - \bar{c}_t(\bar{m}_t^l) + \frac{M_{i1}}{M_{i1}}, t = 0 \tag{37}$$

where again $\bar{c}_t$ incorporates time-varying lender costs $\bar{m}_t^l$.

The estimates from Tables A.5 and A.6 are then used to simulate the transitions of the exogenous states in my model in Section 5.

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 5.1.2. 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 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.7, which is used for my model of $\hat{p}_{t'}$ as in Equation (21).

Table A.7: Logit model of prepayment

	(1)	
	Logit	
prepaid		
init_t	7.446***	(16.59)
init_t_sq	-4.169***	(-12.32)
sato	0.121	(0.62)
sato_sq	-0.765***	(-2.88)
refi_ratediff_gt0	0.348***	(4.55)
refi_ratediff_gtp25	0.345***	(4.08)
refi_ratediff_gtp5	0.599***	(8.31)
refi_ratediff_gtp75	0.322***	(4.69)
refi_ratediff_gt1	0.538***	(5.63)
refi_ratediff_gt1p25	0.144	(1.09)
burnout	-0.0549*	(-1.94)
burnout_sq	0.00182	(1.45)
Constant	-8.264***	(-54.48)
Observations	267603	

t statistics in parentheses

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

Note: The data used in this regression is the Optimal Blue-HMDA-CRISM data from January 2013 to May 2022, for 30-year, fixed-rate, conforming, primary residence mortgages originated in 2013–2019. The sample is further restricted to “TBA likely” mortgages defined as mortgages with a loan amount of at least $150k, loan-to-value ratio less than or equal to 85%, and FICO at origination greater than or equal to 680. The independent variable is an indicator variable for whether the borrower prepaid their mortgage in a given month. The dependent variables include the spread of the mortgage interest rate to the Freddie Mac survey rate at origination (SATO) and its square, as well as categories of rate incentive (the current spread of the mortgage interest rate to the Freddie Mac survey rate). CRISM data is attributed to Equifax Credit Risks Insight Servicing and Black Knight McDash Data.

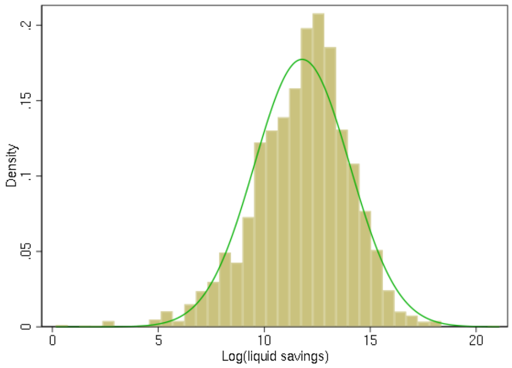
Using the prepayment model from Table A.7 and the interest rate model of Section A.5.1, with the risk-free rate r_{tf} being given as the implied 10 year rate under the Cox, Ingersoll, and Ross (1985) model, I estimate a $\hat{OAS} = 0.22\%$ by minimizing the equally-weighted difference between the observed 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 (21). 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 (2022).

A.5.3 Distribution of liquid assets

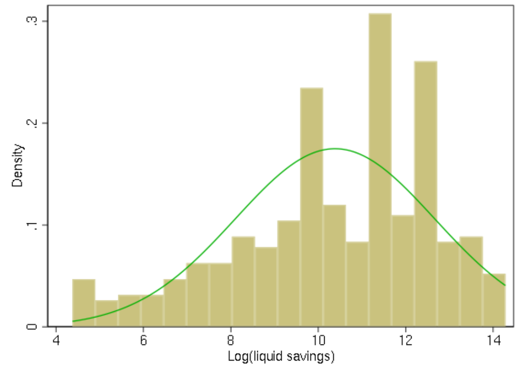
I estimate the distribution of liquid assets in the model using the Survey of Consumer Finances (SCF). To parsimoniously summarize the aggregate distribution of liquid assets, I fit log-normal distributions for liquid assets among Conventional mortgage holders (defined as households with X724=5) using the 2013–2019 SCF data. The data item for liquid assets is "Total Financial Assets," including bank accounts, CDs, mutual funds, stocks, bonds, liquid retirement savings, savings bonds, cash value of whole life insurance, and other managed/financial assets. Black households are defined as households with X6809=2, Hispanic households are defined as households with X6809=3.

Histograms of the those households' log liquid assets, along with the fitted log-normal distribution, are plotted in Figure A.14.

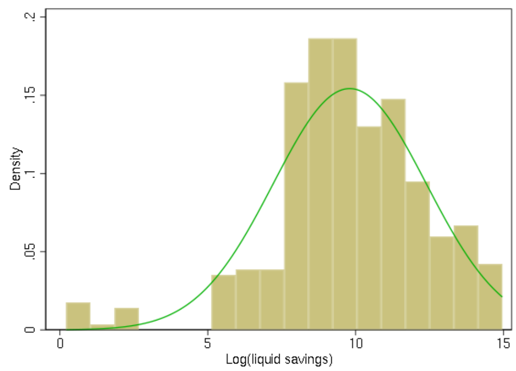
Figure A.14: Histograms of mortgage borrower's log liquid assets along with their fitted distributions



(a) Non-Black, non-Hispanic borrowers



(b) Black borrowers



(c) Hispanic borrowers

Note: The data used in this figure is from the 2013–2019 SCF. The log of the households' total liquid assets, including liquid savings, CDs, mutual funds, stocks, bonds, liquid retirement savings, savings bonds, cash value of whole life insurance, and other managed/financial assets, are plotted for households with a mortgage. In addition, a normal distribution is fit to the log liquid assets. For non-Black, non-Hispanic households with a mortgage, I estimate a mean of 11.78 and a standard deviation of 2.25. For Black households with a mortgage, I estimate a mean of 10.38 and a standard deviation of 2.28. For Hispanic households with a mortgage, I estimate a mean of 9.80 and a standard deviation of 2.59.

A.6 ADL

I investigate the cross-subsidization of low upfront closing cost mortgages from the perspective of lender rate-setting under the model of Agarwal, Driscoll, and Laibson (2013), hereafter referred to as ADL. ADL proposes a model of optimal refinancing based on a Brownian motion interest rate model as well as assumptions on the inflation rate, tax rate, and the probability of moving. The ADL parameters are as follows:

$\rho =$	0.03
$\sigma =$	0.004597
$\pi =$	0.0168
$\mu =$	0.074
$\tau =$	0.22

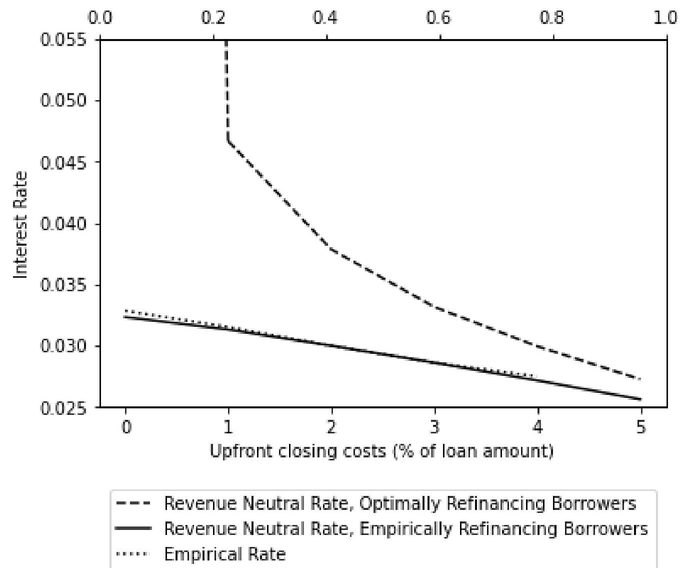
where ρ is the discount rate, σ is the standard deviation of the mortgage rate, π is the expected inflation rate, μ is the expected probability of moving, and τ is the tax rate.

I compute the counterfactual interest rates that the lenders would have to charge to remain revenue neutral if all borrowers behaved as in the ADL model. As an alternative model of optimal refinancing that is relatively easy to compute, this serves as a robustness check to my main calibration result in Figure 6. The result from the ADL model is shown in Figure A.15.

In particular, Figure A.15 shows suggests that optimally refinancing borrowers (in the sense of ADL) receive a substantially lower rate than what they would have received without cross-subsidization: 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, consistent with my main results.

[16] With zero upfront closing costs and optimally refinancing borrowers, I find that lenders would have to charge a rate of 91% to remain revenue-neutral.

Figure A.15: Pricing of mortgages mortgages by upfront closing cost choice, ADL Optimally Refinancing Borrower



Note: Figure A.15 presents the equilibrium rate and upfront closing costs trade-off from the model and compares it to the empirical rate and upfront closing costs trade-off that I estimate from the data. The "Market rate, model implied" solid line refers to the equilibrium rate and closing cost trade-off given the logit prepayment hazard function and an estimated OAS. The "Market rate, empirical" dotted line is the implied rate and upfront closing cost options from MBS TBA prices combined with the PMMS survey rate. I show that the rate and upfront closing cost options implied by MBS TBA prices is consistent with the options actually presented to borrowers in rate sheets from a regression with rate sheet fixed effects in Appendix A.3. Finally, the "No cross-subsidization counterfactual" dashed line refers to the model-implied equilibrium rate and closing cost trade-off in a world where the lender is pricing their mortgages for the calibrated ADL borrower with perfect information on their type.