Refinancing Cross-Subsidies in the Mortgage Market*

Jack Fisher   Alessandro Gavazza   Lu Liu

Tarun Ramadorai   Jagdish Tripathy§

July 2023

Abstract

In household finance markets, inactive households can implicitly cross-subsidize active households who promptly respond to financial incentives. We assess the magnitude and distribution of cross-subsidies in the mortgage market. To do so, we build a model of household mortgage refinancing and structurally estimate it on rich administrative data on the stock of outstanding mortgages in the UK. We estimate sizeable cross-subsidies from relatively poorer households and those located in less-wealthy areas towards richer households and those located in wealthier areas. Our work highlights how the design of household finance markets can contribute to wealth inequality.

*We thank seminar participants at the Bank of England Workshop on Household Finance and Housing, SITE Workshop on Financial Regulation, Imperial College, Kellogg School of Management, King's College Business School, NYU Stern School of Business, the SFS Cavalcade, London Business School, University of Vienna, Washington University in St. Louis, NBER Summer Institute (Household Finance, Real Estate), SITE Workshop on IO of Healthcare and Consumer Finance Markets, Wharton, Federal Reserve Bank of Philadelphia, CEAR-RSI Household Finance Workshop, as well as Nuno Clara, Mikhail Golosov, Pascal Noel, Steffen Andersen and Gregor Matvos for useful comments and discussions. Alessandro Gavazza gratefully acknowledges financial support from the European Research Council (under Consolidator Grant Agreement No. 771004). Any views expressed in this article are not meant to represent the views of the Bank of England, members of its Financial or Monetary Policy Committee, or the Financial Conduct Authority. This work was supported in part through the NYU IT High Performance Computing resources, services, and staff expertise. All errors are our own.

§Fisher: London School of Economics. Email: j.w.fisher@lse.ac.uk. Gavazza: London School of Economics. Email: a.gavazza@lse.ac.uk. Liu: The Wharton School, University of Pennsylvania. Email: lliu1@wharton.upenn.edu. Ramadorai: Imperial College Business School and CEPR. Email: t.ramadorai@imperial.ac.uk. Tripathy: Bank of England. Email: jagdish.tripathy@bankofengland.co.uk.

1 Introduction

In retail financial markets, households often face complex contracts which require prompt action in response to changing financial incentives. Households who do not swiftly respond to these incentives can unwittingly provide revenues to financial firms. Conversely, such products can be beneficial to more sophisticated customers who are quicker to take appropriate action. This can result in regressive cross-subsidies in financial markets that flow from less sophisticated customers, who are often poorer and less educated, to those who are more sophisticated, wealthy, or educated. In this way, the design of household finance products can be a powerful contributor to wealth inequality.

Our paper provides a new structural approach to quantify such household finance cross-subsidies and to identify how they are distributed across the population. We apply the method in the setting of residential mortgage refinancing. Mortgages are the largest household liability (Campbell, 2006; Badarinza, Campbell, and Ramadorai, 2016; Goetzmann, Spaenjers, and Van Nieuwerburgh, 2021), but despite their importance in their budgets, many households do not appropriately manage this debt. A crucial determinant of sound mortgage management is timely refinancing in response to financial incentives, and evidence has built up that lower-income and less-educated households fall short on this dimension (Agarwal, Rosen, and Yao, 2016; Keys, Pope, and Pope, 2016; Andersen, Campbell, Nielsen, and Ramadorai, 2020; Byrne, Devine, King, McCarthy, and Palmer, 2023).

The first step in our approach is to build and estimate a structural model to match a broad set of moments from high-quality administrative data on the entire stock of mortgages in the United Kingdom, an environment uniquely suited to answering these questions. In the UK, as in many other countries, the dominant mortgage form is a "dual-rate" mortgage contract.[1] At origination, borrowers fix their mortgage rate for a short initial period at a low "discounted rate." To fully take advantage of such contracts, it is imperative to swiftly refinance at the point at which the initial fixation period ends, to avoid being rolled on to a significantly more expensive "reset rate."[2] Households who fail to do so pay these higher reset rates, and these higher payments benefit those who quickly refinance back into lower discounted rates.

[1]Such dual-rate contracts are ubiquitous in retail markets, including in credit cards, and cellphone and electricity plans (Armstrong and Vickers, 2012).

[2]This feature of the UK mortgage market has prompted prominent calls for reform which highlight the implicit cross-subsidy (Miles, 2004).

The second step in our approach is to use the estimated structural parameters of the model to compare household outcomes in the status quo with outcomes under a counterfactual scenario. In this counterfactual, borrowers face a simpler contract that pays a single rate until mortgage maturity, with no refinancing required. We conduct these counterfactual comparisons for households across the income distribution and regions of the UK. This procedure allows us to make the invisible cross-subsidies in the current system visible, and to assess their distribution across the population. The approach can be more widely applied to uncover and analyze cross-subsidies in other market settings.

Our model assumes that households are heterogeneous along two dimensions. The first dimension is households' valuation for owned housing (we model renting as an outside option), which allows us to match the loan size distribution in the stock of outstanding mortgages. The second dimension is the costs that households face at the point of refinancing. We model these costs as comprised of both a persistent component as well as temporary fluctuations or shocks around this persistent component. While households in the model rationally optimize subject to these costs, this setup allows us to capture household inertia and inattention to beneficial refinancing opportunities through a realized high refinancing cost shock in any given period. Moreover, at the point of deciding whether to buy a house (and, if they do so decide, their mortgage size), households have only noisy information about their future costs of promptly refinancing. This captures potentially imperfect "self-knowledge," and we consider variants of the model in which, at origination, households have different degrees of precision about their future refinancing costs.

Households face a dual-rate structure in the model, and all households are initially on the discounted rate. When the initial fixation period ends, households choose to refinance when the benefits of refinancing, driven mainly by the difference between the discounted and the reset rates and loan size, outweigh the costs of refinancing whose structure is described above. Larger loans are therefore more likely to pay discounted rates: in the cross-section, these loans correspond to households with a greater valuation for housing; in the time-series, these loans correspond to households who recently originated their mortgages.

The model is built to facilitate easy aggregation of loans, generating intuitive expressions for aggregate mortgage loan balances on the discounted rate and on the reset rate. This allows us to match a rich set of moments in granular and comprehensive data from the UK's Financial Conduct Authority (FCA) which tracks the stock of all outstanding UK mortgage loans issued by all regulated financial institutions in the country at a semi-annual frequency between 2015H1 and 2017H2.[3] The granular nature of the data means that we observe household-level mortgage refinancing behavior; and the comprehensive coverage of the entire mortgage stock allows us to compute cross-subsidies.

The UK mortgage market is a particularly good venue to study these questions for several reasons. First, in contrast with the United States, the refinancing decision is straightforward. Households choose between re-entering a discounted rate contract versus paying the substantially higher reset rate, which, as we show, is a dominated choice even when accounting for any option value of waiting to lock in an advantageous rate.[4] Second, subject to collateral re-verification, many UK mortgages are portable, which means that refinancing in the UK is less affected by households' unobservable moving propensities, unlike in the US and other markets. Third, unlike many other data sources, our UK data provides details on borrower incomes, meaning that we can assess the distribution of cross-subsidies along this important dimension.

Our analysis mainly focuses on the stock in 2015H1, when the total stock of household mortgage debt in our sample equals £470 billion. The majority of this stock (69.8%) pays the discounted rate, but the remaining large share (30.2%) pays the reset rate, with an average rate spread of 52 basis points (bps). Hence, there is an appreciable spread between reset rates and discounted rates, and many households pay these different rates.

We estimate the model parameters assuming that the market is in steady state and match the data very well. Our estimates imply that average refinancing costs equal £3,866 among mortgage borrowers, with a standard deviation equal to £6,641; we later discuss how these money-metric estimates relate to the broader literature on mortgage refinancing.

These estimated parameters are the main input into our cross-subsidy calculations, which compare counterfactual economies featuring a range of different single rates with the dual-rate baseline economy. In these counterfactual scenarios, households adjust both their individual loan sizes (intensive margin) as well as their participation in the housing and mortgage markets (extensive margin), in response to the different paths of mortgage rates and to the elimination of refinancing costs in the single-rate counterfactual economy

[3]In what follows, we denote the first and second observations in each year of our sample by H1 and H2 respectively to denote "half-years".

[4]This feature eliminates reliance on an auxiliary model of the real option, such as that of Agarwal, Driscoll, and Laibson (2013) and allows us to recover households' implicit refinancing costs directly.

relative to the baseline dual-rate economy. These adjustments depend on the degree of self-knowledge that households have about their own persistent refinancing costs (i.e., their general tendency towards inaction). The less self-knowledge that households have, the smaller their mortgage size and mortgage market participation adjustments tend to be, but we find that these adjustments are significant even when households have noisy information about their future refinancing costs.

In the counterfactual single-rate economy with an interest rate equal to the weighted-average rate in the sample, the aggregate mortgage debt increases by 4.13 percent relative to that in the baseline economy.[5] This increase is mainly driven by high-refinancing-cost households, who no longer pay either the punitive reset rate or refinancing costs. Their increased propensity to enter the market translates into an increase in the total number of mortgages in the counterfactual economy. However, the mean initial loan balance falls in the counterfactual equilibrium by 2.37 percent of the baseline average loan size, because the composition of borrowers changes: marginal households who enter the mortgage market in the single-rate economy have smaller loan sizes than inframarginal households whose participation does not change.

We estimate two extended versions of the model to assess the distribution of cross-subsidies across the population of mortgage borrowers. The first estimates parameters separately for different geographical regions in the UK to capture regional heterogeneity in housing preferences and refinancing costs. The second estimates parameters separately for 12 income groups (bottom-eight income deciles, and the top-two deciles each additionally split into two sub-groups). These extended models continue to match the aggregate moments very well, and also feature considerable differences in refinancing propensities across regions and income groups. These differences presage significant variation across both the income distribution and regions in cross-subsidies that are paid and received.

More specifically, we find clear evidence that higher-income households and households in the richer South-West of the UK would pay higher rates under the single-rate structure, and households in the relatively poorer North-East and North-West of the UK would pay lower mortgage rates under the counterfactual single-rate scenario than they do in the dual-rate economy. These redistributive effects depend on heterogeneity in both refinancing costs and

[5]As we later discuss, the implied interest rate elasticity to mortgage debt in our setup is comparable to other estimates in the literature.

valuations for housing. The status quo dual-rate economy penalizes households with high refinancing costs and low valuations for housing. Such households (who are typically lower income) have smaller loan balances, which lowers the refinancing benefit and leads to less time spent on the discounted rate.

Finally, the counterfactual single-rate economy displays striking differences across groups in their endogenous adjustments to mortgage takeup and mortgage sizes. Average mortgage debt shrinks for higher-income groups and wealthier regions in response to the counterfactual single rate, since they no longer have access to the discounted rate. In contrast, the counterfactual single-interest rate economy induces lower-income households to enter the mortgage market because they expect to pay lower rates and incur no refinancing costs. This is evident in increases in the home-ownership rate, mainly driven by low-income households. This "democratization" of mortgage takeup under the counterfactual is an important indicator of the regressive effect of cross-subsidies in the dual-rate economy.

The remainder of this section discusses related literature. Section 2 describes the administrative data and the UK institutional setting. Section 3 lays out the structural model and describes the computation of cross-subsidies. Section 4 discusses parameter estimation, and model fit. Section 5 discusses counterfactual analysis and the flow of cross-subsidies across income groups and regions. Section 6 concludes.

1.1 Related Literature

Our paper contributes to several strands of the literature. First, our work complements many empirical papers that document switching costs, inertia, and inattention in insurance and household finance markets, such as health insurance (e.g., Handel, 2013), car insurance (e.g., Honka, 2014), retirement plans (e.g., Luco, 2019; Illanes, 2016), credit cards (e.g., Ausubel, 1991; Stango and Zinman, 2016; Nelson, 2022), pension contributions (e.g., Choi, Laibson, Madrian, and Metrick, 2002), and portfolio rebalancing (e.g., Brunnermeier and Nagel, 2008) among others.[6] However, none of these papers focuses on documenting regressive cross-subsidies, though this possibility has been raised in theory (e.g., Gabaix and Laibson, 2006; Armstrong and Vickers, 2012).

[6] Farrell and Klemperer (2007) present a survey of the literature on switching costs, with a theoretical focus; Heidhues and Kőszegi (2018) survey the literature on behavioral industrial organization, and Gavazza and Lizzeri (2021) the literature on markets with frictions.

The papers that document inaction and frictions in mortgage refinancing (e.g., Agarwal, Rosen, and Yao, 2016; Keys, Pope, and Pope, 2016; Scharfstein and Sunderam, 2016; DeFusco and Mondragon, 2020; Byrne, Devine, King, McCarthy, and Palmer, 2023; Berger, Milbradt, Tourre, and Vavra, 2023) are more directly related to our work. We advance this literature, developing a novel framework for refinancing inaction that allows us to quantify the magnitudes of cross-subsidies across households through counterfactual analysis. This is a different approach to Andersen, Campbell, Nielsen, and Ramadorai (2020), who model a fixed refinancing cost ("state-dependent inaction"), but with intervals of "time-dependent inaction" where refinancing is not possible, using a periodic "Calvo" shock to borrowers, and Berger, Milbradt, Tourre, and Vavra (2021), who adopt a similar approach in their analysis of US refinancing behaviour. These approaches imply that the costs of refinancing are always higher than the benefits during periods of time-dependent inaction, but do not quantify these costs. In contrast, our model features a household-specific fixed refinancing cost with a time-varying shock, meaning that our structural estimation recovers the full distribution of the costs of inaction across households and over time. Apart from the differences in setting, this different modelling approach explains why the average refinancing costs that we estimate are modestly higher than those in Andersen, Campbell, Nielsen, and Ramadorai (2020) and Berger, Milbradt, Tourre, and Vavra (2021).[7]

Complementary work by Zhang (2022) and Berger, Milbradt, Tourre, and Vavra (2023) studies the distributional effects of refinancing frictions in the US. Using a lifecycle model, Zhang (2022) highlights that US borrowers that do not pay upfront "points" (i.e., closing costs) to reduce mortgage rates are worst affected under the status quo. Intuitively, such borrowers pay higher interest rates for longer, thereby contributing more to lender revenues. Berger, Milbradt, Tourre, and Vavra (2023) also study the US market, and endogenize mortgage rates under simplifying assumptions on the behavior of mortgage borrowers and investors. They consider counterfactuals that include alternative contracts as well as policies that reduce borrowers' frictions. In contrast with these papers, our paper features a more detailed model of household inaction alongside a more stripped-down supply side. Moreover, we document the redistributive consequences of cross-subsidies across income groups and

[7] Andersen, Campbell, Nielsen, and Ramadorai (2020) estimate an average total psychological plus fixed refinancing cost of £1,852 in the Danish mortgage market. Berger, Milbradt, Tourre, and Vavra (2021) estimate an average refinancing cost of $1,934 in the US mortgage market. These are slightly lower than our estimate of the average cost across both refinancing and non-refinancing borrowers, which equals £3,866.

regions. Finally, our focus on the UK market reduces confounds arising from unobservable moving propensities, and enables a simple calculation of refinancing inaction based on the dual-rate structure. The UK setting bears similarities with many mortgage markets outside of the US, making our work also broadly applicable to such markets.

Second, our paper is connected to a growing body of work on the design of mortgage markets around the world (Campbell, 2013; Piskorski and Seru, 2018). For example, several mortgage markets also feature fixed rates for a shorter interval than the maturity of the mortgage. Allen and Li (2020) study borrower refinancing and lender pricing in the Canadian mortgage market; similarly, Thiel (2021) studies a ban on price discrimination between new and existing customers in the Dutch mortgage market. We focus on implicit cross-subsidies across borrowers in the cross-section, whereas Allen and Li (2020) and Thiel (2021) focus on intertemporal price discrimination within borrowers.[8]

Finally, our structural model provides a money-metric assessment of cross-subsidies in an important household finance market, and shows that these cross-subsidies are regressive. This showcases how the design of the financial system can contribute to inequality, connecting our work to the growing literature on wealth inequality (Alvaredo, Chancel, Piketty, Saez, and Zucman, 2017; Benhabib and Bisin, 2018; Fagereng, Guiso, Malacrino, and Pistaferri, 2020; Hubmer, Krusell, and Smith Jr, 2020) and, more specifically, to that on inequality in financial wealth (Campbell, Ramadorai, and Ranish, 2019; Greenwald, Leombroni, Lustig, and Van Nieuwerburgh, 2021). In the process, we document that cross-subsidies vary across regions and devolved administrations of the UK, which shows that regional redistribution can occur directly as a result of differential efficiency in the use of financial products. These results speak to the literature on regional redistribution in housing and mortgage markets (Hurst, Keys, Seru, and Vavra, 2016; Beraja, Fuster, Hurst, and Vavra, 2019).

2 Data and Institutional Setting

Our primary data source is the FCA, which comprehensively tracks the stock of outstanding mortgage loans issued by all regulated financial institutions in the UK. The specific FCA

[8]Our paper also contributes to the growing literature on UK mortgage markets, including Benetton (2021); Robles-Garcia (2022); Cloyne, Huber, Ilzetzki, and Kleven (2019); Best, Cloyne, Ilzetzki, and Kleven (2020); Belgibayeva, Bono, Bracke, Cocco, and Majer (2020); Benetton, Gavazza, and Surico (2021); Liu (2022). Most of these studies focus on the flow of newly originated mortgages, whereas we focus on the stock of mortgages.

dataset that we use is the Product Sales Database 007 (henceforth PSD007), which reports information about the stock of mortgage loans between June 2015 (henceforth 2015H1), and December 2017 (2017H2) at a semi-annual frequency.[9]

At each reporting date, PSD007 records the original loan amount, outstanding balance, original loan term, remaining term to maturity, current interest rate, current monthly payment, and performance status (i.e., whether the loan is in arrears and if so, for how long this has been the case) for each outstanding mortgage. The database also includes information on the property location at the most granular level in the UK (6-digit postcode), and borrower characteristics such as date of birth and the opening date for the bank account associated with the mortgage. Table A.1 in the appendix provides more detailed descriptions of the main variables from the PSD007 dataset used in this paper.

The PSD007 dataset does not include information on borrower incomes. We therefore merge borrowers in the stock data with comprehensive loan-level data on borrower characteristics, including their income, shared with lenders at the time of loan origination. We also measure the current loan-to-value (LTV) ratio on each outstanding loan following a common approach in the literature, dividing the outstanding loan balance by the scaled house price at mortgage origination, using Local Authority district-level house price indices. Appendix A.2 provides details of the procedure used to merge borrower and house characteristics at loan origination to our stock data.

We further complement the PSD007 dataset with data on UK homeownership rates sourced from the Office for National Statistics (ONS) dataset *Dwelling stock by tenure*. These homeownership data allow us to measure households' extensive margin decision of whether to buy a house and take a mortgage, or rent.

Using rich data on the stock of mortgages offers several advantages over using the flow of originations. Notably, the stock allows us to accurately capture refinancing behavior across all mortgage maturities, including mortgages that were originated in the past. Moreover, the parameters of a structural model estimated using the stock of mortgages rather than using the flow depend less on changes in refinancing behavior or refinancing waves over short periods of time. Finally, using the mortgage stock facilitates computing average mortgage

[9]Regulated financial institutions in the UK are legally required to report these details within 30 working days following the end of each calendar half-year. The group of regulated financial institutions in the UK includes deposit-taking institutions (including building societies), as well as some non-bank financial institutions. Our sample focuses on the owner-occupier segment of the mortgage borrowing population, and excludes "buy-to-let" mortgages which are issued mainly to landlords on rental properties.

rates and aggregate lender revenues, which proves useful in our counterfactual analyses.

2.1 UK Mortgage Market: Institutional Features

Our work exploits a few key features of UK mortgage markets apparent in our PSD dataset.

First, the UK mortgage market features posted prices at the national level, with no variation across regions, as Cloyne, Huber, Ilzetzki, and Kleven (2019), Benetton (2021), Robles-Garcia (2022), and Benetton, Gavazza, and Surico (2021) (among others) document. Borrower-specific pricing, common in the US mortgage market, is virtually non-existent in the UK market.

Second (and crucial for our purposes), the vast majority of UK mortgages are issued with discounted interest rates which are fixed for a set time period, usually between one and five years (the modal fixation period is two years), depending on the contract chosen by the borrower. During the discounted period, households typically incur substantial prepayment penalties (between 3-5% of the loan balance), which means that households typically refinance after the end of the fixed period (Cloyne, Huber, Ilzetzki, and Kleven, 2019; Belgibayeva, Bono, Bracke, Cocco, and Majer, 2020). At the end of the discounted period, the mortgage rate automatically rolls over into a higher reset rate known as the "standard variable rate," unless borrowers choose to refinance the mortgage into another discounted rate (for a detailed treatment of the characteristics of the UK mortgage market see Miles, 2004).[10]

This "dual-rate" structure is a feature of many mortgage systems, including Canada, Australia, India, Ireland, Germany, and Spain, meaning that our study is more broadly applicable around the world.[11] We do not study the origins of this rate structure, which likely reflects mortgage lenders' funding structures and price-discrimination strategies between active and inactive borrowers (Ellison, 2005; Gabaix and Laibson, 2006; DellaVigna and Malmendier, 2006), and we focus instead on its implications for borrowers' refinancing. That said, in Appendix D, we follow and extend the analysis of Cloyne, Huber, Ilzetzki,

[10] There is a third type of interest rate known as a tracker rate, paid on around 15% of all mortgages outstanding, which is a floating rate linked to the Bank of England base rate. We exclude such mortgages from our analysis since such mortgages are subject to rate fluctuations and there are rarely transitions from the rest and discounted rate category into this category. Further details in appendix A.1.

[11] Badarinza, Campbell, and Ramadorai (2018) provide information on mortgage interest-rate fixation periods across a broad set of countries and show that many large economies have similar average mortgage-rate fixation periods to the UK.

and Kleven (2019), who perform a thorough comparison between borrowers who pay the discounted and the reset rates, suggesting that the dual-rate structure does not seem designed for lenders to screen borrowers based on their default risk.[12]

This dual-rate contract structure provides strong incentives for households to refinance at the expiration of the fixation period. UK households are free to take advantage of these incentives to refinance, as there are no further credit checks when households refinance with their existing lender, and any upfront fees can be rolled into the loan balance, meaning that liquidity constraints do not inhibit refinancing (Best, Cloyne, Ilzetzki, and Kleven, 2020). In Appendix E, we also rule out the possibility that borrowers rationally stay on the reset rate to exploit the real option of timing their refinancing to coincide with interest rate declines.

Third, an additional feature of the UK setting is that mortgages are portable, meaning that households can retain their existing mortgage contract when they move, subject to the new collateral being re-verified.[13] This feature stands in contrast with the US, where the lack of portability means that moving probabilities are a potentially more important driver of both prepayment/refinancing and contract choice (Stanton, 1995; Stanton and Wallace, 1998; Zhang, 2022).

2.2 Borrowers Ineligible to Refinance

One potential challenge to our empirical analysis and our cross-subsidy calculations is to distinguish between households who can refinance, but do not do so promptly, from households who are constrained and unable to take advantage of refinancing opportunities. To address this potential confounding effect, we filter our data to remove borrowers who are potentially ineligible for refinancing—i.e., borrowers who are "involuntarily" on the reset rate, but who would potentially like to switch if they were allowed to do so.

To identify these ineligible borrowers, we follow studies by the FCA (Financial Conduct Authority, 2019b, 2021) and a 2018 industry agreement that unified and codified refinancing eligibility criteria across major UK lenders (65 lenders, with a market share of around 95%). Passing these eligibility criteria means that a mortgage borrower can refinance into a new contract with their lender, without any affordability assessment, meaning no

[12] This pricing structure with a discount for new or active customers is quite common in many other retail markets, including electricity, telecoms, and magazines, in which default concerns play a negligible role.

[13] Among other countries, Australia, Canada, and Germany share this feature (Lea, 2010).

additional credit or income checks.[14] The criteria are that the borrowers are first-charge owner-occupiers that are existing borrowers of an active lender, up to date with their payments, with a minimum remaining term of 2 years, and a minimum outstanding balance of £10,000 (Financial Conduct Authority, 2019b). We broaden out these eligibility criteria to filter out borrowers potentially ineligible for refinancing, under the assumption that the 2018 agreement ratified pre-existing practice that was prevalent in the 2015H1 stock.

Appendix Table A.6 shows the exact proportion of loans that are potentially ineligible for refinancing using these criteria as well as broader definitions of ineligibility. Borrowers who have very high LTVs greater than 95% comprise approximately 2% of the sample. These borrowers may find it difficult to refinance, even though they are strictly eligible under the industry agreement if they fulfill all the other criteria. Borrowers with small remaining loan balances (loans smaller than £30,000) constitute approximately 6% of the total sample. And about 5% of loans are non-performing (in arrears, or under forbearance or possession orders). Applying these filters together removes around 14% of the mortgage stock in 2015H1-2017H2.

We note here that we estimated our model on both unfiltered and filtered samples. Filtering does not materially affect our main qualitative results on the regressive nature of cross-subsidies, for two main reasons. First, in the filtered sample the share of mortgage debt paying the reset rate is still quite large, and lower-income borrowers are more likely to pay the reset rate than higher-income borrowers. Second, the largest fraction of excluded borrowers are those with small loan balances, for whom refinancing benefits are small, because the refinancing benefit is proportional to the loan balance. Appendix A.4 provides more information on these filtered borrowers.[15]

[14] The UK is somewhere between the US and Denmark in this respect. In the US, a credit check is triggered at the point of refinancing (Keys, Pope, and Pope, 2016), whereas in Denmark, even delinquent borrowers are able to refinance as long as there is no cash out (Andersen, Campbell, Nielsen, and Ramadorai, 2020). The UK system does not trigger a credit check at the point of refinancing as long as the borrower satisfies the eligibility criteria.

[15] A 2018 FCA report of the mortgage market (Financial Conduct Authority, 2019b) studied 2 million reset rate mortgages using the same data that we employ, and concluded that only approximately 30,000 of these mortgages were unable to switch despite being up to date with payments. The report finds that two-thirds of these mortgages were associated with an inactive, failed lender (e.g., Northern Rock, famously subject to a run during the financial crisis); and the remainder were either interest-only mortgages that were subject to changes in lending standards following the financial crisis, or in negative home equity. We expect that our filters catch many of these mortgages.

Table 1: Summary Statistics for the Mortgage Stock in 2015H1

	MEAN	SD	P25	P50	P75
CURRENT LOAN BALANCE (£)	130,871	97,858	71,450	106,967	158,495
CURRENT INTEREST RATE (IN PP)	3.46	0.98	2.53	3.38	4.00
SPREAD TO T-BILL (IN PP)	2.79	1.05	2.05	2.53	3.54
ORIGINAL LOAN BALANCE (£)	142,333	100,661	82,000	118,399	170,000
ORIGINAL TERM (IN YEARS)	23.32	7.07	19.00	25.00	28.00
REMAINING TERM (IN YEARS)	19.27	7.67	13.92	19.00	24.33
REMAINING DISCOUNTED PERIOD	2.11	1.52	1.00	1.83	3.08
BORROWER AGE	41.97	10.02	34.00	41.00	49.00

Notes: The table above shows summary statistics of mortgages from the stock data reported in 2015H1. The sample includes mortgages in two categories, namely, those paying discounted interest rates, and those paying the Standard Variable Rate. The total sample comprises around 3.59 million mortgages, of which 65.0% are discounted rate mortgages at this point in time. Appendix Table A.1 contains a description of the underlying variables.

2.3 Summary Statistics of the Mortgage Stock

Our analysis focuses on the 2015H1 mortgage stock, which comprises 3.59 million mortgages of borrowers eligible to refinance, and for whom we have estimates of current income.[16]

Table 1 shows summary statistics for selected variables in this filtered 2015H1 sample. On average, the mean outstanding balance equal £130,871 and a mean loan balance at origination of £142,333 (the difference is attributable mainly to amortization). This aggregates to a total stock of outstanding mortgage debt of £470 billion.

Taking an equal-weighted average across all mortgages, Table 1 shows that they pay an average interest rate of 3.46% at the end of 2015H1, at a spread of 2.79% over maturity-matched UK Treasuries, and have a remaining term to maturity of 19.3 years on average.[17] 65.0% of the 3.59 million mortgages pay discounted rates in this 2015H1 sample, with an average equal-weighted remaining discounted period of 2.1 years.

[16]The main statistics of the mortgage stock are quite stable between 2015H1 and 2017H2, consistent with the idea that short-run changes have small effects on the stock of long-term debt contracts. Appendix B describes the evolution of the mortgage stock between 2015H1 and 2017H2, which exhibits two main patterns: (1) the fraction of mortgage debt paying the reset rate decreases by 2017H2, and (2) the spread between the average reset rate and the discounted rate increases over the same period. While the first pattern should decrease the magnitude of cross-subsidies across borrowers, the second one should increase them, with a small net effect.

[17]Mortgage spreads are computed with respect to the yield on a nominal zero coupon UK Treasury with maturity matched to the mortgage interest rate fixation period. We use the short-term interest rate for mortgages paying the reset rate. For instance, for a mortgage with t years of fixation, the spread is calculated by subtracting off the spot rate for a UK treasury maturing in t years as at the reporting date.

Table 1 also reveals considerable cross-sectional variation in these variables. Such variation is particularly evident in the outstanding loan balance and the remaining mortgage term. When the outstanding loan balance and/or the remaining term are low, borrowers should be less likely to refinance given the lower financial incentive from any interest rate reduction associated with doing so. The mortgages also vary in terms of the overall interest rate they pay, as well as their spread over the maturity-matched UK treasury rate. There is also demographic variation evident in the age of borrowers, with both relatively young borrowers, aged 34 at the 25th percentile of the cross-sectional distribution, and older borrowers, aged 49 at the 75th percentile of the cross-sectional distribution represented in the mortgage stock.

Two additional statistics not reported in Table 1 constitute important targets for our model. First, the ONS dwelling data report that 63% of households are homeowners in 2015. Second, rate types are highly persistent. Figure B.2 in Appendix B displays the transition probabilities between discounted and reset rates in the PSD007 dataset. Households on a reset or discounted rate are much more likely to stay on the same rate type over the next 24 months than to switch—i.e., 76.2% of 2015H1 borrowers pay the same rate type in 2017H1. However, some borrowers do switch over time. Switches from the lower discounted to the higher reset rate may reflect inattention and inertia, or more generally, refinancing costs, while refinancing benefits decline as the loans amortize. Against this backdrop, switches from the reset to the discounted rate suggest that these costs vary over time within households. Taken together, these switching patterns suggest that a combination of household-specific fixed refinancing costs and time-varying stochastic shocks may capture the high persistence of rate types, as well as the occasional switches across rate types over time.

Table 2 shows summary statistics across quantiles of the income distribution of borrowers; this is an important dimension along which we later evaluate cross-subsidies. The third column of the table shows that the homeownership rate rises monotonically with the level of income—it equals 50% in the lowest-income group and attains 96% in the highest-income group. The remaining columns refer to borrowers. Their loan balance increases with their income, as expected. More importantly, the share of mortgages on the discounted rate (fifth column) also tends to increase with borrower income. These patterns document that lower-income borrowers are less likely to refinance than higher-income borrowers, hinting at the likely direction of cross-subsidies. Table C.1 in Appendix C provides a similar table

Table 2: Summary Statistics for the Mortgage Stock in 2015H1, by Income Quantiles

Quantiles	Income (£)	Homeowners (%)	Balance (£)	Discounted (%)
0-10	24,604	0.50	60,144	0.66
10-20	29,483	0.61	73,839	0.64
20-30	34,564	0.64	84,721	0.64
30-40	39,581	0.68	94,547	0.64
40-50	44,986	0.72	104,950	0.64
50-60	51,327	0.75	116,473	0.64
60-70	59,412	0.80	130,123	0.64
70-80	71,261	0.82	149,041	0.66
80-85	80,290	0.84	169,791	0.66
85-90	94,142	0.86	190,849	0.67
90-95	122,708	0.91	227,788	0.68
95-100	214,486	0.96	345,904	0.69

Notes: The table above shows summary statistics of mortgages from the stock data reported in 2015H1, split by income quantiles of borrowers. Appendix Table A.1 contains a description of the underlying variables.

across UK regions, confirming that borrowers in higher-income regions are more likely to pay discounted rates than those in lower-income regions.

Overall, the summary statistics reported in Tables 1 and 2 document that the UK mortgage market comprises a mix of borrowers paying discounted rates and reset rates. While mortgages on discounted rates constitute the main share, a large fraction (30.2% by loan balance) of the outstanding mortgage stock pays the reset rate, at a spread of 52 bps over the equivalent discounted rate. Our dataset includes mortgages by two large lenders who offered to cap reset rates at 250 bp for mortgages issued up to and during the 2007-09 financial crisis. Excluding these lenders (around 900k observations) pushes up the average rate for reset rate mortgages substantially (with no change in the average rate for discounted mortgages), increasing the spread to 110 bp. We have kept mortgages by these two large lenders in our sample to provide conservative cross-subsidy estimates.

In the next Section, we develop a model that we map to these data features in our structural estimation; we use the model to quantitatively assess the magnitude of the cross-subsidy that the dual-rate structure embeds.

3 Model

We model a mortgage market in which a measure M of households enters in each period. When they enter the market, households choose whether to buy a house with a mortgage or rent a property. If a household i chooses to buy, they pay a one-time origination cost k_i^o and obtain per-period flow utility from their house equal to $v_i h_i^\alpha - m(l_i, r, T)$, where v_i is household i's per-period valuation for housing, h_i is the size of the house that the household i chooses, and $0 < \alpha < 1$ is a parameter governing the utility from housing. $m(l_i, r, s)$ is the per-period mortgage payment of a household with a mortgage with current loan balance l_i, interest rate r, and remaining term s, which follows from the amortization of the loan:

$$m(l_i, r, s) = l_i \frac{r(1+r)^s}{(1+r)^s - 1}. \tag{1}$$

Renting a property yields per-period utility $\bar{u}$, which we assume is common to all households and fixed over time. All households discount the future at the common rate β.

Mortgages are long-term contracts for T periods that pay a discounted rate r_d for an initial time interval T_d, and subsequently pay a reset rate $R > r_d$ following this interval, unless the household refinances back into the discounted rate. To simplify and facilitate evaluating counterfactuals, we take both rates as given constant values. We also assume that T/T_d is a (positive) integer and that households can only refinance at the point at which the discounted rate expires. In what follows, we normalize by the length of this initial fixation period, treating it as a single time unit, i.e., we assume $T_d = 1$ and $T = 15$, and all rates are computed over the period T_d. Moreover, we assume that households do not change their loan balance (i.e., we rule out "cash-out refinancing"), and rule out maturity extensions (i.e., households in the model do not change the maturity of their loan at the point of refinancing). Households receive the loan amount at time $t = 0$, but make the first repayment at $t = 1$, which is also the first refinancing period. Hence, the loan balance of a mortgage with interest rate r evolves over time as follows:

$$l_{i,t+1}(r, l_{i,t}) = l_{i,t}(1+r) - m(l_{i,t}, r, s). \tag{2}$$

Mortgages are fully repaid after T periods. Thus, each household makes T payments over the life of the loan, the same as the duration of the mortgage contract.

At time $t = 0$, if they choose to buy a house, households choose the size of their mortgage loan $l_{i,0}$ to finance their house h_i, where $\omega_i = h_i/l_{i,0}$ denotes the inverse of the loan-to-value at origination. In each subsequent period, households can refinance their mortgage at the discounted rate r_d; to do so, they have to pay refinancing costs equal to $k_{i,t} = k_i\varepsilon_{i,t}$, where k_i is a persistent component of the refinancing cost for household i and $\varepsilon_{i,t}$ is a transitory component. We assume that $\varepsilon_{i,t}$ is a non-negative random variable, independent and identically distributed across households and over time, with mean equal to one, with cumulative distribution function $F(\varepsilon_{i,t})$ and density $f(\varepsilon_{i,t})$. Hence, each household's average refinancing costs equal their persistent component of refinancing costs, i.e., $E(k_{i,t}) = k_i$.

Households are heterogeneous in their per-period valuation for housing v_i (to capture the heterogeneity of initially chosen loan sizes seen in the data) and in their persistent component k_i of the cost of refinancing (to capture the household heterogeneity in refinancing for a given loan balance). We assume that, at the time of originating a mortgage, households perfectly know their valuation for owned housing v_i, but only receive a signal of their persistent component k_i of refinancing costs and thus of their average refinancing costs over time. Specifically, we assume that k_i is correlated with the origination cost k_i^o according to $k_i = k_i^o\varepsilon_{i,0}$, where $\varepsilon_{i,0}$ is a non-negative random variable that is realized after the origination of the mortgage and before the first refinancing opportunity. Thus, the precision of the signal negatively depends on the variance of $\varepsilon_{i,0}$. We assume that $\varepsilon_{i,0}$ is independent and identically distributed across households, with mean equal to one, with cumulative distribution function $F_0(\varepsilon_{i,0})$ and density $f_0(\varepsilon_{i,0})$.

Valuations and origination costs are distributed according to the cumulative joint distribution function $G_o(v_i, k_i^o)$ with density $g_o(v_i, k_i^o)$. Hence, the joint density of valuations and persistent refinancing costs equals $g(v_i, k_i) = \int_0^{+\infty} g_o(v_i, k_i/\varepsilon_{i,0}) f_0(\varepsilon_{i,0}) \frac{1}{\varepsilon_{i,0}} d\varepsilon_{i,0}$.

Intuitively, in the model, households learn about their persistent ongoing mortgage refinancing costs from the costs/hassle that they experience during the process of mortgage origination. The extent to which this initial origination process is informative about ongoing mortgage refinancing costs is dictated by the variance of $\varepsilon_{i,0}$. If this variance is zero, the initial process of mortgage origination perfectly informs households about the future persistent cost of refinancing. Alternatively, if this variance is very high, households learn little about the future process of refinancing from their experience during origination, since k_i is likely quite different from k_i^o.

We now solve the model to determine two household choices: (1) whether or not to refinance at each opportunity; and (2) the optimal size of the initial loan $l^*_{i,0}(v_i, k^o_i)$.

3.1 Optimal Refinancing

Households refinance when their refinancing costs are below a threshold that depends on their loan size. Hence, households with larger loans are more likely to refinance. Similarly, because the loan is amortizing, each household's incentives to refinance decline over time as the outstanding balance decreases; notably, some households (almost) always refinance because they have a low value of the persistent component k_i of the cost of refinancing.

We solve for the optimal refinancing path by backward induction. Consider period T, which is the last refinancing period, and households with a beginning-of-period (i.e., before making a payment) loan balance l_i (we suppress the subscript t for simplicity). Such households refinance if their refinancing cost $k_{i,T}$ is below the benefit of refinancing $k^*_i(T)$:

$$\begin{aligned} k^*_i(T) =& m(l_i, R, 1) - m(l_i, r_d, 1) \\ =& l_i(R - r_d). \end{aligned}$$

The benefit of refinancing depends on the difference between the interest rates $R - r_d$, as well as on the loan balance l_i.

We can define the expected (i.e., prior to the realization of the transitory component $\varepsilon_{i,t}$) value function $V_T(k_i, l_i)$ of a household with persistent cost k_i as the expected payment:

$$\begin{aligned} V_T(k_i, l_i) &= \int_0^{+\infty} \min\left(m(l_i, r_d, 1) + k_i\varepsilon_{i,T}, m(l_i, R, 1)\right) dF(\varepsilon_{i,T}) \\ &= \int_0^{k^*_i(T)/k_i} \left(m(l_i, r_d, 1) + k_i\varepsilon_{i,T}\right) dF(\varepsilon_{i,T}) + \int_{k^*_i(T)/k_i}^{+\infty} m(l_i, R, 1) dF(\varepsilon_{i,T}), \end{aligned} \tag{3}$$

where $k^*_i(T)/k_i$ is the cutoff point in the distribution of the transitory component $\varepsilon_{i,t}$ that determines household refinancing.[18]

Similarly, in the previous period $T - 1$, households' expected value function equals the

[18]When $\varepsilon_{i,T}$ equals $k^*_i(T)/k_i$, then $k_i \cdot \varepsilon_{i,T}$ equals $k^*_i(T)$ and the cost is exactly equal to the benefit of refinancing.

discounted sum of expected future payments:

$$
\begin{aligned}
V_{T-1}(k_i, l_i) = & \int_0^{+\infty} \min \left(m(l_i, r_d, 2) + k_i \varepsilon_{i,T-1} + \beta V_T \left(k_i, l_i(1+r_d) - m(l_i, r_d, 2) \right), \ldots \right. \\
& \left. m(l_i, R, 2) + \beta V_T \left(k_i, l_i(1+R) - m(l_i, R, 2) \right) \right) dF(\varepsilon_{i,T-1}) \\
= & \int_0^{k_i^*(T-1)/k_i} \left(m(l_i, r_d, 2) + k_i \varepsilon_{i,T-1} + \beta V_T \left(k_i, l_i(1+r_d) - m(l_i, r_d, 2) \right) \right) dF(\varepsilon_{i,T-1}) + \\
& \int_{k_i^*(T-1)/k_i}^{+\infty} \left(m(l_i, R, 2) + \beta V_T \left(k_i, l_i(1+R) - m(l_i, R, 2) \right) \right) dF(\varepsilon_{i,T-1}),
\end{aligned}
$$

where

$$
\begin{aligned}
k_i^*(T-1) = & m(l_i, R, 2) + \beta V_T \left(k_i, l_i(1+R) - m(l_i, R, 2) \right) + \\
& - m(l_i, r_d, 2) - \beta V_T \left(k_i, l_i(1+r_d) - m(l_i, r_d, 2) \right)
\end{aligned}
$$

defines the monetary benefits of refinancing, such that households with $k_{i,t} \leq k_i^*(T-1)$ refinance, and households with $k_{i,t} > k_i^*(T-1)$ do not.

In a generic period t, the expected value function equals:

$$
\begin{aligned}
V_t(k_i, l_i) = & \int_0^{+\infty} \min \left(m(l_i, r_d, T-t+1) + k_i \varepsilon_{i,t} + \beta V_{t+1} \left(k_i, l_i(1+r_d) - m(l_i, r_d, T-t+1) \right), \ldots \right. \\
& \left. m(l_i, R, T-t+1) + \beta V_{t+1} \left(k_i, l_i(1+R) - m(l_i, R, T-t+1) \right) \right) dF(\varepsilon_{i,t}),
\end{aligned}
$$

and the benefits $k_i^*(t)$ determine the cutoff point in the cost distribution that characterizes household refinancing decisions.

Therefore, we can describe the optimal refinancing policy as follows:

$$
r(l_i, k_{i,t}) = \begin{cases} r_d & \text{if } k_{i,t} \leq k_i^*(t) \\ R & \text{otherwise.} \end{cases} \tag{4}
$$

Hence, households with a lower persistent component k_i are more likely to refinance and pay the discounted rate r_d than households with a higher k_i. Moreover, the refinancing behavior of each household varies over time depending on the realization of the transitory shock $\varepsilon_{i,t}$. Generally, because the transitory shock $\varepsilon_{i,t}$ is multiplicative, its realization has a smaller effect on the refinancing activity of borrowers with low k_i, and a larger effect on that of borrowers with high k_i.

3.2 Optimal Loan Size

Households choose the loan size that maximizes their value function at origination, given their valuation for housing v_i and origination cost k_i^o. The value at origination equals:

$$W_0(v_i, k_i^o) = \max_{l_{i,0}} \sum_{t=0}^{+\infty} \beta^t v_i(\omega_i l_{i,0})^\alpha - k_i^o - \beta \int_0^{+\infty} V_1(k_i^o \varepsilon_{i,0}, l_{i,0}) dF_0(\varepsilon_{i,0}), \tag{5}$$

where the loan-to-value at origination equals $l_{i,0}/h_i = 1/\omega_i$ and k_i^o is the mortgage origination cost described above. Households do not know the exact value of their future refinancing cost and thus they form their expectations based on the available signal, which is their origination cost.

The optimal loan size $l_{i,0}^*(v_i, k_i^o)$ satisfies the first-order condition

$$\frac{\alpha\omega_i v_i(\omega_i l_{i,0}^*)^{\alpha-1}}{1-\beta} - \beta \frac{\partial}{\partial l_{i,0}} \int_0^{+\infty} V_1(k_i^o \varepsilon_{i,0}, l_{i,0}^*) dF_0(\varepsilon_{i,0}) = 0. \tag{6}$$

Hence, the optimal loan size depends directly on the household valuation for housing v_i, and indirectly on the origination costs k_i^o, because it is correlated with the expected future mortgage payments through the optimal refinancing policy $r(l_{i,t-1}, k_{i,t})$ described above. The refinancing policy in (4) highlights that refinancing costs determine the extent to which households make mortgage payments at the higher reset rate rather than at the lower discounted rate. This is because obtaining the cheaper discounted rate in a greater number of periods requires incurring the refinancing cost $k_{i,t}$ across a greater expected number of refinancing opportunities.

Given the optimal loan size, we can define $v_i^*(k_i^o)$ as the valuation for housing of a household that is indifferent between buying a house and getting a mortgage or renting a property:

$$W_0(v_i^*, k_i^o) = \frac{\bar{u}}{1-\beta}, \tag{7}$$

where $\bar{u}$ is a per-period utility of the outside rental option. This extensive-margin condition determines whether or not households enter the housing market rather than rent:

households with a high valuation v_i and a low cost k_i^o enter the housing and mortgage market.

The precision of information that households have about their future refinancing costs plays into both optimal loan size (the intensive margin described in equation (6)) and whether or not households enter the housing market in the first place (the extensive margin described in equation (7)). On the intensive margin, a higher k_i generates an incentive to scale back the size of the initial loan, and on the extensive margin, a higher k_i may be a deterrent to entering the mortgage market in the first place. Conditional on the other parameters including their housing valuation v_i, the extent to which this effect operates depends on the variance of $\varepsilon_{i,0}$. If this variance is small, households choose an initial loan size that is strongly correlated with the origination costs k_i^o and thus with the persistent component k_i of refinancing costs. If the variance of $\varepsilon_{i,0}$ is larger, households have less precise information at origination to evaluate their future mortgage costs. Hence, their initial loan size will be weakly correlated with the cost k_i. The variance of the transitory component $\varepsilon_{i,t}$ of refinancing costs similarly affects households' optimal initial loan size, because a larger variance of $\varepsilon_{i,t}$ makes it more difficult for households to predict their refinancing activity, and thus the rates of their future mortgage payments.

Equation (7) shows that origination costs k_i^o also capture any household constraints to becoming homeowners. Once again, the precision of households' information at origination, captured by the variance of $\varepsilon_{i,0}$, critically affects this adjustment. This condition will play an important role in our counterfactual analysis as it determines how initial homeownership and mortgage takeup change. We return to these issues in greater detail when evaluating counterfactuals.

3.3 Aggregation: Mortgage Stocks in Steady-State

We calculate the total stock of mortgages that pay the discounted rate and the reset rate, assuming that the economy is in steady state.

It is useful in this calculation to recursively define the endogenous cumulative distribution function $H_t(\cdot)$ and its associated density $h_t(\cdot)$ of loan balances t periods after origination, given the evolution of the loan balances in (2), and the refinancing policy described in (4).

This distribution evolves as follows:

$$H_0(z) = \iint\limits_{\{(v_i,k_i^o):v_i \geq v_i^*(k_i^o) \cap l_{i,0}^*(v_i,k_i^o) \leq z\}} g_o(v_i, k_i^o) dv_i dk_i^o,$$

$$H_t(z) = \int\limits_{\{l_{i,t-1}:l_{i,t}(r,l_{i,t-1}) \leq z\}} h_{t-1}(l_{i,t-1}) dl_{i,t-1}.$$

We next define three groups (0, 1, 2) of mortgages. Group 0 comprises the mortgages of households who took a mortgage of initial size $l_{i,0}^*(v_i, k_i^o)$ and are on their initial discount period. The aggregate number $N_0(r_d)$ and aggregate balance $Q_0(r_d)$ of mortgages of this group equal:

$$N_0(r_d) = M \int_{-\infty}^{+\infty} \int_{v_i^*(k_i^o)}^{+\infty} g(v_i, k_i^o) dv_i dk_i^o, \tag{8}$$

$$Q_0(r_d) = N_0(r_d) \int_0^{+\infty} z h_0(z) dz = M \int_{-\infty}^{+\infty} \int_{v_i^*(k_i^o)}^{+\infty} l_{i,0}^*(v_i, k_i^o) g_o(v_i, k_i^o) dv_i dk_i^o. \tag{9}$$

To gain an intuition for equation (8), recall that a mass M of households enters the market in each time period. (Discounted) mortgage takeup among these households is determined by whether or not they satisfy the extensive margin condition $v_i \geq v_i^*(k_i^o)$, with the outer integral integrating across the k_i^o distribution. Equation (9) follows by weighting these mortgages by their initial loan sizes.

The second group comprises the mortgages of all households who refinanced and pay the discounted rate. In each period $t \in \{1, \ldots, T-1\}$, the number $N_{1,t}(r_d)$ of mortgages in this group equals:

$$N_{1,t}(r_d) = N_0(r_d) \int\limits_{\{l_{i,t}:r(l_{i,t},k_{i,t})=r_d\}} h_t(l_{i,t}) dl_{i,t} \tag{10}$$

Equation (10) combines all borrowers who have refinancing costs $k_{i,t}$ lower than the benefits $k_i^*(t+1)$, and thus have policy functions $r(l_{i,t}, k_{i,t}) = r_d$. Thus, the aggregate number $N_1(r_d)$ of mortgages of this group equals:

$$N_1(r_d) = \sum_{t=1}^{T-1} N_{1,t}(r_d). \tag{11}$$

The aggregate balance of this group is the sum of the balances of the different cohorts who pay the discounted rate r_d. The aggregate balances $Q_{1,t}(r_d)$ of these cohorts evolve as follows:

$$Q_{1,t}(r_d) = N_0(r_d) \int_{\{l_{i,t}:r(l_{i,t},k_{i,t})=r_d\}} l_{i,t} h_t(l_{i,t}) dl_{i,t}.$$

Thus, the aggregate balance equals $Q_1(r_d) = \sum_{t=1}^{T-1} Q_{1,t}(r_d)$.

The third group comprises the mortgages of all households who did not refinance, and pay the reset rate. In each period $t \in \{1, \dots, T-1\}$, the number $N_{2,t}(R)$ of mortgages in this group equals:

$$N_{2,t}(R) = N_0(r_d) \int_{\{l_{i,t}:r(l_{i,t},k_{i,t})=R\}} h_t(l_{i,t}) dl_{i,t}, \tag{12}$$

which is the set of borrowers who have refinancing costs above the benefits $k_i^*(t+1)$, and thus have policy functions $r(l_{i,t}, k_{i,t}) = R$. Thus, the aggregate number of households who pay the reset rate equals

$$N_2(R) = \sum_{t=1}^{T-1} N_{2,t}(R). \tag{13}$$

The aggregate balance $Q_2(R)$ of this group is the sum of the balances of the different cohorts who pay the reset rate R: $Q_2(R) = \sum_{t=2}^{T} Q_{2,t}(R)$, where $Q_{2,t}(R)$ evolves as follows:

$$Q_{2,t}(R) = N_0(r_d) \int_{\{l_{i,t}:r(l_{i,t},k_{i,t})=R\}} l_{i,t} h_t(l_{i,t}) dl_{i,t}.$$

The above expressions can be directly mapped to the empirically observed stock of mortgages in each category, under the assumption that the market is in steady state.

3.4 Cross-Subsidy

To calculate the cross-subsidy across different households, we consider a benchmark case in which all mortgages have a constant interest rate r_c for their entire duration. In Section 5, we consider several values of this constant interest rate.

Under the constant interest rate r_c, households do not need to refinance and their mortgage payments are constant over time. Hence, their optimal loan size $l_{i,0}^{**}(v_i, k_i^o)$

maximizes the value function at origination (5) evaluated at $\varepsilon_{i,t} = 0$ for all $t > 0$, with a constant payment stream $m(l_{i,0}, r_c, T)$. The expression for optimal loan size simplifies to:

$$
\begin{aligned}
l_{i,0}^{**}(v_i, k_i^o) =& \frac{1}{\omega_i}\left(\frac{1-\beta}{\alpha\omega_i v_i}\left(\sum_{t=1}^{T}\beta^t \frac{\partial m(l_{i,0}, r_c, T)}{\partial l_{i,0}}\right)\right)^{\frac{1}{\alpha-1}} \\
=& \frac{1}{\omega_i}\left(\frac{\beta(1-\beta^T)}{\alpha\omega_i v_i}\frac{r_c(1+r_c)^T}{(1+r_c)^T-1}\right)^{\frac{1}{\alpha-1}}.
\end{aligned} \tag{14}
$$

The aggregate number $N(r_c)$ and aggregate balance $Q(r_c)$ of mortgages in this scenario then equals:

$$
\begin{aligned}
N(r_c) &= MT\int_{-\infty}^{+\infty}\int_{v_i^{**}(k_i^o)}^{+\infty} g_o(v_i, k_i^o)dv_i dk_i^o, \\
Q(r_c) &= M\sum_{t=1}^{T}\gamma_{r_c}(t-1)\int_{-\infty}^{+\infty}\int_{v_i^{**}(k_i^o)}^{+\infty} l_{i,0}^{**}(v_i, k_i^o)g_o(v_i, k_i^o)dv_i dk_i^o,
\end{aligned}
$$

where we define

$$
\gamma_{r_c}(t-1) = \frac{l_{i,t}(r_c, l_{i,0})}{l_{i,0}} = \frac{(1+r_c)^T - (1+r_c)^t}{(1+r_c)^T - 1},
$$

as the beginning-of-period-t share of the initial loan still to be repaid, and $v_i^{**}(k_i^o)$ is the valuation of a household that is indifferent between buying a house and getting a mortgage, or renting a property in this constant rate scenario. Thus, households still face the origination cost k_i^o that, as we recount above, includes additional household constraints to homeownership, but no subsequent refinancing costs, i.e., $k_{i,t} = 0$ for $t > 0$.

Based on this counterfactual constant rate r_c, the estimated parameters of the model, and the observed discounted rate r_d and reset rate R, we can calculate the differences in mortgage market outcomes between the current and counterfactual scenarios for each household (v_i, k_i^o). These outcomes include differences in loan sizes and mortgage payments between current and counterfactual scenarios. They also include a measure of the lifetime cross-subsidy paid or received by the household. This can be measured as the household-level reduction or increase (when comparing current and counterfactual scenarios) in the "all-in" interest rate including any refinancing costs. These household-level calculations can be aggregated up at the group level using the baseline model, or indeed, using an extended version of the model in which we estimate group-specific parameters. We describe this extended model next.

3.5 Multiple Groups

The richness of our data allows us to calculate subsidies across different groups based on observable demographic characteristics. We focus on two specific household groupings. The first groups households by income, and the second looks at households located in different UK regions.

Understanding variation in the extent of cross-subsidies paid or received along the income distribution helps us to understand how the design of the financial system contributes to the inequality of financial wealth, to the extent that wealth and income are correlated. We also look at the extent of regional variation in mortgage cross-subsidies given the importance of regional re-distribution through the mortgage market.

We extend the model to accommodate and interpret such heterogeneity. Consider different groups based on observable characteristics and indexed by $j = 1, ..., J$. Let M_j and $G_{oj}(v_i, k_i^o)$ be the measure and the cumulative distribution function of household housing preferences v_i and origination costs k_i^o in group j, respectively. Following the analysis of previous subsections, we can define the variables $N_{0,j}(r_d), Q_{0,j}(r_d), \ldots, Q_{2,j}(R)$ for each group j, and proceed with our counterfactual comparisons as before using this extended model.

We next turn to acquire quantitative estimates of the model's parameters and an assessment of the model-implied cross-subsidy by mapping the model to the data.

4 Quantitative Analysis

The model does not admit an analytic solution for all endogenous outcomes. As a result, we choose the parameters that best match moments of the data with the corresponding moments computed from the numerical solution of the model in steady state. We then study the quantitative implications of the model evaluated at the estimated parameters.

4.1 Estimation

We fix a subset of parameters, often reading them directly from the data, and we estimate the remaining parameters of the model to best match key moments of the mortgage data.

Specifically, we set the unit of time in the model to be $T_d = 2$ years, which is the modal initial fixation period in the UK mortgage market over the sample period; we then set the mortgage maturity at $T = 15$ periods, to give us the modal mortgage origination maturity of 30 years. We set the discount rate at $\beta = 0.95^2 = 0.9025$ to correspond to our assumption on the unit of time.

We read the annual interest rates on discounted and reset rate mortgages directly from the underlying data, using value-weighted averages of the corresponding rates in the 2015H1 sample, and compound them to correspond to $T_d = 2$ years. Annual average discounted and reset rates equal 320 bps and 372 bps in our sample, meaning $r_d = 650$ bps and $R = 759$ bps over two years.

We set the loan-to-value ratio at origination common across households at 80 percent, close to the modal value in our data, so $\omega = 1.25$.

We read market size M from the data, as follows. The total number of mortgages in the model equals:

$$N_0(r_d) + N_1(r_d) + N_2(R) = MT \int_{-\infty}^{+\infty} \int_{v_i^*(k_i^o)}^{+\infty} g_o(v_i, k_i^o) dv_i dk_i^o. \tag{15}$$

Hence, we compute the market size M by dividing the total number of mortgages $N_0(r_d) + N_1(r_d) + N_2(R)$ by their maturity T and by the share of households who own a property $\int_{-\infty}^{+\infty} \int_{v_i^*(k_i^o)}^{+\infty} g_o(v_i, k_i^o) dv_i dk_i^o$.

We estimate all remaining parameters by applying some assumptions about distributions. We assume that households' valuation v_i follows a lognormal distribution, i.e., $\log(v_i)$ follows a normal distribution with mean μ_v and standard deviation σ_v. We assume that the origination cost k_i^o follows a mixture distribution of two lognormal distributions, which allows for a bimodal distribution of origination costs. We model the persistent component of refinancing costs as $k_i = k_i^o \varepsilon_{i,0}$, which is correlated with the origination cost. As a result, the distribution of refinancing costs could be bimodal as well, with some households with low refinancing costs and others with high refinancing costs, and heterogeneity within these two household groups. With probability η, $\log(k_i^o)$ follows a normal distribution with mean

μ_{k1} and standard deviation σ_{k1}; with probability $1-\eta$, $\log(k_i^o)$ follows a normal distribution with mean μ_{k2} and standard deviation σ_{k2}. Without loss of generality, we denote type-1 as households whose average cost is lower than the average cost of type-2 households—i.e., $\exp\left(\mu_{k1}+\frac{\sigma_{k1}^2}{2}\right) \leq \exp\left(\mu_{k2}+\frac{\sigma_{k2}^2}{2}\right)$. We do not restrict the variances of these distributions, and thus the type-2 distribution does not necessarily first-order stochastically dominate the type-1 distribution. Clearly, the assumption of lognormality implies that some type-1 households have higher costs than some type-2 households. We set $\eta = 0.5$, and the correlation between v_i and k_i^o to zero, because the empirical moments that we employ in the estimation do not allow us to separately identify these parameters, as we explain in more detail below.

We further assume that $\varepsilon_{i,t}$ follows a lognormal distribution with parameters μ_ε and σ_ε. We set the mean of $\varepsilon_{i,t}$ equal to one, meaning that $E(k_{i,t}) = k_i$, and hence $\mu_\varepsilon = -\sigma_\varepsilon^2/2$.

We estimate two versions of the model under two different assumptions about the distribution of $\varepsilon_{i,0}$. Version 1 is simply a degenerate distribution, with $\varepsilon_{i,0} = 1$ with probability one. Here, households perfectly know the persistent component of their refinancing costs k_i at origination. Put differently, they know their average refinancing costs because $E(k_{i,t}) = k_i$. Of course, households still face ex-ante uncertainty about their future refinancing costs because they do not know the temporary component of refinancing costs which is governed by $\varepsilon_{i,t}$. Version 2 of the model assumes that $\varepsilon_{i,0}$ is governed by the same distribution as $\varepsilon_{i,t}$.[19] In this version of the model, households obtain a noisy signal of the persistent component of their refinancing costs (and thus of their average refinancing costs) at the point of mortgage origination. Therefore, their loan size and participation decisions will exhibit weaker correlation with their refinancing costs than in version 1. The two versions differ in the precision of borrowers' information at origination, but require the same number of parameters to be estimated.

Finally, our estimation recovers the parameter α of the utility function and the level of the outside option $\bar{u}$.

We search for the vector of 9 parameters $\psi = (\mu_v, \sigma_v, \mu_{k1}, \sigma_{k1}, \mu_{k2}, \sigma_{k2}, \sigma_\varepsilon, \alpha, \bar{u})$ that minimizes the distance between selected moments in the data and the corresponding moments of the model. More specifically, for each combination of these unknown parameters, we

[19] In other words, in this version of the model, if shocks to household attention/distraction (i.e., temporary refinancing costs shocks) are drawn from a high-variance distribution, this also means that households learn less from their mortgage origination costs about their persistent refinancing costs and vice versa.

solve the model shown in Section 3 to find households' optimal policies, characterized by their choice between buying a house with a mortgage or renting a property, and, if they choose to participate in the mortgage market, their mortgage loans at origination $l^*_{i,0}(v_i, k^o_i)$ and their optimal sequence of refinancing. Based on these household policies, we compute the following aggregate moments:

1. the average loan balance for mortgages on the discounted rate;
2. the standard deviation of the loan balance of mortgages on the discounted rate;
3. the average loan balance for mortgages on the reset rate;
4. the standard deviation of the loan balance of mortgages on the reset rate;
5. the average remaining maturity of mortgages on the discounted rate;
6. the standard deviation of the remaining maturity of mortgages on the discounted rate;
7. the average remaining maturity of mortgages on the reset rate;
8. the standard deviation of the remaining maturity of mortgages on the reset rate;

9-14. the shares of mortgages on the discounted rate for the following partition of the loan balance distribution: $[0-5]$ percentile, $(5-25]$ percentile, $(25-50]$ percentile, $(50-75]$ percentile, $(75-95]$ percentile, and $(95-100]$ percentile;

15. the share of mortgages on the reset rate in 2015H1 that paid the discounted rate in 2017H1;
16. the share of homeowners, i.e., the fraction of households that enter the housing market and choose to purchase a house and take on a mortgage.

Section 2 outlines several filters that we apply to the data. One of these filters is that the outstanding mortgage balance exceeds £30,000, and for consistency, we apply the same filter when computing moments 1 to 15 in the model.

The minimum-distance estimator chooses the parameters that minimize the criterion function:

$$(\mathbf{m}(\psi) - \mathbf{m}_S)' \, \Omega \, (\mathbf{m}(\psi) - \mathbf{m}_S),$$

where $\mathbf{m}(\psi)$ is the vector of moments computed from the model at the parameter vector ψ and $\mathbf{m}_S$ is the vector of corresponding sample moments. Ω is a symmetric, positive-definite matrix; in practice, in order for the moments to have a similar scale, we use a diagonal matrix whose elements are those on the main diagonal of the inverse of the matrix $E(\mathbf{m}_S'\mathbf{m}_S)$.

We estimate six cases of the model. Two baseline versions follow from the two different assumptions about the distribution of $\varepsilon_{i,0}$ recounted above. In both cases, we pool together all mortgages in our data and assume that all households can be characterized by a single distribution $G_o(v_i, k_i^o)$, as well as common σ_ε, α, and $\bar{u}$ parameters. This entails estimating 9 parameters using the 16 moments listed above.

We also pursue the estimation in cases with richer borrower heterogeneity. The first one estimates the model separately for different income groups, and the second one estimates the model separately for different geographic areas of the UK. We estimate all these cases with richer heterogeneity in the two versions with the two different assumptions about the distribution of $\varepsilon_{i,0}$. In each case, we set group-specific market sizes M_j and to estimate group-specific parameters of the distributions $G_{oj}(v_i, k_i^o)$, as well as σ_{ε_j}, α_j, and $\bar{u}_j$ for each group j (denoting either income groups or geographical areas).[20] This gives us additional flexibility to capture heterogeneity across groups in preferences, costs, and ultimately refinancing activities. Of course, when we estimate these parameters, we do so using an expanded set of group-specific moments in each case.

We consider 12 income groups based on the following percentiles of the distribution of reported incomes in the PSD: 0-10, 10-20, 20-30, 30-40, 40-50, 50-60, 70-80, 80-85, 85-90, 90-95, and 95-100. We also consider 12 broad regions and devolved administrations of the UK, namely North-East, North-West, Yorkshire and the Humber, East Midlands, West Midlands, East of England, Greater London, South East, South West, Wales, Scotland, and Northern Ireland.[21]

Hence, we estimate a total of 108 parameters (9 parameters for each of the 12 groups) using a total of 192 moments (16 moments listed above for each of the 12 groups).

[20]Hence, in the case of version 2, we implicitly assume that the heterogeneity of refinancing costs across groups is (perfectly) positively correlated with the noise of their information at origination, i.e., groups with smaller temporary shocks to refinancing costs have more precise information about their future refinancing costs at origination than groups with larger temporary shocks to refinancing costs.

[21]These are the 12 NUTS-1 regions of the UK, where NUTS stands for Nomenclature of Territorial Units for Statistics.

4.2 Sources of Identification

The model is highly nonlinear, so (almost) all parameters affect all outcomes. That said, the identification of certain parameters does rely more heavily on particular moments in the data.

More specifically, moments characterizing the distributions of loan sizes on the discounted and the reset rate, those characterizing the distributions of remaining maturities in each mortgage category, and the shares of mortgages in the two categories together identify the parameter α, and the parameters of the distributions of household preferences v_i and the persistent component of costs k_i. Notably, households' initial loan amounts—and, thus over time, their loan balances—depend on their housing preferences v_i, as well as their expected refinancing costs which on average equal k_i. Moreover, for every mortgage, the parameter α affects the sensitivity of the initial loan size to expected mortgage payments, and thus to interest rates, as equations (6) and (14) show.

If the cost $k_{i,t}$ was prohibitively high for all borrowers, almost all mortgages would be on the reset rate, and conversely, if $k_{i,t}$ was extremely low for all borrowers, all mortgages would be on the discounted rate. Hence, the shares of mortgages on the reset rate are informative about the parameters of the distribution of the refinancing cost $k_{i,t}$ and its components.

Given a value of $k_{i,t}$, borrowers have stronger financial incentives to refinance if they have a large loan balance, meaning that the share of mortgages on the discounted rate should be increasing in the loan balance. The rate of change of the share of mortgages on each rate as loan size changes is informative about the heterogeneity in $k_{i,t}$. The increase is fast if the heterogeneity across households is small, whereas it is slow if the heterogeneity is large. Our assumption that k_i follows a mixture distribution allows us to flexibly capture different rates of increase in the share of mortgages on the discounted rate at different percentiles of the loan balance distribution. This means that the change in the share in the two categories of mortgages at different levels of the loan balance contributes to the identification of the refinancing cost heterogeneity parameters σ_{k1} and σ_{k2} of the mixing distribution. Because we allow the support of the two k_i distributions to overlap, it is difficult to separately identify the mixing probability η; therefore, we set it to $\eta = 0.5$. Moreover, identifying any correlation between v_i and k_i would require rich within-borrower moments; our mapping to the stock means that all our moments are cross-sectional (except for the share of mortgages on the reset rate that later pay the discounted rate, moment 15 above), so we set this correlation to zero by assumption.

The share of mortgages that transition from paying the reset rate to paying the discounted rate is informative about the within-borrower heterogeneity in refinancing costs, and thus identifies the parameter σ_ε. If refinancing costs were fixed over time for each borrower, because loan balances decline over time, borrowers' optimal refinancing policy would be deterministic: it would be characterized by a borrower-specific cutoff date $T_{max}(v_i, k_i)$, such that a (v_i, k_i)-borrower always refinances before $T_{max}(v_i, k_i)$ and never does after $T_{max}(v_i, k_i)$. Transitions from the reset rate to the discounted rate violate this deterministic refinancing policy, and therefore identify the within-borrower, transitory variation in refinancing costs governed by σ_ε.

Moreover, our data does not allow us to identify households' information and beliefs at origination about their future refinancing costs, captured by the variance of $\varepsilon_{i,0}$. Hence, we set it to different values in the two versions, and the similarity of the results of these two versions will allow us to establish the robustness of our results to differences in households' information at origination.[22]

Finally, the share of owners versus renters identifies the level of outside option utility $\bar{u}$.

4.3 Parameters and Model Fit

Table 3 reports the parameters of the model for the six cases of the estimated model: aggregate, income group-specific, and geography-specific, each one with the two versions with different assumptions about the precision of households' information at origination. The top of the table reports the fixed parameters, which are common across cases and across groups.

The main body of the table reports the estimated parameters. Columns (1) and (2) report the parameter estimates for the baseline versions that use UK-wide moments, and their asymptotic standard errors in parentheses.[23] The model in column (1) assumes $\varepsilon_{i,0} = 1$

[22] Of course, we could set the value of $\varepsilon_{i,0}$ to alternative, higher values than those that we choose.

[23] To obtain standard errors, we compute the covariance matrix of the moments $\mathbf{m}_S$ by bootstrapping. Specifically, for N_s bootstrap resamples of the data, the covariance matrix of the moments $\mathbf{m}_S$ equals

$$W_{N_s} = N_s^{-1} \sum_{n_s=1}^{N_s} (\mathbf{m}_{n_s} - \mathbf{m}_S)(\mathbf{m}_{n_s} - \mathbf{m}_S)', \tag{16}$$

where $\mathbf{m}_{n_s}$ is the vector of moments in resample n_s. We set the number of resamples N_s at 1,000.

with probability one, and thus at origination households know their persistent component of refinancing costs; the model in column (2) assumes that at origination households receive a noisy signal of their future persistent component of refinancing costs only.

Columns (3) and (4) report the estimates for the case that uses separate moments for each income group, and columns (5) and (6) for the case that uses separate moments for each region and devolved administration. Columns (3) and (5) assume that households have precise information about their persistent component of refinancing costs; columns (4) and (6) assume that households receive noise signals only. In columns (3)-(6), we report the weighted averages of the parameters across groups, as well as the weighted standard deviations of the parameters across groups (in parentheses), where the weights are the estimated market sizes M_j.

The bottom of Table 3 reports the calibrated market size M computed using equation (15); in columns (3)-(6), they correspond to the unweighted averages and standard deviations of M_j across groups. Note that several parameters are not easily comparable across columns. For example, the outside options $\bar{u}$ differ across groups in columns (3)-(6), and affect the estimated parameters of the valuation distribution. Other parameters, such as α, are more easily comparable across columns.

Baseline Models. We focus our discussions on the parameters in column (1) because the differences between those in columns (1) and (2) are small and the two versions of the model have quite similar implications for market outcomes. Nevertheless, we note the key differences between the parameters of the two versions, most notably in origination and refinancing costs, because the two versions differ mainly in these costs.

The estimated parameters in column (1) imply that households' valuation v_i has a median equal to 1.001, a mean equal to 1.012 and a standard deviation equal to 0.152 in the full population of borrowers (homeowners) and non-borrowers (renters). In the model, households with the lowest valuations are less likely to participate in the mortgage market, choosing instead to rent a property. This means that, among borrowers, valuations are higher, with median v_i equalling 1.066, mean 1.077, and standard deviation 0.128.

The estimate of the parameter $\alpha = 0.787$ implies modest concavity in household utility from housing. This value implies that a household with average v_i enjoys a utility flow of $v_i h^\alpha$, i.e., £$10,342$ over a two-year period from a house worth £$125,000$, for example. This

Table 3: Parameters

r	650	R 759	T 15	β 0.902	ω 1.250	η 0.500
	UK-WIDE		INCOME GROUPS		REGIONS	
	(1)	(2)	(3)	(4)	(5)	(6)
μ_v	0.001	0.001	0.001	0.001	0.001	0.001
	(0.004)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
σ_v	0.150	0.157	0.234	0.230	0.349	0.349
	(0.010)	(0.007)	(0.010)	(0.012)	(0.023)	(0.022)
μ_{k1}	4.883	5.763	4.755	4.957	4.925	5.004
	(1.066)	(1.158)	(0.595)	(0.710)	(0.424)	(0.530)
σ_{k1}	2.670	2.618	2.823	2.976	2.333	2.349
	(1.263)	(0.641)	(0.394)	(0.453)	(0.987)	(0.983)
μ_{k2}	9.164	9.196	9.088	9.256	9.188	9.197
	(0.634)	(0.484)	(0.184)	(0.096)	(0.041)	(0.025)
σ_{k2}	0.988	0.987	0.960	0.942	0.987	0.957
	(0.072)	(1.233)	(0.059)	(0.071)	(0.015)	(0.088)
σ_ϵ	1.048	0.926	0.992	0.942	1.041	0.892
	(0.180)	(0.142)	(0.033)	(0.035)	(0.056)	(0.159)
α	0.787	0.786	0.789	0.789	0.788	0.788
	(0.001)	(0.001)	(0.007)	(0.007)	(0.003)	(0.003)
$\bar{u}$	1,190	1,045	1,580	1,231	1,455	1,335
	(235)	(375)	(606)	(447)	(436)	(485)
M	379,145	379,145	27,850	27,850	31,894	31,894
			(10,407)	(10,407)	(15,298)	(15,298)

Notes: This table reports the estimated parameters. In columns (1) and (2), the numbers in parentheses refer to asymptotic standard errors of the parameter estimates. In columns (3)-(6), the numbers in parentheses refer to standard deviations of the parameter estimates across groups. Odd-numbered columns correspond to version 1 of the model, and even-numbered columns correspond to version 2 of the model.

translates into an annual yield of 4.054%, which is slightly lower than the average rental yield for the whole of the UK, but broadly in line with average rental yields reported for London in this period.[24]

In version 1 of the model, in which households know their persistent component of refinancing costs at origination, this persistent component k_i equals the origination costs k_i^o. Among homeowners/borrowers, the median origination cost/persistent component equals £634, its mean equals £$3,866$, and its standard deviation equals £$6,641$. However, households with the highest origination and refinancing costs are less likely to participate

[24] See, for example, Savill's UK Report on Rents and Returns, 2015.

in the mortgage market and choose to rent a property. Because our moments do not report any information on households who do not borrow (except for their share in the population), we obtain the distribution of costs k_i (as well as that of preferences v_i) in the full population by extrapolating those of borrowers out of sample. This leads us to estimate the persistent component k_i of refinancing costs across all households, including those that do not borrow, with a median that equals £3,003, a mean of £10,107, and a standard deviation of £117,213 in the full population. It is worth noting that in the counterfactual exercises, we retain origination costs k_i^o, but remove refinancing costs $k_{i,t}$. This results in a relatively small effect on our calculations of the large $k_{i,t}$ values estimated for non-participants in the baseline dual-rate economy.[25]

In version 2 of the model, in which households have noisy information about their persistent component of refinancing costs at origination, the persistent component k_i differs from the origination cost k_i^o. This version (whose parameters are displayed in column (2) of Table 3) displays slightly higher origination costs and persistent components of refinancing costs relative to version 1 in which households have precise information and these costs are equal. Among homeowners/borrowers, the median origination cost k_i^o equals £1,299, its mean equals £4,334, and its standard deviation equals £7,196; the median persistent component of refinancing cost k_i equals £761, its mean equals £4,049, and its standard deviation equals £9,622.

Interestingly, while we assume that preferences v_i and costs k_i are uncorrelated in the population of households, they are correlated among borrowers because of households' endogenous selection into the mortgage market—borrowers with high k_i enter the market only if their v_i is sufficiently high. The correlation coefficient among borrowers equals 0.340, suggesting that the effect of selection is appreciable.

The estimate of $\sigma_\varepsilon = 1.048$ in column (1) of Table 3 means that the standard deviation of $\varepsilon_{i,t}$ equals 1.413, which implies that the within-household variation in refinancing costs is non-trivial. This estimate of σ_ε means that the ratio $\frac{St.Dev.(k_i)}{St.Dev.(k_{i,t})}$ equals 0.58 in the population and 0.54 among borrowers—that is, the persistent household component k_i (cross-household variation) accounts for a slightly larger share of the standard deviation of the refinancing costs $k_{i,t}$ than the transitory component $\varepsilon_{i,t}$ (within-household variation). The estimate of

[25]We could reduce the large standard deviation by setting an upper bound to k_i equal to the maximum forgone refinancing benefits in the data.

Table 4: Model Fit

	Data	Version 1	Version 2
Mean Loan Balance, Discounted Rate	140,647	143,697	140,525
Standard Deviation Loan Balance, Discounted Rate	105,062	106,551	106,891
Mean Loan Balance, Reset Rate	112,692	113,741	110,977
Standard Deviation Loan Balance, Reset Rate	79,684	76,546	77,779
Mean Remaining Years, Discounted Rate	20.57	18.63	18.83
Standard Deviation Remaining Years, Discounted Rate	7.73	7.91	7.84
Mean Remaining Years, Reset Rate	16.84	15.56	15.54
Standard Deviation Remaining Years, Reset Rate	6.95	7.40	7.39
Share of Mortgages on Discounted Rate, 0-5 Percentile	52.72	52.82	53.03
Share of Mortgages on Discounted Rate, 5-25 Percentile	56.36	58.03	57.19
Share of Mortgages on Discounted Rate, 25-50 Percentile	61.48	60.12	59.96
Share of Mortgages on Discounted Rate, 50-75 Percentile	67.76	63.73	63.69
Share of Mortgages on Discounted Rate, 75-95 Percentile	73.77	72.10	71.72
Share of Mortgages on Discounted Rate, 95-100 Percentile	81.19	83.66	81.90
Transition from Reset Rate to Discounted Rate	16.52	16.42	16.79
Share of Owners	63.13	64.50	63.38
Criterion Function		0.0287	0.0240

Notes: This table reports the values of the empirical moments and of the moments calculated at the estimated parameters reported in columns (1) and (2) of Table 3.

$\sigma_\varepsilon = 0.926$ in column (2) of Table 3 implies that in version 2 the ratio $\frac{St.Dev.(k_i)}{St.Dev.(k_{i,t})}$ equals 0.65 in the population and 0.64 among borrowers. Hence, the transitory component accounts for a smaller share of the standard deviation of borrowers' total refinancing costs $k_{i,t}$ in version 2 than in version 1. Below, we provide more statistics on borrower refinancing costs $k_{i,t}$ and compare them to the benefits of refinancing in the two versions.

The value of the per-period outside option utility $\bar{u}$ equals £1,190, which implies an annual net utility from renting equal to $\frac{\bar{u}}{1+\beta^{1/2}}$=£610. Households with a net utility value (over and above all mortgage payments and refinancing costs) greater than this level from purchasing a house enter the mortgage market.

Table 4 presents a comparison between the empirical moments and the moments calculated from the model at the estimated parameters reported in columns (1) and (2) of Table 3, respectively. Overall, both versions of the model fit the data very well, with version 2 performing slightly better than version 1. Critically, both versions match two features of the data that underscore refinancing incentives across households and over time: on average, mortgages on the discounted rate have higher balances and are closer to issuance (have greater remaining maturity) than those on the reset rate.

Multiple Groups. Columns (3) and (4) in Table 3 report parameters for the model estimated on the 12 different income groups, and Columns (5) and (6) for the model estimated on the 12 UK regions and devolved administrations. As discussed above, some parameters are not easily comparable between Columns (3)-(6) and Columns (1)-(2), though many are similar in magnitude to those reported in Columns (1) and (2) for the UK-wide case. The parameters in Columns (3) and (4) for the different income groups exhibit some differences from those in Columns (5) and (6) because the heterogeneity across and within income groups differs from the heterogeneity across and within regional groups, which in turn affects the average and the standard deviations of some of the parameters.

The parameters that exhibit the most meaningful heterogeneity in the population are those characterizing the distribution of the origination costs k_i^o and thus of the persistent component of refinancing costs k_i. This fact is particularly interesting for our purposes because this heterogeneity affects heterogeneity in refinancing activity across groups, and thus contributes directly to our quantitative assessment of the cross-subsidy across groups.

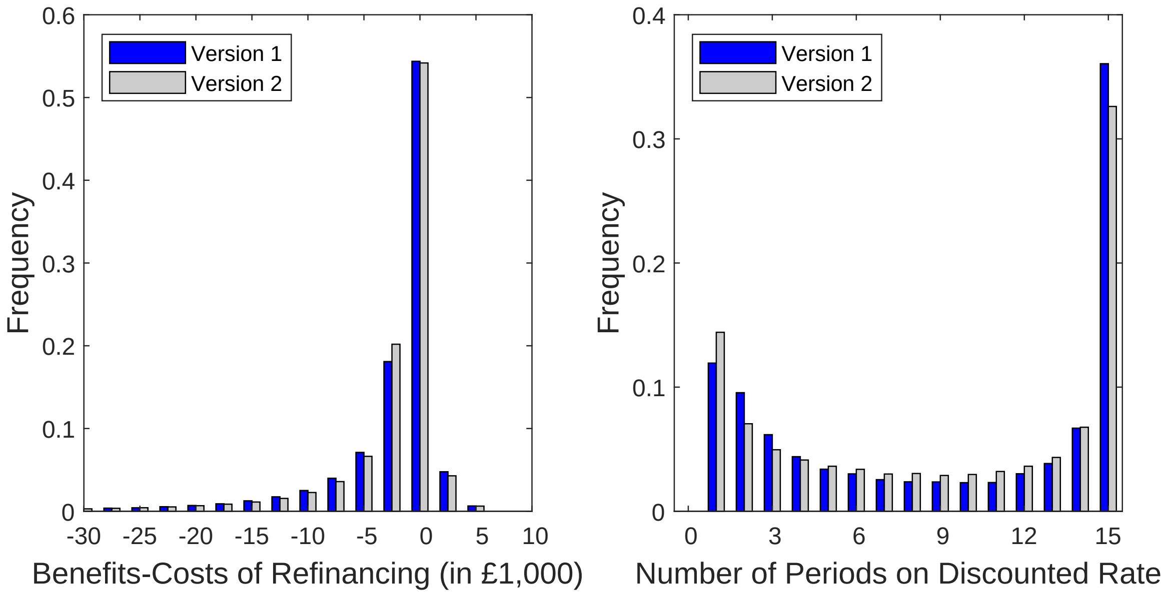
Moreover, the outside option $\bar{u}$ also displays significant heterogeneity in the population. This parameter is a key input into the "extensive margin" decision of households, i.e., whether or not they enter the mortgage market. The heterogeneity in this parameter across groups means that there are different sensitivities across groups of this extensive margin decision to changes in interest rates. This factor contributes to differences between the sensitivity of household participation decisions to interest rates in the multiple-group model and that in the baseline model.

While we do not report measures of goodness-of-fit across groups, we note that the model fits the group-specific moments well. This is perhaps not surprising given that Table 4 shows that the UK-wide model fits the aggregate data well; the same model might therefore be expected to fit as well or better at a lower level of aggregation.

4.4 Refinancing: Benefits and Costs

In this subsection, we discuss borrowers' refinancing behavior in the estimated UK-wide baseline model. We focus on version 1 of the model in which households know their persistent component of refinancing costs, because the qualitative patterns of refinancing behavior are

Figure 1: Distribution of Net Benefits of Refinancing



Notes: The left panel displays the histogram of the net benefits of refinancing. The right panel displays the histograms of the number of periods in which borrowers pay the discounted rate.

very similar in the two versions, noting the main differences between them.

The left panel of Figure 1 displays the full distribution of the net benefits of refinancing $k_i^*(t) - k_{i,t}$. The heterogeneity of net benefits is striking: in 60 percent of refinancing opportunities, borrowers have positive net benefits, thereby matching the aggregate share of mortgages on the discounted rate. The median estimated net benefit is positive (it equals $\pounds 304$), but the average net benefit is negative (it equals $-\pounds 2,619$), driven by the long left tail (the standard deviation equals $\pounds 12,240$). Some borrowers have extremely low measured net benefits, reflecting the fact that the model requires high costs to rationalize the non-refinancing behavior of a small group of borrowers with high loan balances and long maturities that would otherwise be expected to refinance. Moreover, we note that version 2 of the model exhibits slightly lower net benefits of refinancing than version 1: The median estimated net benefit equals $\pounds 255$ and the average net benefit equals $-\pounds 2,846$ (the standard deviation equals $\pounds 15,095$).

The distribution of gross benefits of borrowers who refinance has a median of $\pounds 1,039$, an average of $\pounds 1,357$, and a standard deviation of $\pounds 1,096$; their costs have a median of

£50, an average of £239, and a standard deviation of £466. The corresponding distribution of gross benefits of borrowers who do not refinance has a median of £885, an average of £1,084, and a standard deviation of £841; the costs of these non-refinancing borrowers have a median of £4,208, an average of £9,259, and a standard deviation of £18,030. The comparison of these statistics between borrowers who refinance and borrowers who do not shows that the difference in their respective costs is larger than that in their benefits. Hence, heterogeneity in refinancing costs $k_{i,t}$ is the main driver in the model of the heterogeneity in refinancing behavior observed across borrowers.

The heterogeneity of refinancing behavior is also apparent in the right panel of Figure 1, which displays the distribution of the number of periods on the discounted rate across individuals. No borrower always pays the reset rate because all of them receive the discounted rate at origination, in period $t = 0$. Approximately 10 percent of borrowers never refinance thereafter, many borrowers refinance occasionally, and 36 percent of borrowers always refinance. This heterogeneous distribution obtains because borrowers with low values of their persistent component k_i of refinancing costs (almost) always refinance, whereas borrowers with high values of k_i refinance only when they receive a temporary shock $\varepsilon_{i,t}$ that is low enough. The distribution of the share of periods on the discounted rate across individuals in model 1 first-order stochastically dominates that in version 2, because, as we recount above, the estimated net benefits of refinancing are slightly lower in version 2 than in version 1.

While our primary focus is on refinancing behavior, we should also point out that the heterogeneity of borrower refinancing propagates into substantial heterogeneity in the elasticities of their initial loan size with respect to discounted and reset rates. The mean borrower elasticity with respect to the discounted rate r_d equals -1.566 and its standard deviation equals 0.722. Borrowers with a lower k_i are more elastic to the discounted rate (and less elastic to the reset rate) than borrowers with a higher k_i because they are more likely to refinance regularly and thus pay the discounted rate—the elasticity to the discounted rate of lowest-k_i borrowers equals -2.603. The mean borrower elasticity with respect to the reset rate R among borrowers equals -0.466 and its standard deviation equals 0.813. These estimates are comparable to estimates from the US market, where DeFusco and Paciorek (2017) find elasticities between 1-2 percent for total mortgage debt (and 2-3 percent when ignoring substitution via second mortgages). Interestingly, in our setting some borrowers display a positive elasticity with respect to the reset rate, because if the reset rate increases (while keeping the discounted rate fixed), the benefits of refinancing increase, and thus some borrowers are more likely to refinance. This leads to a lower average expected interest rate, meaning that such borrowers will increase their initial loan size in response to a higher reset rate.

5 Counterfactual Analyses: Constant Interest Rate

We compare the outcomes for households in our estimated models under the dual rate structure, with a counterfactual in which all households simply pay a constant interest rate and have no need to refinance. We perform these comparisons for four different values of the constant interest rate, namely:

1. The average discounted rate, i.e., $r_c = 650$ bps.

2. The weighted average of the discounted and the reset rates, i.e.,

$$r_c = \frac{r_d(Q_0(r_d) + Q_1(r_d)) + RQ_2(R)}{Q_0(r_d) + Q_1(r_d) + Q_2(R)}. \tag{17}$$

   We calculate this weighted average using the aggregate balances in the data and obtain $r_c = 683$ bps.

3. The rate that yields the same revenue as the composite of the populations on the discounted rate and the reset rate.

   More precisely, in the baseline case aggregate lender revenues from all mortgages (on both discounted and reset rates) equal:

$$r_d(Q_0(r_d) + Q_1(r_d)) + RQ_2(R). \tag{18}$$

   Under the assumption of aggregate lender revenues remaining constant across the two scenarios, the interest rate r_c must satisfy:

$$r_cQ(r_c) = r_d(Q_0(r_d) + Q_1(r_d)) + RQ_2(R). \tag{19}$$

In practice, this equality yields $r_c = 700$ bps.[26]

4. The average reset rate, i.e., $r_c = 759$ bps.

The values of the constant interest rate in cases 1 and 4 likely represent lower and upper bounds to interest rates in a counterfactual market with constant rates, respectively, whereas cases 2 and 3 use intermediate values. Because these intermediate values seem more plausible to us than the other values, we focus on these cases more extensively below.

We note here that our model focuses on cross-household differences in borrowers' inaction—i.e., the demand side of the mortgage market. Our counterfactual scenarios attempt to capture a range of differences in the magnitudes of borrower cross-subsidies. Clearly, changes in the profile of interest rates affect lender profits and revenues as well, and their supply-side responses could constitute an important ingredient for further analysis.[27]

Table 5 reports the results of the counterfactual mortgage market outcomes for the different combinations of interest rates (in different panels) and estimated models (in different columns) as ratios of their respective baseline values (i.e., in the dual-rate economy). Odd-numbered columns correspond to version 1 of the model, and even-numbered columns correspond to version 2 of the model.

Perhaps not surprisingly, Panel A shows that all reported statistics increase in a market with a constant interest rate equal to the average discounted rate. Interest rates decline for all borrowers, except for those who always refinance, thereby boosting mortgage debt both at the extensive margin (increasing the total number of mortgages) and at the intensive margin (increases in average initial loan sizes and average loan balances). These outcomes are remarkably similar across the different estimated models, i.e., single versus multiple groups, with the slight difference that the multiple-group models display a higher sensitivity

[26] When working with multiple groups, we perform the cross-subsidy calculation using the interest rate that satisfies:

$$r_c \sum_{j=1}^{J} Q_j(r_c) = \sum_{j=1}^{J} \left(r(Q_{0,j}(r_d) + Q_{1,j}(r_d)) + RQ_{2,j}(R) \right), \tag{20}$$

where $Q_j(r_c)$ is the aggregate mortgage debt of group j when the interest rate is fixed at r_c. The difference between equations (19) and (20) is that aggregate revenues are calculated using the heterogeneous parameters across groups. In practice, the difference between the interest rates that satisfy equations (19) and (20) is only a few bps, with minimal effects on the counterfactuals reported in Table 5.

[27] Among others, Gurun, Matvos, and Seru (2016), Guiso, Pozzi, Tsoy, Gambacorta, and Mistrulli (2022), Benetton, Gavazza, and Surico (2021), Allen and Li (2020), and Thiel (2021) study supply-side incentives in mortgage markets.

Table 5: Market Outcomes with Constant Interest Rates

	UK-wide		Income Groups		Regions	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Constant Interest Rate=650 bps						
Number of Mortgages	1.10	1.11	1.15	1.10	1.12	1.10
Mean Initial Loan Amount	1.06	1.05	1.05	1.06	1.05	1.05
Standard Deviation Initial Loan Amount	1.05	1.05	1.01	1.02	1.04	1.04
Mean Loan Balance	1.05	1.04	1.05	1.05	1.05	1.05
Standard Deviation Loan Balance	1.06	1.05	1.02	1.03	1.04	1.04
Consumer Surplus	1.12	1.12	1.12	1.12	1.11	1.11
Panel B: Constant Interest Rate=683 bps						
Number of Mortgages	1.06	1.06	1.10	1.07	1.07	1.06
Mean Initial Loan Amount	0.98	0.97	0.96	0.97	0.97	0.97
Standard Deviation Initial Loan Amount	0.96	0.96	0.92	0.92	0.95	0.95
Mean Loan Balance	0.98	0.97	0.97	0.97	0.97	0.97
Standard Deviation Loan Balance	0.97	0.96	0.93	0.94	0.96	0.96
Consumer Surplus	1.04	1.04	1.03	1.03	1.03	1.03
Panel C: Constant Interest Rate=700 bps						
Number of Mortgages	1.04	1.03	1.07	1.05	1.04	1.04
Mean Initial Loan Amount	0.94	0.94	0.92	0.92	0.93	0.93
Standard Deviation Initial Loan Amount	0.91	0.91	0.87	0.88	0.90	0.90
Mean Loan Balance	0.94	0.94	0.93	0.93	0.94	0.93
Standard Deviation Loan Balance	0.92	0.92	0.89	0.89	0.92	0.91
Consumer Surplus	1.00	1.00	0.99	0.99	0.99	0.99
Panel D: Constant Interest Rate=759 bps						
Number of Mortgages	0.95	0.95	0.95	0.98	0.94	0.95
Mean Initial Loan Amount	0.83	0.82	0.80	0.79	0.82	0.81
Standard Deviation Initial Loan Amount	0.77	0.78	0.75	0.74	0.77	0.77
Mean Loan Balance	0.84	0.84	0.81	0.80	0.83	0.82
Standard Deviation Loan Balance	0.80	0.80	0.77	0.76	0.79	0.79
Consumer Surplus	0.88	0.88	0.87	0.86	0.88	0.87

Notes: This table reports the statistics on the mortgage market in counterfactual markets with constant interest rates, as ratios of those of the estimated market with dual interest rates. The statistics in Panel A are calculated using a constant interest rate equal to the average discounted rate. The statistics in Panel B are calculated using a constant interest rate equal to the average interest rate equal to (17). The statistics in Panel C are calculated using a constant interest rate equal to the interest rate that satisfies the equal-revenue equation (19). The statistics in Panel D are calculated using a constant interest rate equal to the average reset rate. Odd-numbered columns correspond to version 1 of the model, and even-numbered columns correspond to version 2 of the model.

of household participation decisions to interest rates relative to the UK-wide case, and thus a larger increase in the number of mortgages as interest rates decline.

Similarly, Panel D shows that all reported statistics decrease in a market with a constant interest rate equal to the average reset rate, because interest rates increase for all borrowers.

We note that the model estimated separately for each income group reported in columns (3)-(4) features a similar aggregate number of mortgages to the UK-wide case of columns (1)-(2), which in combination with the results in Panel A suggests that allowing for heterogeneity between groups pushes toward slightly greater sensitivity to downward rate movements than upward rate movements.

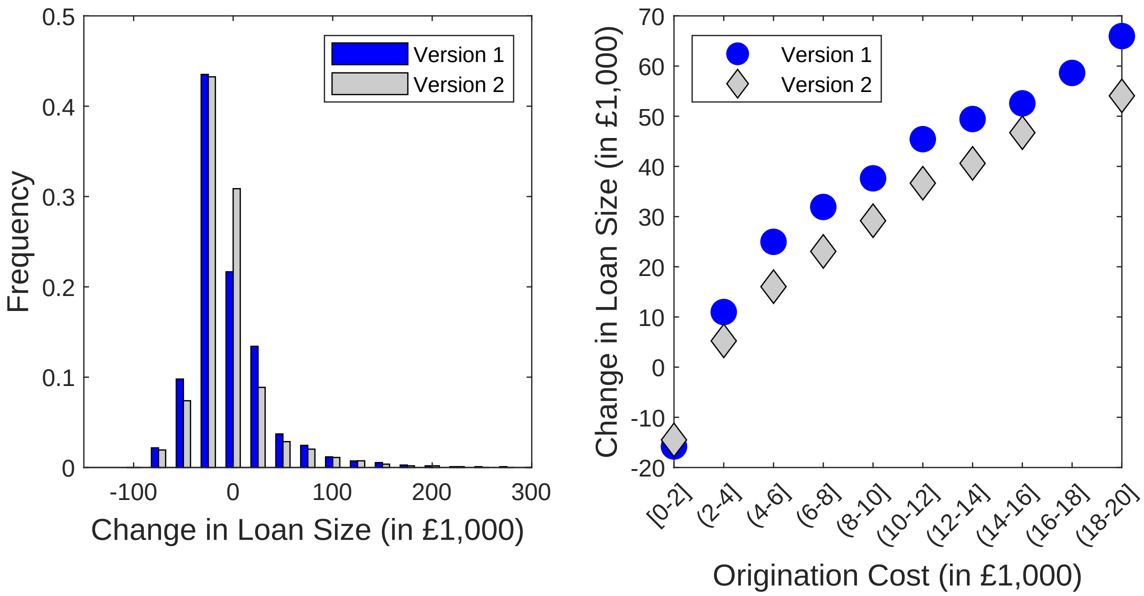
Panels B and C report several interesting outcomes which appear robust across different estimated models. Panels B and C are overall quite similar, we therefore focus our discussion on Panel B. We first describe the changes for the UK-wide case, and then for the multiple-group cases.

Panel B, UK-wide case. Panel B reports that the change in the profile of interest rates to a single-rate structure yields two main aggregate adjustments in opposite directions: the number of mortgages increases, but the average loan size decreases. Again, our discussion focuses on version 1 of the model in column (1), noting the key differences with version 2 in column (2).

More precisely, the first row of Panel B in Table 5 reports that the number of mortgages increases by 6.49 percent relative to the number of mortgages in the baseline economy. The reason for this increase is that there are many households with valuation v_i and with moderate or high costs $k_i^o = k_i$ just below the entry threshold $v_i^*(k_i^o)$ in the baseline economy who switch from renting a property in the dual-rate economy to taking a mortgage to buy a house in the counterfactual single-rate economy. These households would rarely refinance, and thus expect to pay an average rate close to the reset rate in the baseline dual-rate economy, which raises the costs of taking on a mortgage. These households, therefore, choose to rent in the dual-rate world, but since they pay a lower rate in the counterfactual single-rate economy, they choose to buy a house by taking on a mortgage in the counterfactual. The mass of these households, on net, is greater than the mass of households with low $k_i^o = k_i$ who pay an average rate close to the discounted rate in the baseline economy but pay a higher rate in the single-rate economy. Such low k_i^o households switch from owning with a mortgage in the dual-rate economy to renting a property in the counterfactual single-rate economy, but their exit from the mortgage market is more than offset by new entrants into the single-rate mortgage market.

The second row of Panel B shows that the average initial loan size decreases by 2.37

Figure 2: Change in Loan Size at Origination



Notes: The left panel displays the distribution of the changes in loan sizes at origination between the counterfactual economy with constant interest rates and the baseline economy with discounted and reset rates. The right panel displays the average change in loan sizes for households with different values of their origination costs (in bins of £2,000). All statistics displayed are computed including only households who either participate in the mortgage market in the baseline dual-rate economy, or in the counterfactual single-rate economy, or in both.

percent of the average loan size in the baseline case, corresponding to a mortgage size reduction of £4,653. The main reason for this decline is the change in the composition of borrowers: marginal households who enter the mortgage market in the single-rate economy have smaller loan sizes than inframarginal households whose participation does not change.

More generally, the change in the average loan size combines borrowers who increase their mortgage amounts with borrowers who decrease them. The left panel of Figure 2 shows the full distribution of the changes in mortgage amounts of those households who participate in the mortgage market in the baseline dual-rate economy, in the counterfactual single-rate economy, or both. The heterogeneity of the changes in mortgage amounts at origination is apparent, with decreases in mortgage amounts more concentrated than increases.

The right panel of Figure 2 helps to rationalize the asymmetric adjustment in loan sizes. It displays how the average change in mortgage size varies with the origination cost k_i^o of refinancing costs. Borrowers with the lowest $k_i^o = k_i$ pay an interest rate close to 650 bps in the estimated mortgage market, because they almost always refinance, but they pay 683 bps in the counterfactual market with a constant interest rate. This higher rate induces them to reduce their loan sizes. In contrast, borrowers with the highest k_i^o pay an interest rate close to 759 bps in the baseline market, because they never refinance, but pay 683 bps in the counterfactual market. As a result, these borrowers increase their loan sizes. The increases in loan sizes are more dispersed than the decreases in loan sizes, because there is a bigger difference between the rates in the dual- and single-rate worlds paid by households with high k_i^o than that between the interest rates in the economies paid by those with low k_i^o.

The right panel of Figure 2 shows that version 2 displays similar qualitative patterns to those of version 1. The key difference is that the magnitudes of all adjustments in version 2 are smaller than those in version 1. The reason is that households forecast their future interest rates when they choose their loan size at origination. The uncertainty over the level of refinancing costs is larger in version 2 than in version 1, and thus households' interest rate forecasts converge to an intermediate value between the discount and reset rate. These forecasts resemble the counterfactual interest rate that we consider, and thus households respond less in their intensive margin of adjustment in version 2 than in version 1.

The third row of Panel B in Table 5 reports that the standard deviation of initial loan sizes declines quite substantially, by 4.04 percent of the standard deviation of the initial loan size (corresponding to £4,978) in the estimated baseline model. The reason is that one dimension of household heterogeneity, namely k_i, contributes to the determination of the loan size in the baseline model with refinancing. However, this dimension of heterogeneity becomes irrelevant when interest rates are constant. More specifically, the previous arguments suggest—and Figure 2 shows—that borrowers with larger loans in the baseline economy decrease their loan sizes in the counterfactual, whereas borrowers with smaller loans in the baseline economy increase their loan sizes in the counterfactual with constant interest rates and no refinancing. A common interest rate thus pushes loan sizes to be more homogeneous.

The decline in initial loan size and the increase in the number of mortgages together combine to increase aggregate mortgage debt by 4.13 percent relative to the model with dual rates. Cross-subsidies are eliminated in the counterfactual, and one consequence of this change is that the mortgage market increases in size, although the effect is tempered by the opposing effects on the extensive and intensive margins.

The fourth and fifth rows of Panel B report that the patterns in the initial loan size distribution described above transfer to the aggregate loan balance distribution (i.e., including different cohorts of mortgages), with one additional subtle effect. In the baseline economy, on average, borrowers who originate large loans pay lower rates than borrowers with small loans. Hence, as loans amortize over time, the loan balances of borrowers with large loans tend to decline at a faster rate than the loan balances of borrowers with small loans, which compresses the distribution of loan balances over time. This force is absent in the counterfactual single-rate economy as all borrowers pay the same rate. Hence, the standard deviation of loan balances (normalized by that observed in the baseline economy) reported in the last row is slightly larger than the standard deviation of initial loan balances (also normalized with respect to the baseline economy) reported in the third row.

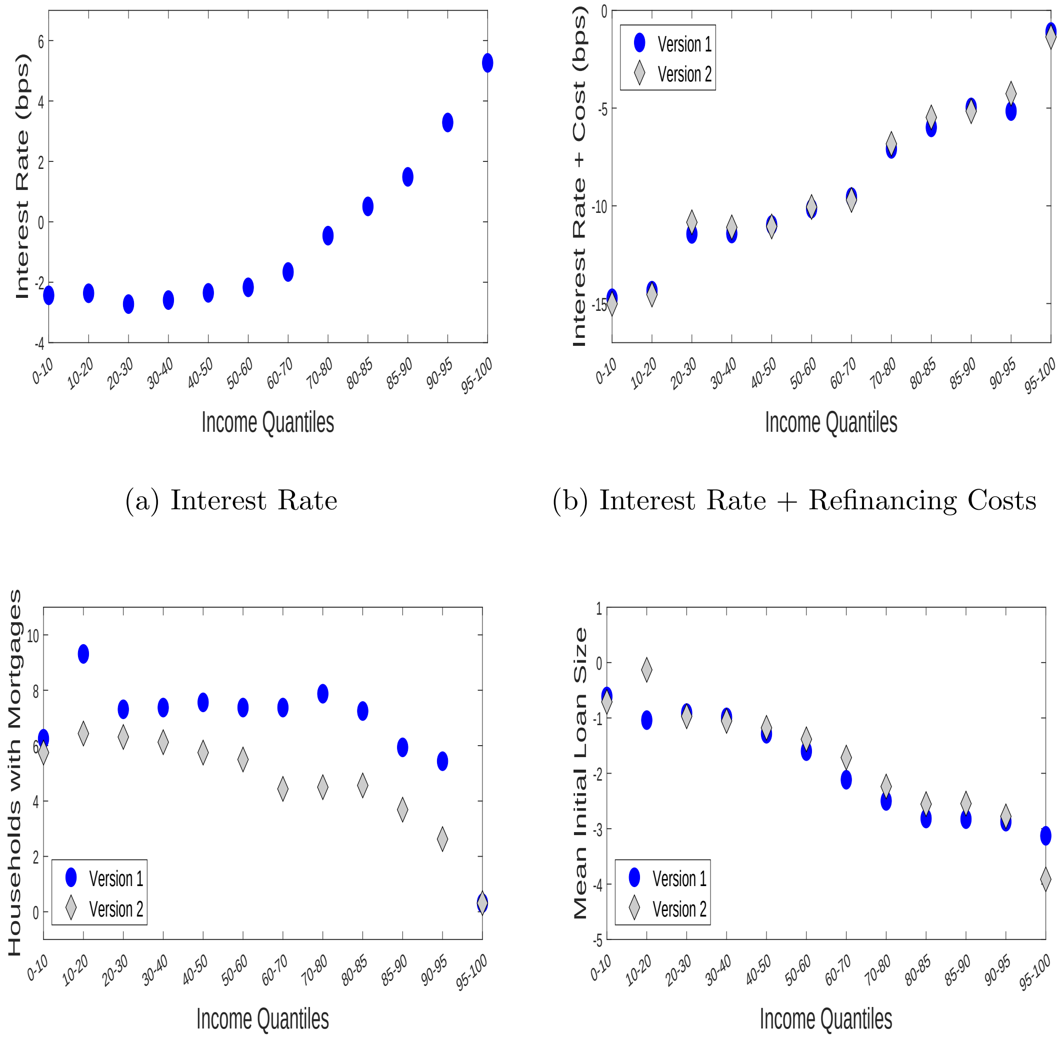
Finally, the last row of Panel B summarizes all the changes in a single money-metric ex-ante measure of consumer surplus, calculated for each household as $\max\left(W_0(v_i, k_i^o), \frac{\bar{u}}{1-\beta}\right)$. Consumer surplus increases by 3.94 percent in the single-rate economy relative to the dual-rate economy.

Panel B, Income Groups. The cases with multiple groups allow us to explain some of the observable heterogeneity in refinancing rates across income groups and geographies of the UK with heterogeneity in preferences v_i and costs k_i. These richer cases help us to evaluate whether and how the shift to a single mortgage rate structure leads to different outcomes for households in these groups.

Columns (3) and (4) of Table 5 report aggregated counterfactual estimates when the model is estimated using moments for different income groups. When we compare these aggregate statistics with those of the UK-wide model in columns (1) and (2), the differences appear small. The main difference is that column (3) exhibits a slightly larger adjustment in the extensive margin (i.e., the number of mortgages) than column (1), whereas the differences between columns (4) and (2) are minor. We now analyze how the results differ across income groups.

Figure 3 plots selected changes to mortgage market outcomes for each income group. The top-left panel shows that interest rates (in bps) are lower in the counterfactual economy for income groups up to roughly the 80th percentile of the income distribution in the sample, and are higher for the very highest income groups. This pattern is consistent with the

Figure 3: Changes in Market Outcomes by Income Groups



(a) Interest Rate

(b) Interest Rate + Refinancing Costs

(c) Households with Mortgages

(d) Initial Loan Size

Notes: The top-left panel displays the change in interest rates (in bps); the top-right panel displays the change in mortgage costs, calculated as interest rates net of refinancing costs $k_{i,t}$; the bottom-left panel displays the percentage-point change in the share of households with mortgages; and the bottom-right panel displays the percentage change in the average initial loan size for each income group in the counterfactual case with a constant interest rate equal to $r_c = 683$ bps relative to the baseline case. Dark dots correspond to the model of column (3) and light diamonds correspond to the model of column (4) in Panel B of Table 5 regressive nature of the cross-subsidies in the dual-rate economy. The highest income group pays higher interest rates in the single-rate economy than the average rates they pay in the dual-rate economy. This is primarily because high-income households have larger loans, which gives them greater incentives to refinance promptly in the dual-rate economy.

The top-right panel adds refinancing costs to interest rates to calculate an all-inclusive mortgage cost. The pattern across groups is broadly similar to that seen in interest rates, with lower-income households paying lower mortgage costs under the counterfactual single-rate economy, whereas higher-income households pay mortgage costs similar to the costs that they pay in the baseline dual-rate economy.

The bottom-left panel shows that these changes in interest rates translate into an aggregate increase in the share of households with mortgages. Critically, the percentage-point increase is larger for lower-income groups, and minimal for the highest-income groups. Table 2 reports that the homeownership rate among low-income households in the baseline dual-rate economy is low (it equals 50%) and rises steeply with income, and our model suggests that the design of the mortgage market may be a contributing factor to these patterns. In the single-rate market, there is a substantially greater entry of these low-income households into the housing and mortgage markets. As expected, version 2 displays smaller adjustments than version 1 for all income groups, because, when originating their mortgages in the baseline dual-rate economy, households can tailor their loan amounts to their future costs less precisely in version 2 than in version 1, because of their noisier information about their future costs. However, the qualitative pattern of the adjustments across income groups appears robust to the differential household information in the two versions.

The bottom-right panel plots the average percent differences between initial loan sizes in the single-rate economy and those in the dual-rate economy. While there are also important changes within groups, the across-group comparison highlights that higher-income groups adjust their average initial loan size downward more than lower-income groups. The adjustment in the average initial loan size of the highest-income group is sizable, because many of these borrowers—i.e., a larger fraction than among lower-income groups—almost always refinance and thus suffer a substantial increase in the interest rate that they pay, from $r_d = 650$ to $r_c = 683$ bps. These loan size adjustments across income groups are very similar in the two versions of the model.

Overall, these panels suggest that the richer model with greater household heterogeneity across the income distribution implies that higher-income households pay lower rates and lower all-in mortgage costs than lower-income households in the current dual-rate structure. These patterns are consistent with the idea that the dual-rate structure fosters regressive cross-subsidies. The bottom panels of Figure 3 suggest that different income groups would respond to a single-rate structure with different types and levels of adjustments on both the intensive and extensive margins. In particular, raising participation in the mortgage market is the main adjustment for lower-income groups, whereas lowering initial loan sizes is the main adjustment for higher-income groups.

Finally, the changes in consumer surplus confirm that all income groups would benefit in the single-rate economy relative to the dual-rate economy, because they either pay lower interest rates or they save the refinancing costs $k_{i,t}$. Lower-income groups attain larger percentage-change increases in consumer surplus than higher-income groups.

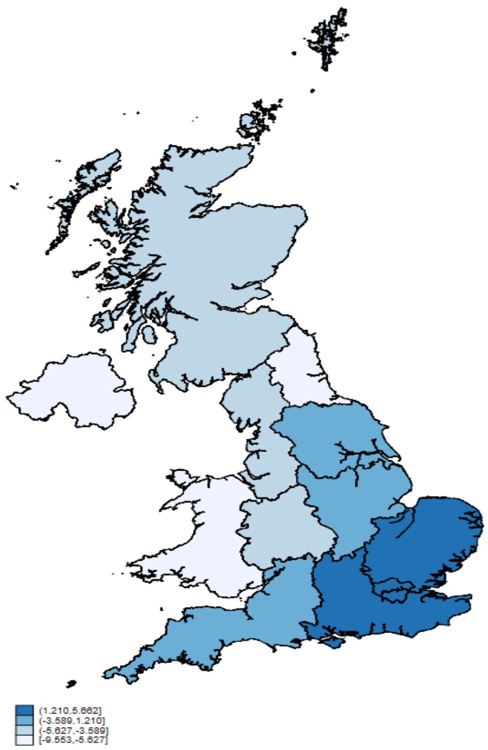
Panel B, Regions. Columns (5) and (6) of Table 5 report aggregated counterfactual estimates when the model is estimated with parameters and moments for different UK regions. Once again, as with the model which incorporates greater heterogeneity across income groups, the aggregate statistics reported in columns (5) and (6) are remarkably similar to those of the UK-wide estimation reported in columns (1) and (2).

Figures 4 and 5 present maps that display some of the changes to mortgage market outcomes across different UK regions, for the models with more- and less-precise information about the persistent component of future refinancing costs at origination (i.e., versions 1 and 2), respectively. In each panel, darker colors indicate larger (positive) changes in the counterfactual market with constant interest rates when compared with the baseline dual-rate economy with discounted and reset rates.

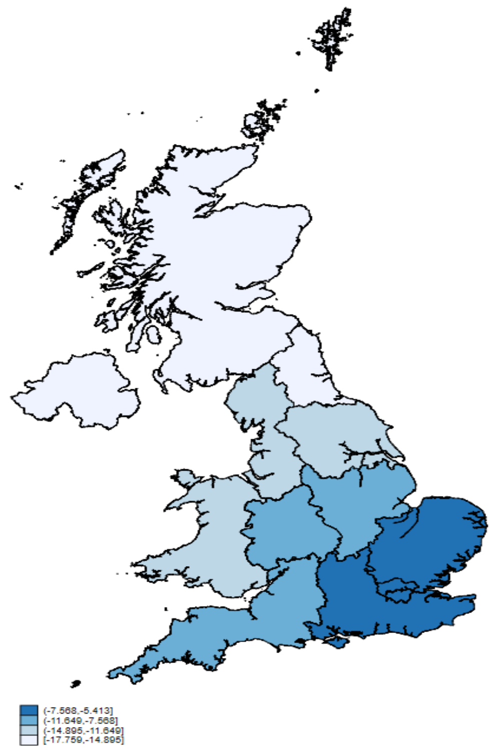
The top-left map displays the change in average interest rates paid on mortgages, reported in bps. Households in the more prosperous regions of Greater London, the South East of England, and the East of England experience the largest increases in mortgage rates, whereas households in relatively less well-off regions and devolved administrations such as Northern Ireland, Wales, and the North East of England would experience the largest decreases in rates when moving to a single rate. These regional patterns are consistent with the dual-rate structure featuring regressive cross-subsidies across UK regions.

The top-right plot displays all-inclusive mortgage costs that sum (paid) refinancing costs

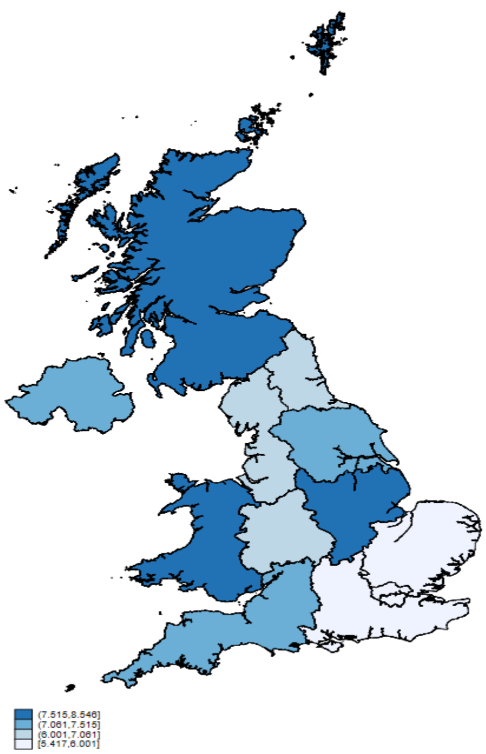
Figure 4: Regional Changes, Version 1



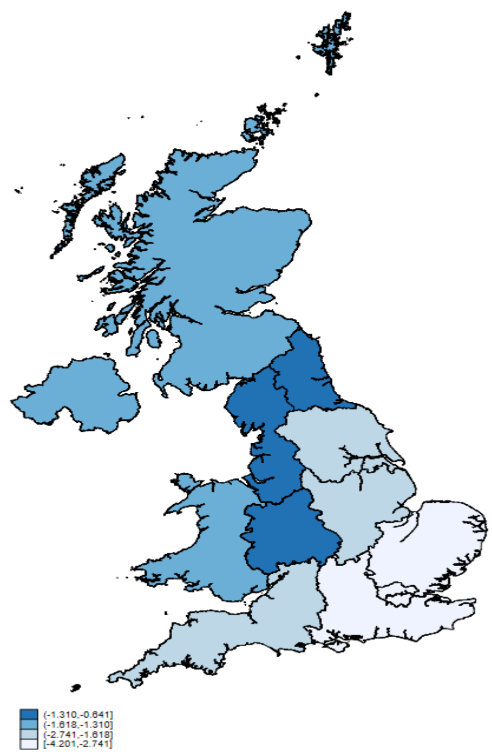
(a) Interest Rate



(b) Interest Rate + Refinancing Costs

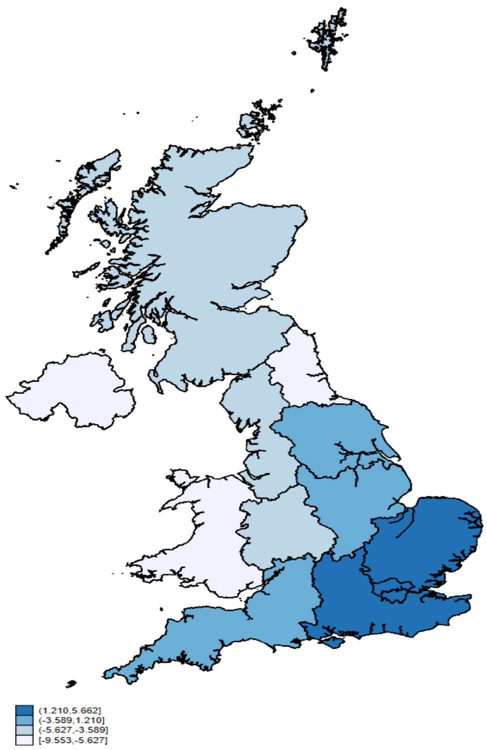


(c) Number of Mortgages

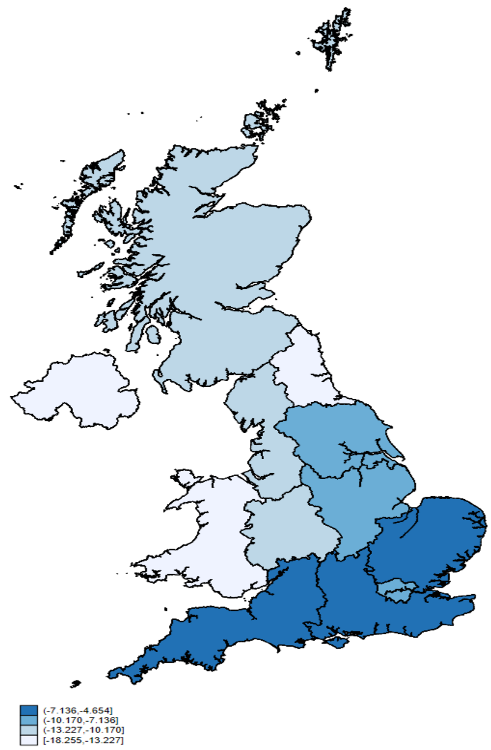


(d) Initial Loan Size

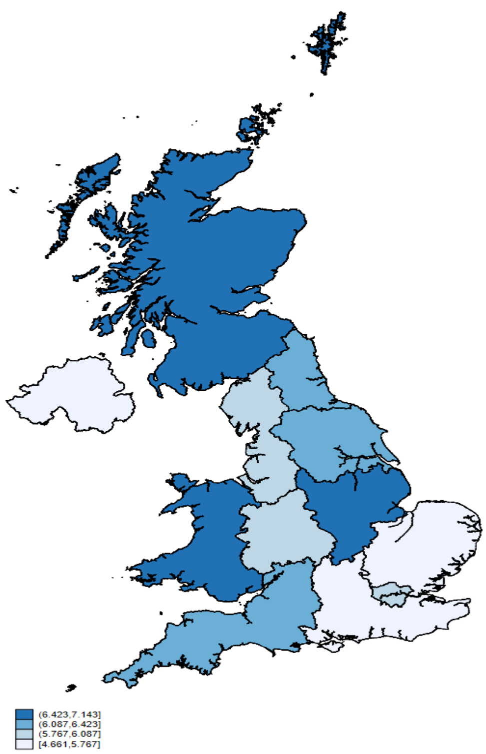
Figure 5: Regional Changes, Version 2



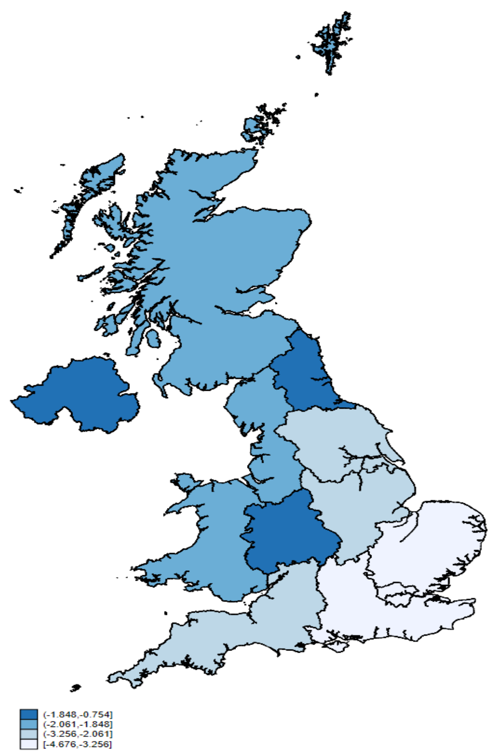
(a) Interest Rate



(b) Interest Rate + Refinancing Costs



(c) Number of Mortgages



(d) Initial Loan Size

and interest rates. The geographic patterns once again point to the regressive patterns found for interest rates in the top-left panel, with Scotland, Northern Ireland, and North East of England paying lower mortgage costs, whereas southern regions of the UK pay higher mortgage costs, in the counterfactual with a single rate.

In the counterfactual equilibrium, as seen earlier in the case of income groups, households endogenously adjust their mortgage market participation as well as their mortgage amounts. The bottom-left plot shows the change in the number of mortgages, which broadly increases the most in regions and devolved administrations that experience the largest decrease in mortgage rates and costs, such as Scotland and Northern Ireland. In contrast, southern regions of the UK experience smaller adjustments to mortgage market participation.

The bottom-right map displays the changes in the average initial loan size rate across regions. The differences in these averages when moving to the single-rate world mask larger within-region changes. That said, once again, southern regions' average initial loan sizes do shrink considerably more than less well-off regions' average initial loan sizes in the counterfactual single-rate world. Overall, the bottom maps confirm the pattern that the change in the profile of interest rates affects mostly the extensive margin in lower-income regions, and mostly the intensive margin in higher-income regions.

Finally, all regions would enjoy higher consumer surplus in the single-rate economy relative to the dual-rate economy, because they either pay lower interest rates or they save the refinancing costs $k_{i,t}$, with lower-income regions experiencing higher percentage-change increases in consumer surplus than higher-income regions.

6 Conclusion

We develop a model of mortgage refinancing and structurally estimate it on rich data from the UK mortgage market. Our model matches broad features of the data, and the parameters reveal considerable heterogeneity in mortgage refinancing costs across households, echoing findings in prior literature. We use the estimated parameters to uncover regressive cross-subsidies in this market by conducting a counterfactual comparison with an alternative mortgage contract that features a constant interest rate and no need for refinancing.

This approach allows us to quantify cross-subsidies in this market setting. Using 2015 data, we set annual interest rates in our main counterfactual single-rate equilibrium to lie approximately 25bps above the average discounted rate and 30bps below the average reset rate that borrowers are routinely rolled on to at the expiration of the discounted rate fixation period. These are material changes given the importance of mortgages to household budgets.

The counterfactual scenario features different adjustments by low- and high-income groups. Low-income households enter the mortgage market in greater numbers, and raise their loan balances in response to the lower interest rates that they pay on average and the elimination of refinancing costs. Essentially, low-income groups are penalized by the dual-rate structure because they have smaller loan balances. Hence, they are more likely to pay the high reset rate in the dual-rate economy compared with high-income households. In contrast, high-income households mainly take on smaller loans in the single-rate counterfactual economy in response to their inability to take advantage of the discounted rate.

These findings highlight an important and novel dimension of inequality that would be invisible without our structural approach: we find that changes in mortgage rates increase entry into the housing and mortgage markets for low-income households; they also tend to push loan sizes to be more uniform across high- and low-income households. The economic size of these responses is substantial, even when we conservatively assume that households only observe noisy signals about their ongoing refinancing costs. Our results suggest that simplifying the design of mortgage refinancing and eliminating the costs associated with refinancing can cause forward-looking (even if imperfectly informed) households to participate more extensively in the mortgage market.

Our work has both methodological and economic contributions beyond the specific context that we study. First, we believe that our structural approach to estimating financial cross-subsidies by comparing the current and counterfactual market structures is a useful way to provide a money-metric assessment of the impacts of heterogeneity in household inaction. This has potentially wider implications for the field of household finance, where such heterogeneity is widely prevalent in many markets including credit and insurance. Our findings on the regressive nature of these cross-subsidies highlight that other household finance settings where high-income households benefit more due to their larger stakes and their greater propensity to take action may also contribute to inequality. In a broader sense, our results on the distribution of financial cross-subsidies in this important market show that studying household finance is helpful for the agenda of identifying the sources and consequences of wealth inequality, a continuing concern for society.

References

Agarwal, S., J. C. Driscoll, and D. I. Laibson (2013): "Optimal Mortgage Refinancing: a Closed-form Solution," *Journal of Money, Credit and Banking*, 45(4), 591–622.

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

Allen, J., and S. Li (2020): "Dynamic Competition in Negotiated Price Markets," Staff Working Paper 2020-22, Bank of Canada.

Alvaredo, F., L. Chancel, T. Piketty, E. Saez, and G. Zucman (2017): "Global Inequality Dynamics: New Findings from WID.world," *American Economic Review*, 107(5), 404–409.

Andersen, S., J. Y. Campbell, K. M. Nielsen, and T. Ramadorai (2020): "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market," *American Economic Review*, 110(10), 3184–3230.

Armstrong, M., and J. Vickers (2012): "Consumer Protection and Contingent Charges," *Journal of Economic Literature*, 50(2), 477–493.

Ausubel, L. M. (1991): "The Failure of Competition in the Credit Card Market," *American Economic Review*, 81(1), 50–81.

Badarinza, C., J. Y. Campbell, and T. Ramadorai (2016): "International Comparative Household Finance," *Annual Review of Economics*, 8, 111–144.

——— (2018): "What Calls to ARMs? International Evidence on Interest Rates and the Choice of Adjustable-rate Mortgages," *Management Science*, 64(5), 2275–2288.

Belgibayeva, A., T. Bono, P. Bracke, J. Cocco, and T. Majer (2020): "When Discounted Rates End: The Cost of a Taking Action in the Mortgage Market," Occasional Paper 54, Financial Conduct Authority.

Benetton, M. (2021): "Leverage Regulation and Market Structure: An Empirical Model of the UK Mortgage Market," *Journal of Finance*, 76(6), 2297–3053.

Benetton, M., A. Gavazza, and P. Surico (2021): "Mortgage Pricing and Monetary Policy," Discussion Paper 16456, Centre for Economic Policy Research.

Benhabib, J., and A. Bisin (2018): "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 56(4), 1261–1291.

Beraja, M., A. Fuster, E. Hurst, and J. Vavra (2019): "Regional Heterogeneity and the Refinancing Channel of Monetary Policy," *The Quarterly Journal of Economics*, 134(1), 109–183.

Berger, D., K. Milbradt, F. Tourre, and J. Vavra (2021): "Mortgage Prepayment and Path-Dependent Effects of Monetary Policy," *American Economic Review*, 111(9), 2829–2878.

Berger, D., K. Milbradt, F. Tourre, and J. Vavra (2023): "Refinancing Frictions, Mortgage Pricing and Redistribution," Working paper, Northwestern University.

Best, M. C., J. S. Cloyne, E. Ilzetzki, and H. J. Kleven (2020): "Estimating the Elasticity of Intertemporal Substitution Using Mortgage Notches," *The Review of Economic Studies*, 87(2), 656–690.

Brunnermeier, M. K., and S. Nagel (2008): "Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-evidence on Individuals," *American Economic Review*, 98(3), 713–736.

Byrne, S., K. Devine, M. King, Y. McCarthy, and C. Palmer (2023): "The Last Mile of Monetary Policy: Inattention, Reminders, and the Refinancing Channel," Working Paper 31043, National Bureau of Economic Research.

Campbell, J. Y. (2006): "Household Finance," *The Journal of Finance*, 61(4), 1553–1604.

——— (2013): "Mortgage Market Design," *Review of Finance*, 17(1), 1–33.

Campbell, J. Y., T. Ramadorai, and B. Ranish (2019): "Do the Rich Get Richer in the Stock Market? Evidence from India," *American Economic Review: Insights*, 1(2), 225–240.

Chen, A. H., and D. C. Ling (1989): "Optimal mortgage refinancing with stochastic interest rates," *Real Estate Economics*, 17(3), 278–299.

Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2002): "Defined Contribution Pensions: Plan Rules, Participant Choices, and the Path of Least Resistance," *Tax Policy and the Economy*, 16, 67–113.

Cloyne, J., K. Huber, E. Ilzetzki, and H. Kleven (2019): "The Effect of House Prices on Household Borrowing: A New Approach," *American Economic Review*, 109(6), 2104–2136.

DeFusco, A. A., and J. Mondragon (2020): "No Job, No Money, No Refi: Frictions to Refinancing in a Recession," *The Journal of Finance*, 75(5), 2327–2376.

DeFusco, A. A., and A. Paciorek (2017): "The Interest Rate Elasticity of Mortgage

Demand: Evidence from Bunching at the Conforming Loan Limit," *American Economic Journal: Economic Policy*, 9(1), 210–240.

DellaVigna, S., and U. Malmendier (2006): "Paying not to Go to the Gym," *American Economic Review*, 96(3), 694–719.

Ellison, G. (2005): "A Model of Add-on Pricing," *The Quarterly Journal of Economics*, 120(2), 585–637.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2020): "Heterogeneity and Persistence in Returns to Wealth," *Econometrica*, 88(1), 115–170.

Farrell, J., and P. Klemperer (2007): "Coordination and Lock-in: Competition with Switching Costs and Network Effects," in *Handbook of Industrial Organization*, ed. by M. Armstrong, and R. Porter, vol. 3, pp. 1967–2072. Elsevier.

Financial Conduct Authority (2019a): "Changes to mortgage responsible lending rules and guidance—feedback on CP19/14 and final rules," Policy Statement PS 19/27, Financial Conduct Authority, Accessed: 2021-10-19.

——— (2019b): "Mortgage Market Study: Final Report," Market Study MS 16/2.3, Financial Conduct Authority.

——— (2020a): "FCA issues research on mortgage switching," https://www.fca.org.uk/news/news-stories/fca-issues-research-mortgage-switching, Accessed: 2021-10-19.

——— (2020b): "Removing barriers to intra-group switching and helping borrowers with maturing interest-only and part-and-part mortgages," Policy Statement PS 20/11, Financial Conduct Authority, Accessed: 2021-10-19.

——— (2021): "Mortgage Prisoners Review," Policy Paper CP 576, Financial Conduct Authority.

Gabaix, X., and D. Laibson (2006): "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *The Quarterly Journal of Economics*, 121(2), 505–540.

Gavazza, A., and A. Lizzeri (2021): "Frictions in Product Markets," in *Handbook of Industrial Organization*, ed. by K. Ho, A. Hortaçsu, and A. Lizzeri, vol. 4, pp. 433–484. Elsevier.

Goetzmann, W. N., C. Spaenjers, and S. Van Nieuwerburgh (2021): "Real and Private-Value Assets," *The Review of Financial Studies*, 34(8), 3497–3526.

Greenwald, D. L., M. Leombroni, H. Lustig, and S. Van Nieuwerburgh (2021): "Financial and Total Wealth Inequality with Declining Interest Rates," Working Paper 28613, National Bureau of Economic Research.

Guiso, L., A. Pozzi, A. Tsoy, L. Gambacorta, and P. E. Mistrulli (2022): "The Cost of Steering in Financial Markets: Evidence from the Mortgage Market," *Journal of Financial Economics*, 143(3), 1209–1226.

Gurun, U. G., G. Matvos, and A. Seru (2016): "Advertising Expensive Mortgages," *The Journal of Finance*, 71(5), 2371–2416.

Handel, B. R. (2013): "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *American Economic Review*, 103(7), 2643–82.

Heidhues, P., and B. Kőszegi (2018): "Behavioral Industrial Organization," in *Handbook of Behavioral Economics - Foundations and Applications 1*, ed. by B. D. Bernheim, S. D. Vigna, and D. Laibson, vol. 1, pp. 517–612. Elsevier.

Honka, E. (2014): "Quantifying Search and Switching Costs in the US Auto Insurance Industry," *The RAND Journal of Economics*, 45(4), 847–884.

Hubmer, J., P. Krusell, and A. A. Smith Jr (2020): "Sources of US Wealth Inequality: Past, Present, and Future," in *NBER Macroeconomics Annual 2020, volume 35*. University of Chicago Press.

Hurst, E., B. J. Keys, A. Seru, and J. Vavra (2016): "Regional Redistribution through the US Mortgage Market," *American Economic Review*, 106(10), 2982–3028.

Illanes, G. (2016): "Switching Costs in Pension Plan Choice," Working paper, Northwestern University.

Keys, B. J., D. G. Pope, and J. C. Pope (2016): "Failure to Refinance," *Journal of Financial Economics*, 122(3), 482–499.

Lea, M. (2010): "International Comparison of Mortgage Product Offerings," Discussion paper, Research Institute for Housing America.

Liu, L. (2022): "The Demand for Long-Term Mortgage Contracts and the Role of Collateral," Working paper, Imperial College London.

Luco, F. (2019): "Switching Costs and Competition in Retirement Investment," *American Economic Journal: Microeconomics*, 11(2), 26–54.

Miles, D. (2004): "The UK Mortgage Market: Taking a Longer-Term View," Working paper, Imperial College London.

Nelson, S. (2022): "Private Information and Price Regulation in the US Credit Card Market," Working paper, University of Chicago.

Piskorski, T., and A. Seru (2018): "Mortgage market design: Lessons from the Great Recession," *Brookings Papers on Economic Activity*, 2018(1), 429–513.

Robles-Garcia, C. (2022): "Competition and Incentives in Mortgage Markets: The Role of Brokers," Working paper, Stanford University.

Scharfstein, D., and A. Sunderam (2016): "Market power in mortgage lending and the transmission of monetary policy," Working paper, Harvard University.

Stango, V., and J. Zinman (2016): "Borrowing High versus Borrowing Higher: Price Dispersion and Shopping Behavior in the US Credit Card Market," *Review of Financial Studies*, 29(4), 979–1006.

Stanton, R. (1995): "Rational Prepayment and the Valuation of Mortgage-Backed Securities," *The Review of Financial Studies*, 8(3), 677–708.

Stanton, R., and N. Wallace (1998): "Mortgage Choice: What's the Point?," *Real Estate Economics*, 26(2), 173–205.

Thiel, J. (2021): "Price Discrimination, Switching Costs and Welfare: Evidence from the Dutch Mortgage Market," Working paper, Copenaghen Business School.

Zhang, D. (2022): "Closing Costs, Refinancing, and Inefficiencies in the Mortgage Market," Working paper, Harvard University.

Online Appendix for "Refinancing Cross-Subsidies in the Mortgage Market"

A Dataset Construction

A.1 Primary Data Source: Stock of Owner-Occupier Mortgages

PSD007 includes loan-level information on the universe of mortgages in the owner-occupier or residential segment of the mortgage market. The owner-occupier segment includes first-time-buyers, home-movers, and refinancers who obtain mortgages from regulated financial institutions, such as deposit-taking lenders and building societies. All regulated financial institutions are mandated by law to share this data with the FCA at a semi-annual frequency.

We have data on 6 PSD007 snapshots, reported half-yearly from mid-2015 to end-2017. Table A.1 provides a description of the loan-level variables reported in PSD007 relevant to our study. In each snapshot, we observe the loan balance, original size of the loan, remaining term to maturity, original maturity, and interest rate for each mortgage recorded on the reporting date. The database also includes information on the type of interest rate and whether the mortgage is incentivized (i.e., on a discounted rate), and if so, the remaining period under the incentivized or discounted rate. The types of interest rates reported in the dataset are teaser, discounted, capped, standard variable rate, tracker, and an unclassified other category. In some of our summary statistics, we use the reported interest rate to calculate a spread of the discounted rate over the yield on a nominal zero coupon UK Treasury maturing over the horizon over which the interest rate is fixed.[A.1]

Table A.2 shows the overall balance of mortgages in 2015H1 by interest rate type and incentivized status. The table shows that a vast majority of the mortgages reported as being incentivized are also reported to be under teaser rates. Most mortgages under discounted and capped interest rates are also reported as being incentivized. However, there are a few discounted and capped mortgages that are reported as being non-incentivized and appear

[A.1] Discounted mortgages are fixed-term mortgages with a specified period under the discounted rate. In our model and in the data, mortgages automatically switch to the reset rate at the end of the discounted period. For example, in the case of a mortgage with a year remaining on the discounted rate, the spread is calculated over the yield on a nominal zero coupon UK Treasury bill maturing in a year. Reset rate mortgages are variable rate mortgages; the spread for reset rate mortgages is calculated based on the yield on short-term (6 month) UK Treasury bill.

Table A.1: Description of Variables

Variable	Description
Current Loan Balance	Balance as on the date of reporting
Current Interest Rate	Interest rate charged on the mortgage
Spread	Spread over the yield on a nominal zero coupon bond maturing over a horizon comparable to the fixation period for interest rates (0 for mortgages under reset rate).
Original Loan Balance	Original size at the time of mortgage account opening date.
Original Term	Original term to maturity at the time of mort. account opening date.
Remaining Term	Remaining term to maturity.
Remaining Discounted Period	Remaining period under discounted rates.
Borrower Age	Borrower age as on the date of reporting.

Notes: The table above provides a brief description of mortgage level variables reported in PSD007 data relevant to our study.

to have anomalous interest rates (we explain further below). We exclude such mortgages from our sample, and pool all incentivized mortgages reported as teaser, discounted, and capped interest rates into our discounted category in the paper.

Table A.3 shows the average interest rate by interest rate type and incentivized status in the 2015H1 snapshot. Mortgages that we classify as discounted, i.e., incentivized mortgages on teaser, discounted, and capped interest rates have lower average interest rates. Mortgages on reset rates (or Standard Variable Rates, SVRs) have higher average interest rates than these categories. There is a small group of mortgages on reset rates which are also reported as being incentivized, which bear interest rates comparable to that of the non-incentivized reset rate mortgages.[A.2] We treat all instances of mortgages on reset rates as non-incentivized.

Tracker mortgages are the remaining large category of mortgages. Their interest rates are benchmarked to the contemporaneous Bank of England base rate or LIBOR. Table A.3 shows that the average interest rate of mortgages in this category are lower than other mortgage types. However, this category is distinct from the discounted rate mortgages, as these are not teaser rates fixed for a duration; they are subject to rate fluctuations, and there are rarely transitions from the reset and discounted rate category into this category. As the tracker category is relatively isolated from the other two categories and outside of our model, we restrict our study on cross-subsidies to mortgages under the discounted and reset rate categories.[A.3]

[A.2] This is a data issue only in the 2015H1 snapshot.

[A.3] The total number of mortgages, average interest rates, and outstanding balances reported in Tables

Table A.2: Mortgages in 2015H1: Total Balance by Interest Rate Type and Incentivized Status (in £ billions)

	Incentivized		
	No	Yes	Total
Teaser	11.4	442.6	454.0
Discount	1.3	7.3	8.7
Capped	0.0	0.5	0.5
SVR	208.5	6.1	214.6
Tracker	121.7	90.6	212.3
Other	39.0	0.1	39.1
Total	381.9	547.3	929.2

Notes: The table above shows the total balance in £ billions by type of interest rate, and whether the mortgage is reported as being incentivized in the mortgage snapshot for 2015H1.

Table A.3: Mortgages in 2015H1: Average Interest Rate by Interest Rate Type and Incentivized Status

	Incentivized		
	No	Yes	Total
Teaser	5.83	3.35	3.48
Discount	3.04	3.31	3.26
Capped	4.02	2.91	2.99
SVR	3.79	3.63	3.79
Tracker	2.22	2.16	2.19
Other	2.88	2.80	2.88
Total	3.39	3.15	3.26

Notes: The table above shows the average interest rate by type of interest rate, and whether the mortgage is reported as being incentivized in the mortgage snapshot for 2015H1.

A.2 Borrower Incomes and House Prices

The PSD007 dataset does not include information on current borrower incomes, which are typically reported at mortgage origination. We obtain information at origination from the PSD001 dataset (a dataset similar to PSD007, but from earlier years). We use the same variable used to merge information across stock snapshots (since it uniquely identifies a mortgage) to merge the stock data with the loan origination data. We obtain the latest income reported to the lender at the time of origination (usually the first instance of the

A.2-A.3 are before the data filtering and cleaning steps that we describe in this appendix and in the paper.

mortgage being issued, occasionally captured in a subsequent refinancing round), and scale it using local-area level income indices obtained from the Office of National Statistics to an estimate for 2015H1.

Importantly, the distribution of the year of the reported income recorded at the time of origination does not vary across regions or across income bins. This helps to validate the quality of our loan-level income data, and provides reassurance that our cross-subsidy estimate by income-quantiles in section 5 is not affected by any differential quality of reported income across UK regions.

Similar to income, house prices are typically appraised at origination. In order to compute current house prices, and in particular to obtain current loan-to-value ratios (LTV), we scale house prices observed at origination using local-area-level house price indices, reported by HM Land Registry. These house price indices are available at monthly frequency. In order to match the reporting frequency of the mortgage stock (PSD007), we use house price indices reported in June (for H1), and December (for H2 data). This approach is standard in the literature, and is consistent with lenders' own adjustments of loan-to-value ratios when households refinance.

A.3 Data Cleaning

In the preceding section, we discussed filtering out mortgages with anomalous interest rate types, tracker mortgages, and mortgages under an unspecified "other" category. We implement further data cleaning steps to filter out observations with anomalous or inconsistent data on remaining discounted period, balance, interest rate, remaining term and borrower age, as well as borrowers who may not be able to refinance because they are underwater, in arrears, highly leveraged or have a loan balance that is too small.

Reported Remaining Discounted Period. Table A.4 shows summary statistics for the remaining period on discounted rates (in years) for discounted mortgages across the six snapshots. The mean and standard deviation of the remaining discounted period is consistent across the snapshots, except 2015H1. This is driven primarily by the misclassification of reset rate mortgages as discounted mortgages in 2015H1, resulting in a mass of mortgages with remaining discounted periods greater than 10 years. These are cases where the remaining term is reported as the remaining discounted period, something not seen in other snapshots.

Table A.4: Remaining Discounted Period in Years

(a) Raw database

	mean	sd	p10	p25	p50	p75	p90
Rem. discounted period (2015H1)	2.89	5.37	0.42	1.00	1.83	3.33	4.58
Rem. discounted period (2015H2)	2.21	2.27	0.33	0.92	1.75	3.00	4.42
Rem. discounted period (2016H1)	2.15	2.23	0.33	1.00	1.75	2.92	4.33
Rem. discounted period (2016H2)	2.09	2.20	0.42	0.92	1.67	2.83	4.25
Rem. discounted period (2017H1)	2.04	2.19	0.33	0.83	1.50	2.75	4.33
Rem. discounted period (2017H2)	2.11	2.21	0.33	0.83	1.58	3.00	4.50

(b) Filtered database

	mean	sd	p10	p25	p50	p75	p90
Rem. discounted period (2015H1)	2.11	1.52	0.42	1.00	1.83	3.08	4.25
Rem. discounted period (2015H2)	2.09	1.53	0.33	0.92	1.75	3.00	4.33
Rem. discounted period (2016H1)	2.06	1.51	0.42	1.00	1.75	2.92	4.25
Rem. discounted period (2016H2)	2.01	1.52	0.42	0.92	1.67	2.75	4.25
Rem. discounted period (2017H1)	1.96	1.57	0.33	0.83	1.50	2.75	4.25
Rem. discounted period (2017H2)	2.03	1.65	0.33	0.83	1.58	2.92	4.50

Notes: The above tables show summary statistics for the remaining discounted period in months for discounted mortgages across the PSD007 snapshots. Panel (a) shows the summary statistics for the raw database; panel (b) shows the summary statistics after the filtering steps described in Section A.3.

We reclassify the reset rate (or SVR) mortgages with misreported remaining discounted period as being not incentivized.

In addition, across all snapshots, there are few mortgages with remaining discounted periods less than -1 years, and greater than 11 years. We drop all such observations from the sample. Table A.4 (b) shows that the distribution of the remaining discounted period is similar across the different snapshots after implementing the filtering steps described above.

Reported Balance. Table A.5 reports summary statistics for loan balances across snapshots. It shows that all the moments (including mean and standard deviation) for loan balances in 2017H2 are higher than that for other snapshots. This difference is driven by discounted rate mortgages; the loan balance moments for reset rate mortgages are stable across the snapshots.

We find that the high mean and standard deviation for discounted mortgages in 2017H2 is driven by misreported loan balances for two lenders. Hence, for the discounted mortgages issued by these two lenders in 2017H2, we replace the reported loan balance in 2017H2 with

Table A.5: Loan Balances for Raw and Filtered Databases

(a) Raw database

	mean	sd	p10	p25	p50	p75	p90
Balance (2015H1)	118,143	108,109	29,300	59,534	98,043	149,398	219,929
Balance (2015H2)	119,800	115,850	25,000	57,743	98,198	151,763	227,112
Balance (2016H1)	124,175	121,525	28,302	59,246	99,952	155,683	235,932
Balance (2016H2)	128,213	126,975	29,279	60,250	101,966	160,238	244,876
Balance (2017H1)	130,608	127,003	30,000	60,775	103,092	162,999	250,191
Balance (2017H2)	143,369	148,222	29,357	61,902	108,069	178,562	286,897

(b) Filtered database

	mean	sd	p10	p25	p50	p75	p90
Balance (2015H1)	123,325	98,092	38,770	64,821	101,620	152,765	223,988
Balance (2015H2)	127,332	105,483	38,758	65,237	103,424	157,122	233,834
Balance (2016H1)	130,092	111,117	38,309	65,061	104,278	160,214	241,596
Balance (2016H2)	133,558	116,289	38,336	65,680	106,060	164,432	250,009
Balance (2017H1)	134,998	117,715	37,984	65,622	106,807	166,905	254,782
Balance (2017H2)	140,451	125,369	37,953	66,386	109,479	173,774	269,100

Notes: The above tables show summary statistics for the outstanding balance for mortgages across the PSD007 snapshots. Panel (a) shows the summary statistics for the raw database; panel (b) shows the summary statistics after the filtering steps described in Section A.3.

the estimated amortized loan balance based on the reported loan balance, remaining term, and discounted interest rate of 2017H1.[A.4]

Reported Interest Rate, Remaining Term, and Age. We drop all instances of negative interest rates, and winsorize interest rates at the 99.9 percentile point for each snapshot to address outliers that clearly arise from misreporting (for instance, interest rates of >1000%). We drop all instances of negative remaining terms, and winsorize the distribution at the 99.9 percentile point for each snapshot to address outliers that clearly arise from misreporting (for instance, remaining term of 9999 months). Finally, we drop all instances of reported negative age of borrowers.

[A.4] We estimate an amortized loan balance for 2017H2 only for discounted mortgages with at least 6 months on discounted periods in 2017H1. Further, we do this estimation only for mortgages that are on a capital and interest payment plan; i.e. we do not restate the 2017H2 loan balance for the small balance of interest only discounted mortgages in the stock of the two aberrant lenders. Further, we do not restate the loan balance of the newly issued discounted rate mortgages issued by these two lenders which leads to a slightly higher average loan balance and LTV for the 2017H2 snapshot. The higher loan balance for 2017H2 has no bearing on the cross-subsidy estimates in our paper based on data moments from the 2015H1 snapshot.

Table A.6: Proportion of Loans Ineligible for Refinancing

	2015H1	2015H2	2016H1	2016H2	2017H1	2017H2
All	100%	100%	100%	100%	100%	100%
(1) LTV>=95	2.3%	1.9%	2.2%	2.4%	2.4%	3.6%
(2) Balance<=30000	6.5%	6.5%	6.7%	6.7%	6.9%	6.9%
(3) Non-performing	5.5%	5.0%	3.9%	3.9%	3.8%	3.6%
All excl. (1),(2),(3)	86.4%	87.2%	87.7%	87.4%	87.4%	86.3%

Notes: The table above shows the share of borrowers who fall under (1) to (3). "LTV" refers to the current loan-to-value ratio; "Balance" to the current loan balance; and "Non-performing" includes short-term arrears, loans in forbearance, and loans with a possession order.

A.4 Borrowers Potentially Ineligible to Refinance

Table A.6 shows the proportion of mortgages/loans that are potentially ineligible for refinancing using specific criteria based on loan characteristics such as high LTV, low loan-balances or default status discussed in section 2 in the paper. Together, these filters account for 14% of mortgages in 2015H1, a figure that ranges from 13-14% across the snapshots.

Table A.7 presents an alternate view of the effect of these criteria by showing the proportions of mortgages remaining after filtering by these criteria. The first row titled 'All' corresponds to the sample of loans in different snapshots for which we have non-missing data for all variables, and row 4 corresponds to the fraction of this sample that remains after applying the filtering criteria to exclude borrowers that are potentially ineligible for refinancing. The 86.4% figure for 2015H1 corresponds to the sample on which we compute the cross-subsidies in our paper. Table A.7 (b) shows that the spread between the average interest rate paid on reset rate and discount rate mortgages goes down with the application of these criteria.

A.5 Summary of Dataset Construction Steps

Table A.8 shows a reconciliation from the owner-occupier mortgage stock of £ 989 mn in 2015H1 in the UK, to the sample of £ 470 mn mortgages relevant to our study. Column 2 in rows 2-8 corresponds to the effect in £ mn of each step discussed in Sections A.1 - A.4, while columns 3 and 4 show the size of the remaining data in £ mn and mn observations, respectively.

Table A.7: Sample With and Without Mortgages Unlikely to Be Able to Refinance

(a) Percentage of mortgages

	2015H1	2015H2	2016H1	2016H2	2017H1	2017H2
All	100%	100%	100%	100%	100%	100%
All excl. (1) LTV>=95	97.7%	98.1%	97.8%	97.6%	97.6%	96.4%
All excl. (2) Balance<=30000	93.5%	93.5%	93.3%	93.3%	93.1%	93.1%
All excl. (3) Non-performing	94.5%	95.0%	96.1%	96.1%	96.2%	96.4%
All excl. (1), (2), (3)	86.4%	87.2%	87.7%	87.4%	87.4%	86.3%

(b) R-r (spread)

	2015H1	2015H2	2016H1	2016H2	2017H1	2017H2
All	0.50	0.61	0.72	0.61	0.77	0.98
All excl. (1) LTV>=95 (3)	0.49	0.62	0.75	0.66	0.82	1.00
All excl. (2) Balance<=30000 (5)	0.51	0.62	0.73	0.62	0.79	0.99
All excl. (3) Non-performing (7)	0.46	0.56	0.68	0.57	0.73	0.94
All excl. (1), (2), (3)	0.45	0.58	0.72	0.62	0.80	0.96

Notes: The tables above show the share of borrowers (panel (a)) and the average interest rate spread between mortgages under reset (R) and teaser (r) rates (panel (b)) of mortgages excluding those under the filters (1) to (3). "LTV" refers to the current loan-to-value ratio; "Balance" to the current loan balance; and "Non-performing" includes short-term arrears, loans in forbearance, and loans with a possession order.

Table A.8: Steps in Dataset Construction

	Drop (£M)	£Mn	Mn
Mortgage stock in 2015H1		**989**	**8.7**
Drop if non-unique postcode + d.o.b.	60	929	7.8
Drop if not reset/discounted	251	678	5.7
Drop discounted if reported as non-incentivised	13	665	5.6
Drop if not reset/discounted in future snapshots	68	597	5.0
Drop if inconsistent/anomalous	32	565	4.6
Drop if unreported income/region	53	512	4.1
Drop if ineligible to refinance	42	470	3.6
Sample for cross-subsidies		**470**	**3.6**

Notes: The above shows the total size of the mortgage stock in 2015H1 (row 1), and the effect of each step in data construction leading up to the sample used for cross-subsidies (row 9).

B Merging Across Stock Snapshots

The high-quality disaggregated information in our database allows us to track mortgages across snapshots. In particular, we use the loan-level information on borrower date of birth and the 6-digit postcode to track mortgages across snapshots. These variables, when combined, provide a unique identifier for each mortgage.

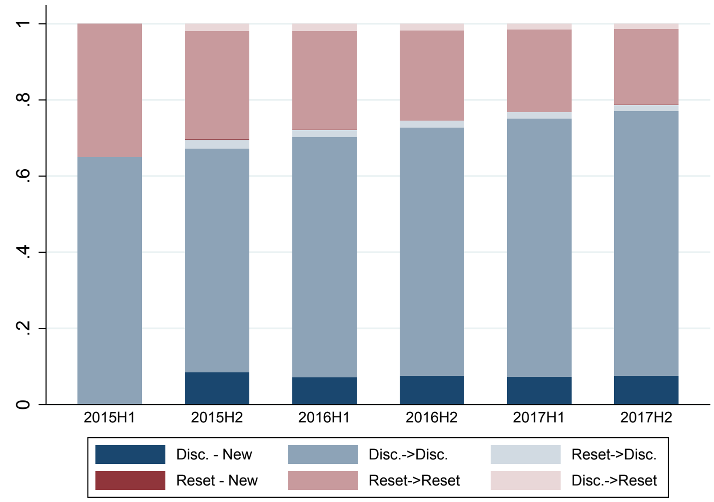
We start with the 2015H1 snapshot as the base, and merge data from subsequent snapshots using this unique identifier. For each mortgage, we can track whether it is discontinued between specific snapshots and whether it originated in any of the snapshots. Exploiting our ability to observe mortgages across snapshots, we also track whether a mortgage transitions across categories (discounted-to-reset rate or reset rate-to-discounted) between snapshots, or whether it continues in the same interest rate category. Across all 6 snapshots, the data track 6.00 million unique mortgages.

Figure B.1 provides a breakdown of discounted and reset rate mortgages in a given snapshot to show the cross-flows between these two mortgage groups across consecutive snapshots. For each snapshot starting 2015H2, we show the discounted (in blue) and reset (in red) mortgages by whether they are new (darkest shade), in the same category in the previous snapshot (lighter shade, e.g. discounted-to-discounted), or cross-flow from a different category in the previous snapshot (lightest shade, e.g. discounted-to-reset).

Observing these cross-flows by tracking mortgages across snapshots underlies one of the key moments we use to estimate our model—the transition probability of reset rate mortgages to be on discounted rates over a 2-year window. Figure B.2 shows the transition probabilities across different categories over 6-month to 24-month horizons for mortgages in the cleaned data for the 2015H1 snapshot (i.e., from 2015H1 up to 2015H2 and 2017H1, respectively). The transition from reset to discounted rate mortgages (R to r in the figure, light blue) increases from 6.73% of the mortgages after 6-months to 16.52% after 24-months.

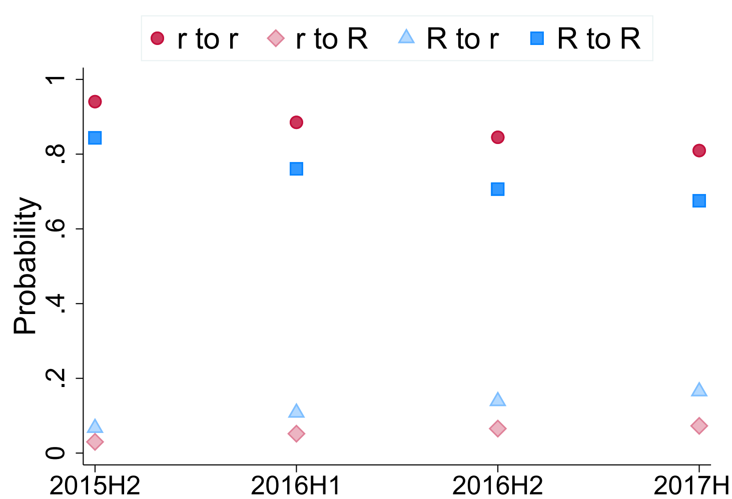
In our sample, on average, around 345,000 mortgages are originated, and around 178,000 mortgage accounts are discontinued every 6 months. Given that new mortgages in each snapshot are predominantly discounted rate mortgages, this leads to an increase in the share of discounted mortgages from 65.0% in 2015H1 to 78.7% in 2017H2. We discuss potential arguments explaining this shift in Section B.1. In comparison, the proportion of mortgages that flow across the two groups between snapshots is relatively stable.

Figure B.1: Mortgage Flows across Discounted and Reset Rate Categories



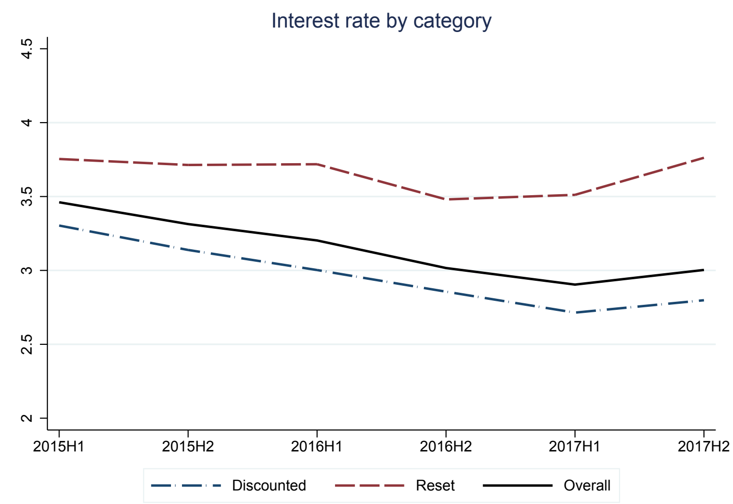
The figure above shows the proportion of mortgages under discounted rates (blue) and the reset rate (red) from the mortgage stock as reported at a half-yearly period from 2015H1 to 2017H2. The proportion of mortgages that are new to a snapshot are shown using a darker shade; and the proportion of mortgages that cross categories across snapshots are shown in a lighter shade.

Figure B.2: Transition Probabilities Over Time



Notes: This figure displays transition probabilities between rate types, namely from discounted to discounted, discounted to reset, reset to discounted, and discounted to discounted, over time. The transition probabilities are measured between the share of mortgages in the 2015H1 stock, and 6 to 24 months later (2015H2, 2016H1, 2016H2, 2017H1). Omitted transition probabilities are from discounted or reset rate to account closure.

Figure B.3: Average Interest Rate of Mortgages under Discounted and Reset Rates



Notes: The figure above shows the equal-weighted interest rate for mortgages under discounted/reset rates in the snapshots of mortgage stock reported at a half-yearly period from 2015H1 to 2017H2.

B.1 Time-Series Evolution of Mortgage Stock

The UK mortgage market has experienced a number of changes from 2015H1 to 2017H2. In particular, Figure B.1 shows a decline in the share of mortgages paying the reset rate. We note that a number of factors could explain the drop in the number of mortgages on reset rates since 2015H1. First, Figure B.3 displays a decline in the average discounted rate, and an increase in the spread between average reset rates and discounted rates after 2015H1. Second, as reported in Financial Conduct Authority (2019b), there has been an increase in lenders' focus on retaining existing customers through internal switching, and an increased role of intermediaries in prompting borrowers to undertake beneficial switches.[B.1]

Table B.1 shows the averages of selected variables in the dataset across snapshots from 2015H1 to 2017H2. The average size of discounted rate loans has risen steadily over time, while the average size of loans on the reset rate has decreased. This is consistent with the change in refinancing incentives over time highlighted above. We also observe an increase in cash-out refinancing as evident in the average loan balance of discounted-to-discounted rate

[B.1] We direct interested readers to more recent changes in the UK mortgage market aimed at facilitating switching at the time of refinancing. For instance, Financial Conduct Authority (2020a) reflects on increased use of technology and other remedies to facilitate switching; and recent policies have made it easier for for financial groups to switch customers from a group's closed book or lender to an active one (Financial Conduct Authority, 2020b), with the objective to make intra-group switching easier), and modified affordability assessments while refinancing for borrowers with up-to-date payments (Financial Conduct Authority, 2019a).

Table B.1: Summary Statistics over Mortgage Snapshots

Snapshots	2015H1	2015H2	2016H1	2016H2	2017H1	2017H2
Average loan size in £						
Discounted	140,647	143,611	145,431	147,815	149,792	152,278
Reset	112,692	111,176	109,285	108,468	107,038	105,799
Average remaining term in years						
Discounted	20.57	20.68	20.71	20.82	20.87	20.90
Reset	16.84	16.53	16.15	15.93	15.62	15.23
Average remaining term (value-weighted) in years						
Discounted	21.67	21.85	21.95	22.11	22.22	22.30
Reset	17.00	16.68	16.30	16.07	15.72	15.33
Average remaining discounted period in years						
Discounted	2.11	2.09	2.05	2.00	1.97	2.03
Average remaining discounted (value-weighted) period in years						
Discounted	2.10	2.10	2.05	1.99	1.95	2.02
Average interest rate						
Discounted	3.30	3.14	3.00	2.86	2.71	2.80
Reset	3.75	3.71	3.72	3.48	3.51	3.76
Average interest rate (value-weighted)						
Discounted	3.20	3.03	2.89	2.75	2.60	2.68
Reset	3.72	3.69	3.69	3.45	3.48	3.74
Average borrower age						
Discounted	41	41	41	41	41	41
Reset	44	45	45	46	46	46

Notes: The table above share summary statistics of mortgages for the stock snapshots from 2015H1 to 2017H2. The sample includes mortgages under two categories - those under discounted rates, and under the reset rate. Please see Appendix Table A.1 for a description of the underlying variables.

refinanced mortgages. The average remaining term on discounted rate mortgages rises from around 20.6 to 20.9 years (21.7 to 22.3 value-weighted), while the average remaining term on reset rate loans decreases through the sample period, from around 16.8 years to 15.2 years (17.0 to 15.3 value-weighted). The average remaining discounted period on discounted loans is 24 to 25 months in all snapshots of the data, reflecting the modal discounted period of 2 years observed in the data.[B.2] Finally, we observe an increase in the average interest rate gap between loans on reset rate and discounted rates over time, from 45bp (52bp value-weighted) in 2015H1, to 96bp (106bp value-weighted) in 2017H2. In all sample periods, the loan-balance weighted rate spread is higher than the equal-weighted rate spread. This effect stems mainly from larger discounted rate mortgages having lower rates on average, which is consistent with wealth-based heterogeneity in mortgage refinancing efficiency.

B.1.1 Legacy reset rate mortgages

As mentioned in Section 2.3, our data includes mortgages by two large lenders who offered to cap the reset rate for mortgages issued up to and during the 2007-09 financial crisis at 250 bp. In 2015H1, roughly 90% of all reset rate mortgages for these lenders were on these historically low rates, while the rest of the lenders had raised the interest rate paid when moving to reset rate mortgages to more than 400 bp. Consequently, excluding mortgages by these lenders more than doubles the spread between the average interest paid on reset and discounted mortgages—around 110 bp in 2015H1 (52 bp in our sample including the two large lenders) to 144 bp two years later in 2017H1.

[B.2] Mortgages under the discounted period are essentially fixed-rate loans. At origination, the most common discounted period is 2 years, followed by 5-year fixed-rate loans.

C Mortgage Moments by Region

Table C.1 shows summary statistics across UK regions and devolved administration; we evaluate cross-subsidies by UK regions in Section 5.

The regions are ordered by the average borrower income in the third column. The second column reports the population size, which shows that higher-income regions, such as London and the South East, are also the most populous. Higher-income regions tend to have higher balances as well a larger share of borrowers on the discounted rate, consistent with the patterns that we document across income quantiles in Table 2. One noticeable difference with the patterns in Table 2 is that the highest-income region, London, has the lowest homeownership rate, suggesting rich heterogeneity within and across regions.

Table C.1: Summary Statistics for the Mortgage Stock in 2015H1, By Region

Regions	Population (1,000)	Income (£)	Homeowners (%)	Balance (£)	Discounted (%)
Northern Ireland	1,852	46,236	0.69	88,790	0.59
Wales	3,099	46,443	0.67	100,026	0.62
North East (England)	2,625	46,465	0.61	93,488	0.60
Yorkshire and The Humber	5,390	47,138	0.63	100,650	0.64
East Midlands (England)	4,677	49,331	0.67	106,786	0.64
North West (England)	7,175	49,439	0.64	103,406	0.63
West Midlands (England)	5,755	50,270	0.65	110,089	0.61
Scotland	5,373	51,463	0.60	102,084	0.61
South West (England)	5,472	55,248	0.67	128,260	0.67
East of England	6,076	62,041	0.67	146,888	0.69
South East (England)	8,949	68,143	0.67	165,072	0.69
London	8,667	85,598	0.49	207,592	0.69

Notes: The table above shows summary statistics of mortgages from the stock data reported in 2015H1, split by UK regions. Regional population for June 2015 is in 1,000s. Appendix Table A.1 contains a description of the underlying variables.

D Prompt and Sluggish Refinancers

The goal of this Appendix is to check whether the dual-rate structure serves the function of screening borrowers on the basis of risk. To do so, we build on the analysis of Cloyne, Huber, Ilzetzki, and Kleven (2019) by evaluating observable differences in the characteristics of the borrowers on different rates.

Table D.1 compares households who are on the reset and discounted rates, respectively. The table further splits the data on whether the current loan balance is above median (columns 1 and 2), or below median (columns 3 and 4).

We are particularly interested in any differences between borrowers on the discounted and reset rates evident in mortgage contract characteristics and borrower attributes that capture riskiness. Table D.1 shows that borrowers on the reset rate have smaller original and current loan balances and shorter remaining terms, consistent with rationally lower incentives to refinance.[D.1] However, the two sets of borrowers have comparable loan-to-value ratios and if anything, reset rate borrowers have slightly higher incomes and slightly lower loan-to-income ratios. These patterns are consistent for borrowers with both below and above median loan balances, suggesting that they are not just driven by smaller loan balances that reflect older contracts and/or lower leverage. These summary statistics are consistent with lenders earning greater risk-adjusted returns from borrowers on the reset rate, since they pay higher average rates, but have similar LTVs and LTIs, which are standard indicators of borrower riskiness.

We further study households on discounted and reset rates by assessing the sub-sample of households whose fixed-rate period expires in a subsequent snapshot of the mortgage stock. We then compare borrowers who refinance within 6 months of the fixed-rate expiration window ("prompt"), and "sluggish" borrowers who delay refinancing past this point.[D.2] We conduct this comparison using borrowers whose fixed-rate contract expired between 2015H1 and 2015H2. Table D.2 shows that households who refinance promptly have typically larger (origination and current) loan balances (cross-sectional means are displayed with cross-sectional standard deviations below in parentheses), slightly longer remaining terms,

[D.1] By definition, as a result of being on the reset rate, the current interest rate paid by these borrowers in the table is higher than that of borrowers on the discounted rate.

[D.2] This definition of prompt and sluggish will not necessarily correspond to optimal and sub-optimal refinancing, as optimality depends on borrowers' specific circumstances, as we describe in more detail in the model section.

Table D.1: Characteristics by Rate Type and Loan Balance

	Disc. (high)	Reset (high)	Disc. (low)	Reset (low)
Original Loan Balance	205,634	187,929	82,467	86,351
	(117,247)	(97,522)	(27,204)	(32,456)
Original Term	24.84	23.97	21.35	23.04
	(6.52)	(6.42)	(7.59)	(6.94)
Current Loan Balance	196,766	176,047	72,114	69,137
	(112,791)	(90,371)	(20,954)	(21,058)
Current Interest Rate	3.22	3.72	3.41	3.78
	(0.93)	(0.97)	(0.96)	(0.98)
Remaining Term	22.56	17.53	18.15	16.37
	(7.02)	(6.80)	(7.86)	(7.01)
Borrower Age	39.47	44.35	42.07	44.49
	(8.78)	(9.32)	(10.63)	(10.56)
Current Income	74,855	78,754	38,665	41,894
	(261,183)	(109,242)	(37,069)	(48,369)
Current LTV	62.76	62.64	51.46	53.32
	(17.01)	(17.07)	(21.61)	(22.20)
Current LTI	2.98	2.65	2.21	1.97
	(1.61)	(2.37)	(1.16)	(1.06)
N	1,283,633	511,486	1,051,121	743,988

Notes: The table above compares means and standard deviations (in parentheses) for households across rate types, split by current loan balance. Columns 1 and 2 report values for households who have above median loan balances, while columns 3 and 4 report values for households who have below median loan balances.

higher income levels, but comparable LTI and LTV ratios.

Table D.2: Refinancing Decisions Due Between 2015H1 and 2015H2

	Prompt	Sluggish	Repaid/Moved
Original Loan Balance	155,221	120,994	158,886
	(107,043)	(81,381)	(113,046)
Original Term	23.82	23.33	23.15
	(6.97)	(7.79)	(7.78)
Current Loan Balance	143,670	111,216	143,441
	(102,432)	(77,468)	(107,417)
Current Interest Rate	3.33	3.75	3.38
	(0.98)	(1.10)	(1.03)
Remaining Term	20.79	20.14	19.77
	(7.47)	(8.36)	(8.26)
Borrower Age	40.66	41.13	41.13
	(9.42)	(11.05)	(10.73)
Current Income	59,104	51,332	65,300
	(64,198)	(120,794)	(69,529)
Current LTV	57.35	57.10	54.16
	(17.95)	(20.18)	(19.43)
Current LTI	2.66	2.50	2.48
	(1.44)	(1.40)	(1.36)
N	207,320	49,586	19,482

Notes: The table above compares means and standard deviations (in parentheses) for households whose fixed-rate contract was due to expire between 2015H1 and 2015H2. Column 1 reports households who refinanced within 6 months after their contract expiration date ("prompt"), column 2 reports households who did not refinance in that window ("sluggish"), and column 3 reports households who prepaid the loan and leave the sample.

E Option Value of Staying on Reset Rate

In this Appendix, we evaluate to what extent households not refinancing reflects an optimal option exercise, rather than their refinancing cost. As we explain below, reset rates do not come with prepayment penalties and so allow households to flexibly refinance in case interest rates go down. If there is an option value associated with waiting on the reset rate, it is difficult to interpret k as a behavioral parameter, the need for which is eliminated in the counterfactual. In order to evaluate the option value, we apply a standard approach in the literature and customize it to see if it can deliver an option value that justifies paying the reset rate for short periods, and find that it cannot, for reasonable parameter values.

E.1 Reset Rates and the Option To Refinance

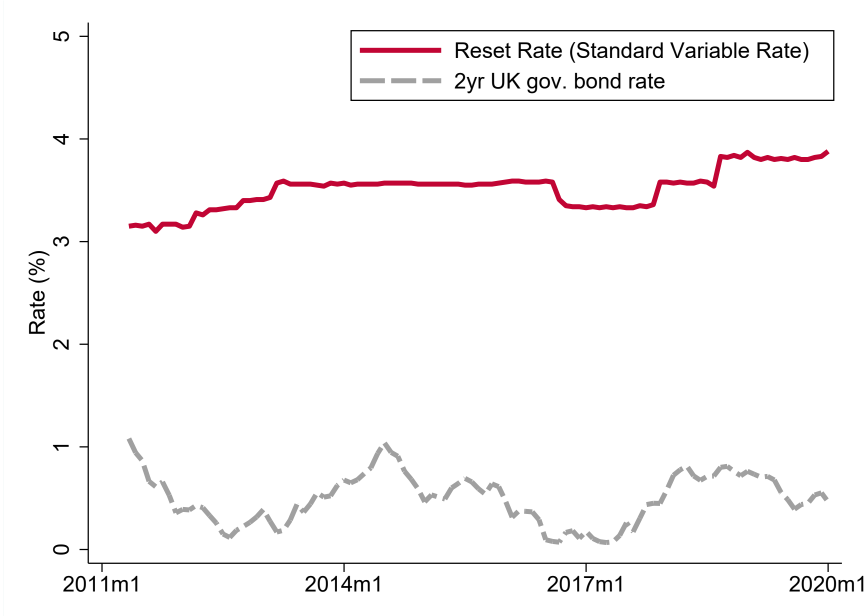
UK mortgages have fixed rates and substantial prepayment penalties over the discounted fixation period, such that households typically lock in the rate and refinance next at the end of the fixation period. The reset rate, also called the "Standard Variable Rate" (SVR), in contrast, allows households to retain the option to refinance flexibly when interest rates go down. Figure E.1 shows the average reset rate and variation in 2-year government bond yields since 2011.

While reset rates do appear linked to underlying UK Treasury rates, they adjust at a slower rate and have less variability. Households incur a cost to retain the option to refinance, which is the difference between this slow-to-adjust and high reset rate R and the discounted rate r that they can currently access in the market. In what follows, we conduct a quantitative assessment of the net benefit of this refinancing option.

E.2 Optimal Refinancing Differential

In a standard refinancing framework with long-term fixed-rate mortgages, households rationally evaluate the present value of interest payments that they make under the new rate into which they refinance, and compare the payments they would make on this rate with those on the rate they would otherwise be in, accounting for any refinancing costs incurred, plus any difference between the value of the refinancing option that they give up, and the value of the new refinancing option that they acquire (Chen and Ling, 1989; Agarwal, Driscoll, and Laibson, 2013). Households optimally exercise their option to refinance when

Figure E.1: Reset Rates Over Time



Notes: This figure shows the average Standard Variable Rate and 2-year UK Treasury bond yield at monthly frequency, as reported by the Bank of England Database.

the new rate is sufficiently lower than the rate they would otherwise bear (the "old rate"). This decision can be characterized using a "threshold," which is a specific value of the differential between the new and old rates beyond which it is rational to refinance.

Agarwal, Driscoll, and Laibson (2013) derive an analytical solution to this class of refinancing problems. They propose that households should refinance when the difference between the current mortgage interest rate (r_t) and the old rate (r_0), denoted by Δr, is greater than the optimal threshold Δr^*

$$\Delta r^* \equiv \frac{1}{\psi} \left(\phi + W\left(-\exp(-\phi)\right)\right),$$

where $W(.)$ is the principal branch of the Lambert W-function, $\psi = \frac{\sqrt{2(\rho+\lambda)}}{\sigma_r}$, and $\phi = 1 + \psi(\rho + \lambda)\frac{\kappa/M}{(1-\tau)}$. The optimal threshold depends on the real discount rate ρ, the expected real rate of exogenous mortgage repayment λ, the standard deviation of the mortgage rate σ_r, and κ/M, the ratio of refinancing cost and outstanding loan balance.

To make this framework applicable for UK borrowers considering whether or not to refinance into a discounted rate mortgage, we set κ to the median persistent component k_i

of borrowers, £634, a conservative estimate of total refinancing cost, and M as the average loan balance £130,871. σ_r is set to the historical standard deviation of 2-year UK real rates, at 0.0193. We further follow Agarwal, Driscoll, and Laibson (2013) to compute the rate of mortgage repayment as

$$\lambda = \mu + \frac{r_0}{\exp\left[r_0 T\right] - 1} + \pi,$$

where μ is the (annual) probability of prepayment, T is the remaining loan maturity, which we set to 30 years, and π is the rate of inflation, which we set to 2%. We set μ to 0.5 to capture the expected holding period of 2 years.[E.1]

For this representatively calibrated household, the optimal refinancing differential that we obtain is 111 basis points. This estimate provides a quantitative sense of by how much interest rates need to decrease for the refinancing option to be valuable, i.e. to be "in the money". For comparison, Figure E.2 shows how the optimal refinancing differential would vary by loan size. Intuitively, since the refinancing cost is fixed, the refinancing differential is decreasing in loan size, as the refinancing benefit is scaled by the loan balance. The majority of loans in the stock of mortgages have loan balances that are smaller than £200,000, and so require differentials between around 100 to 150 basis points.

Next, we simulate how likely it is that this refinancing threshold is hit, i.e., how likely the option is to be in the money, and the expected value of the option.

E.3 Simulation of Option Value of Refinancing

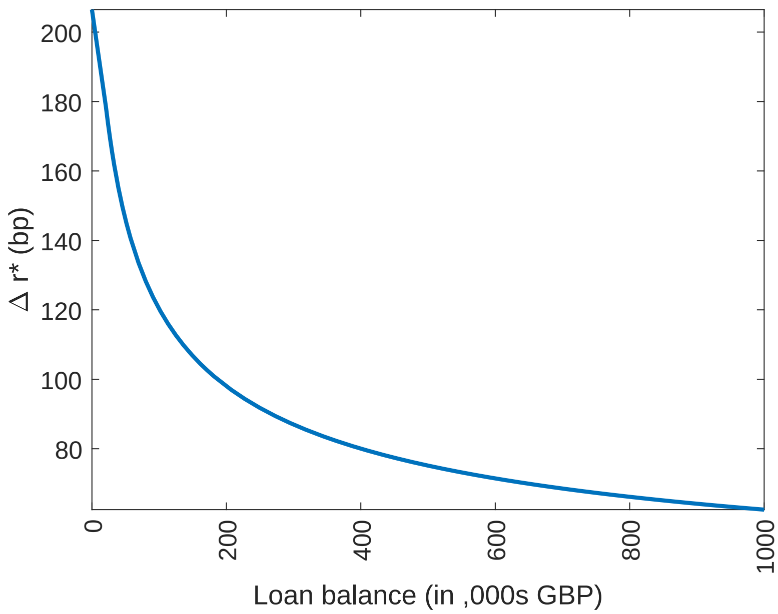
For a given old interest rate r_0, the optimal refinancing threshold calculated in the previous subsection (i.e., 111 basis points in our calibration) characterizes the level of the current mortgage rate r^* below which it is optimal to refinance. The expected value of the refinancing option is then

$$\begin{aligned} &\int_{-\infty}^{+\infty} (r_0 - r_s)\mathbb{I}_{(r_s \leq r*)} f(r_s) dr_s \\ &= \int_{-\infty}^{r_s \leq r^*} (r_0 - r_s) f(r_s) dr_s. \end{aligned} \tag{E.1}$$

We can simulate the expected value of this expression by specifying a data-generating

[E.1] In this framework, the time over which the refinancing benefit accrues is primarily captured via the prepayment probability μ, rather than T.

Figure E.2: ADL Threshold Under Different Loan Sizes



Notes: This figure plots the optimal refinancing threshold using the formula by Agarwal, Driscoll, and Laibson (2013) under the main calibration, when varying the loan balance.

process for interest rates. Suppose interest rates follow a standard AR(1) process:

$$r_t = (1 - \rho_r)\mu_r + \rho_r r_{t-1} + \epsilon_t, \tag{E.2}$$

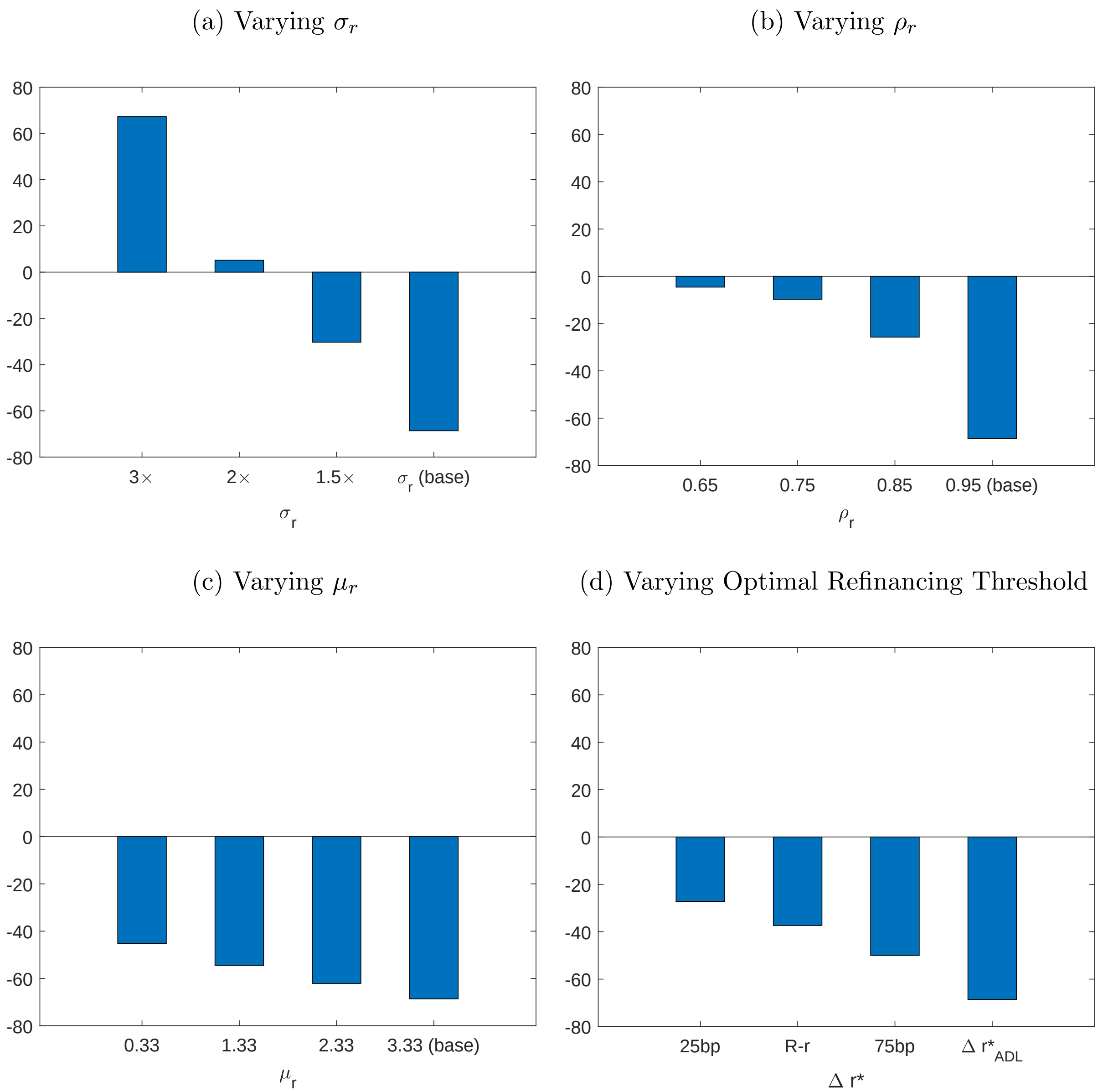
where ϵ_t is a normally distributed white noise shock with mean zero and variance σ_ϵ^2, and ρ_r is the autocorrelation coefficient. The variance of the white noise shock is related to the variance of interest rates σ_r via $\sigma_\epsilon = \sqrt{\sigma_r^2 \cdot (1 - \rho_r^2)}$.

Households in our setting compare two options. The first is to take out a new 2-year fixed-rate contract right away (subsequently refinancing again after 2 years). The second is to stay on the reset rate and then refinance into a new 2-year fixed-rate contract whenever the optimal refinancing threshold is met. We simulate the expected NPV of such an option exercise, assuming that households can choose to wait up to 2 years to exercise the option to refinance. We then evaluate the NPV over a period of 4 years to ensure that we compare all households' NPV of option exercise over a similar period, regardless of when they fix again within the two-year window.

Figure E.3 shows simulations of the expected NPV of the option under different calibrations of the interest rate process and optimal refinancing thresholds. Panels (a), (b) and (c) show that the NPV is only positive if interest rates are substantially more volatile than they have been historically, by a factor of 2 to 3. Varying the persistence or long-run average of the interest rate process within reasonable ranges are not sufficient to yield a positive value. In addition, Figure E.3d shows the NPV for different assumptions about the refinancing threshold—lowering it to 25 and 75bp rather than calibrated ADL value of 111 bp. We also consider an alternative benchmark, which is a simple rule-of-thumb to refinance whenever the current interest rate makes up for the reset rate differential $R - r$. Under all these scenarios, we find that the expected value of the option remains negative.

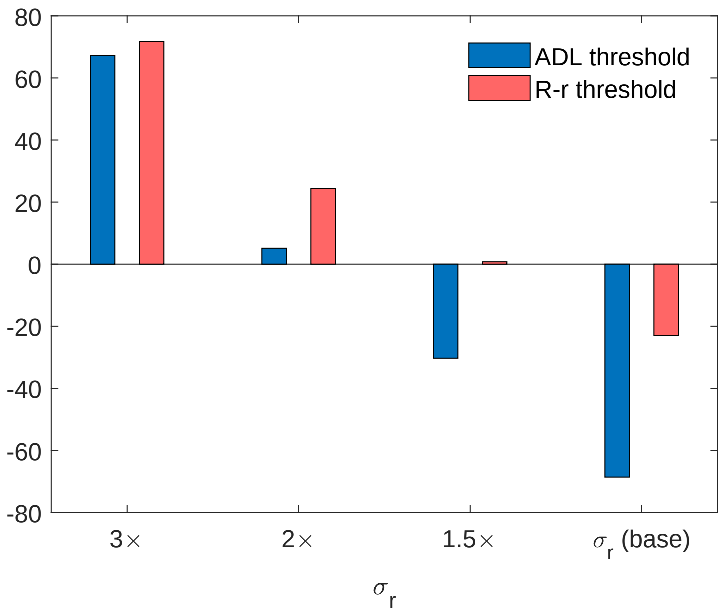
Lastly, Figure E.4 shows that the expected value of the refinancing option is around 20 basis points when households refinance based on the $R - r$ threshold, under a counterfactual interest rate volatility that is twice the historical average. Based on these simulations, we conclude that the option value of staying on the reset rate is not economically significant under a benchmark calibration of interest rates. This is because the cost of retaining the option is very high (this is the spread of the reset rate over the discounted rate) which penalizes waiting, and because the window over which the option can be exercised is relatively short, corresponding to the typical fixation window of 2 years.

Figure E.3: Expected NPV of Refinancing Option on Reset Rate



Notes: This figure displays the simulated net present value of the option to refinance when staying on the reset rate, under different calibrations of the interest rate process and the optimal refinancing threshold.

Figure E.4: Varying Optimal Refinancing Threshold and σ_r



Notes: This figure displays the simulated net present value of the option to refinance when staying on the reset rate, under different calibrations of interest rate volatility and comparing the ADL optimal refinancing threshold with a threshold that corresponds to the difference between the reset and discounted rate.