Financial Conduct Authority
Occasional Paper 33

May 2018

Choices of dominated mortgage products by UK consumers

Zanna Iscenko

FCA occasional papers in financial regulation

FCA Occasional Papers

The FCA is committed to encouraging debate on all aspects of financial regulation and to creating rigorous evidence to support its decision-making. To facilitate this, we publish a series of Occasional Papers, extending across economics and other disciplines. The main factor in accepting papers is that they should make substantial contributions to knowledge and understanding of financial regulation. If you want to contribute to this series or comment on these papers, please contact Karen Croxon at Karen.Croxon@fca.org.uk.

Disclaimer

Occasional Papers contribute to the work of the FCA by providing rigorous research results and stimulating debate. While they may not necessarily represent the position of the FCA, they are one source of evidence that the FCA may use while discharging its functions and to inform its views. The FCA endeavours to ensure that research outputs are correct, through checks including independent referee reports, but the nature of such research and choice of research methods is a matter for the authors using their expert judgement. To the extent that Occasional Papers contain any errors or omissions, they should be attributed to the individual authors, rather than to the FCA.

Authors

Zanna Iscenko

Zanna Iscenko works at the FCA and is a PhD candidate at University College London.

Acknowledgements

I would like to thank Richard Blundell, João Cocco, Ian Preston, and participants in the 2016 FCA/Imperial College Mortgage Conference for comments, suggestions and challenges. I am also very grateful to Peter Andrews, Nick Cook, Carmen Suarez, Damien Fennell, Jeroen Nieboer, Simone Pedemonte, Teresa Bono, Adiya Belgibayeva and the rest of the Mortgages Market Study Team, as well as many other past and present employees at the FCA, as this research would not have been possible without their support and feedback.

All our publications are available to download from www.fca.org.uk. If you would like to receive this paper in an alternative format, please call 020 7066 9644 or email publications_graphics @fca.org.uk or write to Editorial and Digital Department, Financial Conduct Authority, 25 The North Colonnade, Canary Wharf, London E14 5HS.

Contents

1	Executive summary	3
2	Background on the UK mortgage market	8
	Product design	8
	Price posting and search	9
	Tools and policy context	9
3	Data	11
	Datasets	11
	Sample construction and representativeness	12
	Descriptive statistics	13
4	Introduction to dominance analysis	16
	Identification of dominated choices	16
	Cost of dominated choices	18
	Headline results on frequency and costs of dominated mortgage choices	20
5	Demographic variation	22
	Univariate trends	22
	Regression specification	22
	Marginal demographic effects	23
6	Dominating products	27
	Salience of price features	27
	Product eligibility requirements	27
7	Familiarity	29
	Intermediary search	29
	Direct search	30
8	Robustness	32
	Prime borrowers only	32
	Cost specification	33
	Sample representativeness	34
	Definition of strong dominance	34
	Are choices better than random?	34
9	Discussion and conclusion	36
A	Figures and tables	40

1 Executive summary

With housing debt accounting for over 80% of total UK household liabilities, choosing a mortgage is one of the most important financial decisions consumers make.[1] It is also a very difficult one to get right. Even after deciding which house to buy and how much to borrow, a consumer is faced with comparing dozens, sometimes even hundreds, of mortgage products with many different features, such as the duration of the fixed-rate period, and multiple price components. It is hardly a surprise that regulators around the world, including the FCA, are interested in how well consumers can make decisions about their mortgages and what outcomes they experience.[2]

This paper adds to the evidence on these topics by investigating a uniquely granular combination of lending transaction reports, detailed product information and credit files for nearly 700,000 UK households that took out a mortgage between January 2015 and July 2016.[3] I also develop and implement a new methodology for measuring consumer outcomes in the mortgage market or other circumstances with multidimensional product prices and features. The proposed methodology — dominance analysis — focuses on choices of unambiguously 'dominated' products by assessing, for each observed borrower, whether they were eligible for another product with the same features as their chosen mortgage but at a strictly lower price (i.e. having rates or fees that are lower with none that are higher). This makes it possible to detect consumers who are likely to not be searching the market effectively without trade-offs between different price components or additional assumptions about consumer preferences for different mortgage types.

I find that nearly one in three UK consumers (29.9%) in 2015/16 chose mortgage products that were dominated by apparently available alternatives and paid £550 more per year on average as a result. I also provide insights on demographic characteristics, product features and consumer search patterns that could be contributing to these widespread dominated product choices.

Methodology

The dominance approach used in this paper is, in short, intuitively similar to:

- retrospectively running an extremely detailed price comparison website search (including eligibility checks) for each of the observed households
- checking which of the products consumers actually purchased were strictly worse in terms of cost than some other available alternative with exactly the same non-cost features (eg fixed rate duration) and, finally,
- identifying consumers who took out these 'dominated' products and calculating the avoidable excess costs each of them incurred as a result

The benefit of looking at dominated products rather than just comparing borrowing costs (eg APR) across consumers is that it tackles a common challenge to measuring consumer

[1] Bank of England, Money and Credit Statistics January 2018 release, and Student Loans Company, Student Loans Balance FY16/17.

[2] Examples of recent regulatory activity in the mortgages sector include FCA's Mortgages Market Study (FCA, 2018), CFPB's Know Before You Owe initiative and the ongoing Residential mortgage products price inquiry by the Australian Competition and Consumer Commission.

[3] Section 3 describes the datasets and their representativeness in more detail.

outcomes when price and product suitability can often vary at the same time and may offset each other. It allows me to detect poor consumer outcomes on cost while being reasonably sure that suitability is held constant:

- Using extensive data about borrowers and eligibility criteria for different mortgages makes it possible to identify the relevant alternatives for which the borrower satisfied observable lending requirements. (Assessing factors including, for example, loan amount, loan-to-value ratio, property type and value, borrower income, employment status, age, history of credit impairments and credit score.)
- Only comparing products with identical substantive non-cost features means consumers are not just paying extra to get a particular type of mortgage they prefer.
- Finally, by requiring that the superior 'dominating' alternative be at least as good on all price dimensions separately, I can ensure that it is not unsuitable due to the borrower's circumstances, such as inability to pay higher up-front fees or desire to keep early redemption charges low due to possible risk of moving.

More information about on the dominance methodology can be found in Section 4.[4]

Findings

I find that nearly a third (29.9%) of chosen mortgage products were dominated on cost by apparently available alternative mortgage products. The average cost savings forgone as a result of a dominated choice ('excess cost') are £550 per year, equivalent to 12.7% of borrowers' annual mortgage costs, over the course of their initial deal period. Variation in excess costs is considerable: a quarter of borrowers pay less than £180 per year while top 25% pay excess costs above £580. To focus the demographic analysis on consumers with material extra costs, I also define *strongly* dominated mortgage choices as those with excess costs above both £250 per year and 5% of the borrower's total annual mortgage costs. Over 17% of all borrowers (around 60% of those with dominated products) fall into this category.

Why do so many consumers make dominated choices? The rest of this paper investigates several potential explanations for the apparent difficulties consumers have with finding good deals.

First, the ability and willingness to search for products effectively varies a lot across different consumer demographics and is shaped by both borrowers' own financial capability and the complexity of their circumstances. Section 5 describes demographic findings in more detail, and explains the methods used to reach them.[5]

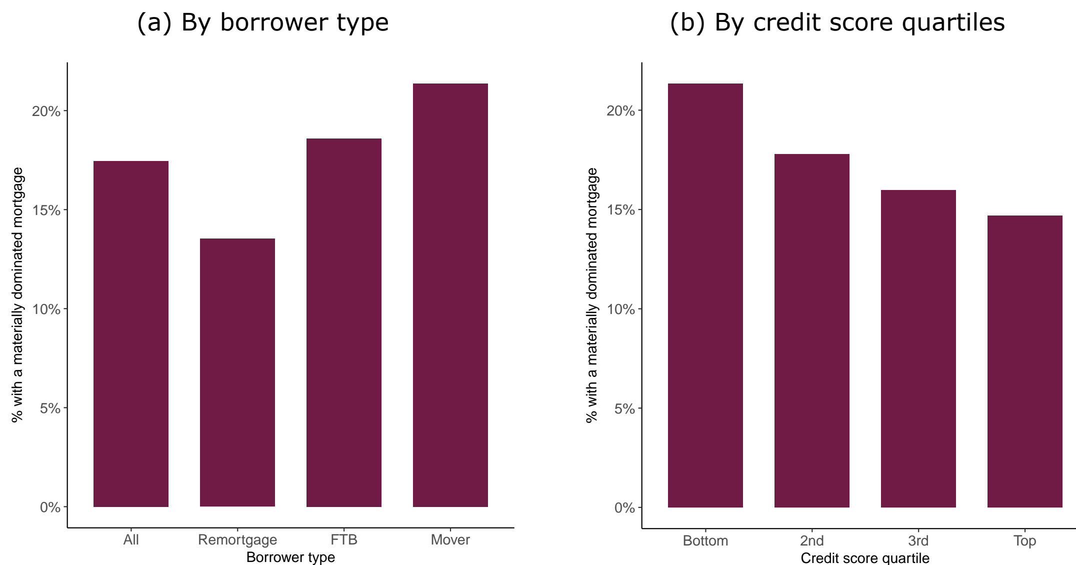
Even after I control for eligibility and the effects of other factors, consumers are more likely to make dominated choices when they have characteristics associated with lower financial capability, such as low income, old age or poor credit history (reflecting the ability to manage finances in general). For example, strongly dominated (costly) choices are about

[4] Because of the strict conditions required for dominance, when it takes into consideration all significant eligibility criteria, this methodology is likely to give a lower bound on the share of consumers who are not minimising their borrowing costs effectively. These are not necessarily errors, however. Complexity of the search, time pressure or gaps in available information (if present) could make it reasonable for some of the consumers to 'settle'.

[5] Unless stated otherwise, the paper reports results based on combined data covering both intermediated and direct transactions. Exploring the variation in demographic effects within these two groups separately is outside of the scope of this research. For some information on how APR costs vary with consumer demographics and other factors for intermediated transactions, see Belgibayeva and Majer (2018).

Figure 1.1: Rates of strongly dominated product choices, by demographics

(a) By borrower type (b) By credit score quartiles



a third less likely among consumers in the top 25% by credit rating compared to those in the bottom 25%.

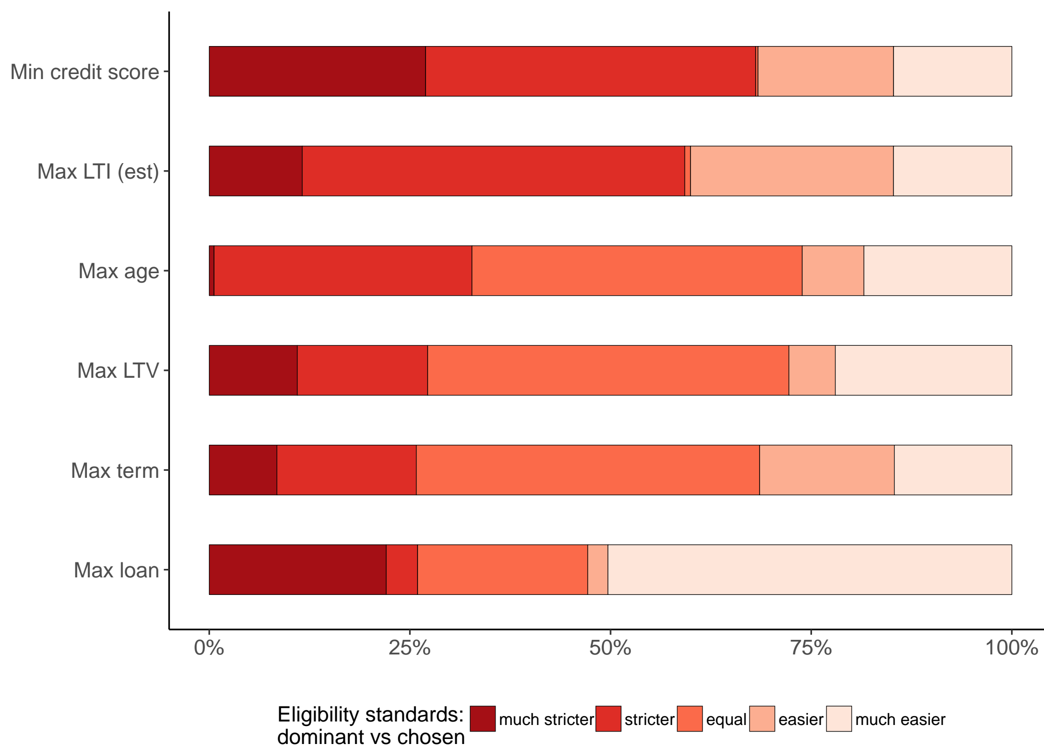
I also find that factors which are likely to make it more difficult to search for a mortgage increase the probability of choosing a dominated product. For instance, home movers (who are often buying and selling simultaneously) are likely to have a more complicated housing transaction to deal with than first-time buyers, who in turn are more likely to be distracted by the demands of a house purchase than consumers who are simply remortgaging. The rates of strongly dominated choices for borrower types reflect this intuition — only 13.5% of remortgagers choose a strongly dominated mortgage product, compared to 18.5% of first-time buyers and 21.3% for borrowers moving house. Other common complicating factors — being self-employed, buying a newly-built or unusual property or even simply having more mortgage alternatives to compare — also increase dominated choices after robust controls for other factors.

Next, in Section 6, I review product characteristics in search of systematic features that might lead lower cost alternatives to be overlooked in favour dominated options. Although some behavioural economics models suggest that borrowers might make poor choices because they focus on headline prices and overlook other fees and charges, this is not common for dominated mortgages. For 92% of dominated choices, a superior alternative had a lower initial (headline) rate than the chosen dominated product. Nor did consumers appear to focus on comparing any other single price feature in their product choice.

The situation is more ambiguous for differences in eligibility criteria between the chosen mortgages and alternatives that dominate them on cost. I find that in general better products do not have higher *explicit* borrowing criteria than mortgages they dominate. On the contrary, compared to the chosen mortgage, almost 3 in 4 of dominating products had equal or more lenient requirements on age, mortgage term, loan amount and LTV. Higher explicit standards are therefore not what is putting qualifying borrowers off cheaper deals.

Dominating mortgages did, however, tend to be somewhat stricter than borrowers' chosen products on less observable criteria, which I am able to derive from data on what credit scores and loan-to-income (LTI) ratios of applicants the banks accepted for each product. Some of the dominated choices could plausibly be a result of consumers being cautious

Figure 1.2: Eligibility criteria of dominating products compared to chosen dominated option



in the face of uncertainty. They might not be selecting the cheapest mortgage they, in fact, qualified for due to not knowing or being unsure that they satisfied all the lending requirements.

Finally, Section 7 explores whether dominated choices are likely to arise from borrowers or their intermediaries only considering a subset of available alternatives and thus overlooking cheaper products.

It is clear from the data that decision-makers tend to gravitate towards familiar alternatives. Intermediaries appear to prefer dealing with small sets of familiar lenders: an average intermediary observed in the data used just 5 mortgage lenders for 75% of their business.[6] Even for larger intermediary networks, this number of regularly used lenders was only around 10. In a half of dominated choices made through intermediaries, none of the strictly cheaper alternatives were offered by the intermediary's regular lenders, so it is plausible that they may not have been identified or considered due to lack of familiarity.

The apparent tendency to choose familiar options is even more pronounced in non-intermediated transactions. Around 60% of those consumers took out a mortgage with one of the lenders with whom they already had financial product (current account, loan or credit card), despite those lenders accounting for less than 20% of products for which the relevant consumer was eligible. Moreover, for 70% direct consumers who ended up with dominated mortgages, the superior alternatives were less 'familiar': not offered by lenders with whom they had an existing relationship.

In fact, after controlling for a variety of demographic and situational factors, choosing a lender with whom the consumer already had a financial product increased the probability of getting a strongly dominated mortgage by 3 percentage points (to 21% from the sample average of 17.5%). Initial evidence suggests use of familiar lenders by intermediaries as-

[6] It is not the case that all intermediaries used *the same* 5 lenders, however.

sociated with somewhat higher rate of dominated choices, although this relationship needs further research.

Robustness and limitations

This paper finds that a significant minority of consumers in the UK mortgage market have recently ended up with deals that are strictly inferior on cost to their other apparent alternatives. Looking into these choices in more detail is useful for learning about how outcomes vary across different mortgage borrower subgroups, and what barriers to searching the market consumers might be facing. Any detailed policy implications of these findings, however, need to be interpreted in the broader context of the market as discussed in the FCA Mortgages Market Study (FCA, 2018).

The dominance methodology developed in this paper offers a way of detecting poor consumer outcomes on cost that is more robust to uncertainty about consumer preferences and needs than traditional cost comparison metrics such as APR benchmarking. This robustness comes at the price of only being able to 'count' choices that *cannot* be explained on cost grounds under any assumptions about the consumer's future behaviour or preferences for product features. There may be cases where a borrower has alternatives that are not strictly superior on all price characteristics to their chosen product, but would still have been better for that individual consumer given their preferences.[7] This means that some of the 70% households that aren't making dominated choices could still be having difficulties with searching for a good mortgage deal.

On the flipside, the level of dominated choices could be lower than reported in this paper if lenders impose additional eligibility criteria that are not reported in the data and cannot be inferred from observed transactions (in the way that for example minimum acceptable credit scores can be). If that were the case, some of the superior mortgage products might not have been available to the consumer when they appeared to be.

In practice, however, this concern is unlikely to be materially distorting the headline results reported in the paper. The rate of dominated choices is 26%, close to the general sample, even among the very prime borrowers—those with relatively high incomes, low LTVs, and no factors that could typically complicate a mortgage application (self-employment, unusual property, etc). Those types of borrowers would be less likely to be investigated closely by the lenders and are much less likely to have their options restricted by idiosyncratic criteria as well. The high rate of dominated choices even among this very prime set of borrowers suggests that any additional unobservable lending criteria are unlikely to be a major driver of the economically significant results of this research. This and other robustness checks are discussed in more detail in Section 8.

[7] For instance, the alternative might have slightly higher fixed fees but a lower interest rate than the chosen product, and the specific borrower has no strong need to minimise up-front costs and their loan is large enough loan for interest rate savings to offset the higher fees. The chosen product might have been a better option for another household, however, with a higher preference to minimise borrowing costs early on or a much smaller loan.

2 Background on the UK mortgage market

There is limited research on price dispersion in the mortgage market. Woodward and Hall (2012) and Allen, Clark, and Houde (2014a; 2014b) are the main contributions on this topic, focusing on the US and Canadian mortgage markets respectively. The UK mortgage market has not received much attention in the literature on price dispersion specifically or household finance issues more generally. Institutions in the UK mortgage market are rather different from those in other countries more commonly covered in the mortgage literature. To help interpret the methodology and findings, I first outline some of the relevant key institutional features of the UK market below.

Product design

The first distinctive feature of the UK market is that mortgages with interest rates that are fixed for terms longer than 5 years, let alone until maturity, are not typically offered by lenders. The predominant mortgage types involve interest rates that are fixed for a relatively short period by other countries' standards (2-5 years) and then reset to standard variable rate (reversion rate) - an interest rate which can be varied by the lender at their discretion, but in practice tends to move in line with the Bank of England monetary policy rate. Two-year fixed-rate contracts are by far the most common mortgage type, accounting for almost two-thirds of all new mortgage lending.

The small remaining minority of transactions (around 10%[8]) involve fully variable rate mortgages that typically track Bank of England base rate in some form for the contractual term. However, they too often offer more attractive terms in the early years of the contract, for example in the form of a lower spread over the underlying rate. The initial years of the mortgage during which the borrower benefits from special terms, such a fixed rate or a discount on the variable rate is called the 'deal period'.

Once their deal period expires, UK households do not appear to display systematic inertia in responding to refinancing incentives to quite the same extent that has been recently reported in household finance literature in the USA (eg Agarwal, Rosen, and Yao, 2015; Keys, D. G. Pope, and J. C. Pope, 2016) and Denmark (Andersen et al., 2015). In fact, according to the FCA (2018), a significant majority of borrowers on initial fixed rate deals — around 77% — tend to refinance their mortgages within 6 months of their deal period ending. Borrowers are often approached by their own lender or a previously used intermediary with offers to take out a new deal around that time. The 'exit costs' after the deal period expiry are typically low, just covering administrative charges of closing an account. Refinancing before the deal period expires involves high penalty charges — of around 1-4% of the outstanding loan amount. Hence, refinancing before deal period expiry is highly uncommon in the UK, although it can occur in cases of unexpected location or lifestyle changes.

There is also a lot of product heterogeneity on additional product features: for instance,

[8] Here and henceforth, the descriptive statistics are based on the dataset used for research, which covers all borrowers buying a new property and those refinancing an existing property with a new lender, but does not cover borrowers refinancing an existing mortgage with the same lender. For this reason, market statistics may be slightly different from those presented in other sources.

whether the mortgage contracts allows the borrower to repay the loan faster than implied by their standard monthly instalments or to pay less for a short period of time.

Price posting and search

Another feature of the UK mortgage market that is different from settings covered in the mortgage literature is the absence of bilateral negotiation between individual borrowers and the lender (or the lender's agent).

Instead, at any given point in time, a UK lender posts a 'menu' of contracts they offer, each associated with specific eligibility criteria (eg borrower loan-to-value ratio (LTV)), features (eg length of the fixed-rate period and the degree of flexibility in repayments) and the corresponding set of prices, which normally include an up-front fee, the interest rate that applies during the deal period, the rate after the deal period ends, exit fees and any early refinancing penalty charges.

For most segments of the market, lenders offer many permutations of deal period term, product type and price structure to choose from. Combined with a relatively large number of lenders operating in the market, this leads to very large choice sets. For instance, in 2015/16 an average borrower was eligible for nearly 500 mortgage products.[9]

For each borrower, their required loan amount, the property they intend to purchase and their demographic and financial characteristics determine which subset of the menu they are eligible for. Each borrower selects a product form this subset, and applies for that particular mortgage with no scope for varying price terms or fees, which for example is described as the main driver of price dispersion in the Canadian market (Allen, Clark, and Houde, 2014a).

Following the application, the lender evaluates whether this particular borrower fits their posted (and internal) risk criteria for the chosen mortgage and accepts or rejects the application. The lender does not normally suggest alternative products when rejecting an applicant. In verifying borrower eligibility, lenders normally rely on extensive information collected from the borrower, as well as access to detailed credit files that pool information about the borrower's creditworthiness across creditors and industries.

Tools and policy context

During the recent time period covered by this research, the predominant majority (over 97%) of UK mortgage borrowers received mortgage advice before they applied for a mortgage loan. The primary function of advice is to help the consumer identify a suitable product in terms of its features — for instance, the length of the fixed-rate period. Advice does not need to be independent or to compare options across the whole of the mortgage market. In fact, many lenders have their own advisers who help the consumer decide which product from the lender's range is best for them or how long the term of their mortgage should be. Borrowers need to take additional steps to opt out of getting advice. As a result, consumers who take out mortgages without advice are lively to have higher incomes and a lot of confidence in their financial capability.

When searching for a mortgage, most consumers (over 70%) use an independent inter-

[9] Sample mean for the size of the 'available product set' as defined in definition 1 in section 4.1.

mediary rather than going to a lender directly. Intermediaries provide advice as described above, but they are also able to search for mortgage deals across multiple lenders and help the borrower complete their application. Some intermediaries charge the borrower a fixed up-front fee, but commission from lenders for a successful referred applicant is typically a larger source of revenue.

Use of intermediaries in not a major focus of this paper, although they are relevant for understanding the different ways in which borrowers search for and compare mortgage products. Iscenko and Nieboer (2018) focus specifically on the effects of intermediation and advice, and provide more institutional context about these services.

3 Data

Datasets

Most of the research in this paper relies on the combination of the following datasets.

FCA Product Sales Data The main source of data for this paper is the Product Sales Data (PSD), a confidential administrative dataset of new residential mortgage transactions reported by all UK lenders to the Financial Conduct Authority (FCA). Each transaction report is anonymised but contains detailed information about the characteristics of the loan (eg date, identity of the lender, loan amount, interest rate type), of the property used as collateral (postcode, purchase price) and of the borrower (eg age, date of birth, gross income, adverse credit events in the past).[10]

Moneyfacts I supplement the PSD data with a more granular commercial dataset on product characteristics available from one of the leading UK product data aggregators, Moneyfacts, to obtain detailed product information needed to understand total borrowing costs and constructing choice sets, such as lenders' eligibility criteria. These product data provide daily snapshots of all mortgage products available on the market from 2011 to mid-2016 and contain detailed information on each product's complete price structure, extra features and lending standards.

Credit reference data I also have access to selected variables from borrowers' credit files at the time of their mortgage application. This confidential information has been provided to the FCA by one of the UK's big 3 credit reference agencies (credit bureaus in the US) and covers over 90% of all mortgage transactions recorded in regulatory data. In addition to showing credit scores for the mortgage application, the available credit file variables also provide a rich picture of the borrower's creditworthiness and financial behaviour. The credit bureau has provided credit scores and other credit metrics for each borrower as they were around the time of mortgage application, so this information is the best approximation to the lender's information set after the credit check on the applicant borrower.

Other sources of data Some of the postcode level information about the demographic profile of the borrower's primary residence comes from other public and quasi-public sources. First, I rely on HM Land Registry's Price Paid Data to identify whether a mortgaged property is newly built, in cases when this information not available in the PSD. This dataset contains addresses of all properties sold in England and Wales since 1995. If a mortgage is taken out on a property in a new postcode that starts appearing in Land Registry data only around the time of the transaction, this is a strong indication that the mortgaged property is a new-build. Second, I use UK 2011 Census data on average employment rates, educational attainment and socio-demographic status in the borrower's postcode as a proxy for demographic information that is not available on an individual basis, most importantly the borrower's education.

[10] For more detail about the PSD and variables it covers, see: www.fca.org.uk/firms/systems-reporting/gabriel/system-information/data-reference-guides/psd/psd001 , version 1.3.

Sample construction and representativeness

The construction of my sample starts with 1.3 million PSD mortgage transactions recorded for England and Wales between 1 January 2015 and 30 June 2016. Given the mandatory nature of regulatory reporting, these PSD data cover the universe of regional mortgage lending over those 18 months for the main three borrower groups: first-time buyers, home movers (existing mortgagors who are taking out mortgage to purchase a new property) and many remortgagors (borrowers refinancing loans on their existing property with a new lender).[11]

I restrict the sample to mortgages with LTV$\geq$20%, loan value$\geq$£20,000 and maturity of 5 or more years to avoid transactions that are likely to have low stakes and be unrepresentative of normal mortgage borrower behaviour. I also exclude mortgages that are part of government initiatives to support home ownership, and interest-only mortgages, where no principal repayment is required until the loan maturity date. Those types of loans cannot be included as they typically involve nuanced conditions that are not well captured in the data. Loans that satisfy these additional requirements represent the majority of all UK lending (87.5% of all transactions reported in England and Wales during the observation period).

I am able to match 64% of observations that satisfy the sample selection rules above to all three main additional datasets. I primarily use Moneyfacts, to obtain additional characteristics of the borrower's own mortgage and other mortgages that were available at the time the choice was made, credit bureau data for borrower creditworthiness and local area statistics, and Land Registry data to identify whether the property being purchased is a newly constructed building (new-build).

The comparison of summary statistics for the original population and the matched sample in Table A.1 suggests that distributions of most demographic and financial variables of interest are very similar. Any material sampling bias from the imperfect merges to the additional data appears highly unlikely. However, as there are some groups that are somewhat over-represented in the sample, for instance remortgagers (40.6% compared to 36% in the population) or those with fixed-rate loans (92.1% compared to 89.9% in the population), I also explore the robustness of my findings to reweighting the sample to correct for these divergences in Section 8.3.

The final sample used for all results in this paper unless otherwise stated contains c.695,000 mortgage transactions. This is a uniquely granular dataset that provides detailed information about the individuals and market circumstances under which they made mortgage decisions. A combination of PSD and Moneyfacts alone has only recently started being used for research (Benetton, 2017; Best et al., 2015; Cloyne et al., 2017), but this paper is the first to supplement those data with further extensive credit file information that offers a much better understanding of the individual risk of the borrowers.

[11] Due to the limitations of the mortgage transaction reporting data and of Moneyfacts data on products for existing mortgage clients, internal remortgagors, who refinance their mortgage with their current lender, are not included in this research.

Descriptive statistics

Demographic overview

Table A.2 provides summary statistics on borrowers' demographic and financial characteristics.

The sample is roughly evenly split between the three main mortgage borrower groups — first-time buyers, home movers and remortgagors. These groups have interesting differences in the complexity of their housing transaction that may be simultaneous with the mortgage choice. In order of increasing complexity:

- remortgagors typically do not have an accompanying housing transaction to worry about and can focus on the loan
- first-time buyers are choosing the mortgage alongside trying to secure a house and navigate the legal process for its purchase for the first time, and, finally,
- home movers would usually be trying to sell their existing property, as well as going through the new house purchase (with the associated legal processes)

As the whole sample, by definition, consists of households that qualified for a mortgage loan, borrowers' credit scores are high relative to the population overall, with the mean of 63.[12] Many borrowers are also similar in their creditworthiness: credit scores are within just 5 points from the mean in 50% of transactions. Lenders' transaction reports indicate mortgage borrowers that have a material credit impairment within the past 5 years, such as County Court Judgments (a type of court order that is registered in case of failure to repay a specific debt). This sub-group has a significantly lower mean credit score of 50, but its credit score distribution nonetheless has a large overlap with the rest of the population. This highlights the importance of factors beyond material past impairments in assessing their desirability as a mortgage borrower.

There is no information about borrowers' educational attainment in regulatory transaction reports or credit bureau data. I therefore use census data about the proportion of individuals in low skill occupations in the postcode of borrowers' residence to provide a rough proxy for borrower's level of education. The average proportion of low-skilled workers is 29%, but it ranges from 0 to 93% depending on location.

Product choice and borrowing costs

In line with the UK market as a whole, a predominant majority of borrowers in my sample, 92%, take out a fixed-rate mortgage. Nearly all of fixed rate mortgages have either 2-year (58%) or 5-year (32%) deal periods, after which the mortgage reverts to a form of variable rate. Only 2% of observed mortgage transactions involve loans with fixed rate periods of over 5 years. Around 15% of the variable rate mortgages also have a discounted rate deal period (typically for 2 years).

Table A.3 summarises the distribution of main components of borrowing cost for the borrowers in the sample. Over a third of borrowers do not pay up-front fees (consistently with frequent remortgaging in the UK), but fees are material when they are charged, averaging £930. There is also considerable variation in the interest rates. For instance, even though

[12] To avoid disclosing the identity of the credit bureau or any of their propitiatory information, I normalise credit scores to fall between 0 and 100.

the interquartile range for the initial interest rate is around one percentage point, this is equivalent to almost £2k difference in annual payments for the mean loan size in the sample (£180k). The difference between the bottom 25% and the top 25% by initial annual mortgage cost is over £6,500 per annum. An average household in the sample faces annual mortgage costs of nearly £12k a year during their deal period. Mortgage costs are thus a big part of household budget even in the current low interest rate environment, accounting for a quarter of post-tax household income.

Instead of comparing initial interest rates, it is more useful to aggregate the different components of the price into a standard metric for summarising multidimensional borrowing costs — the Annual Percentage Rate (APR).[13]

I calculate APRs on two bases in this paper. The first is the deal period APR, which is a rate that covers only the fixed (or discounted) period, implicitly assuming that the borrower will remortgage once this period expires. This is a plausible assumption in the UK market as according to FCA (2018) calculations, as only around 20% of borrowers stay on their reversion rate for more than one year after the period expires. However, deal periods vary, and APRs calculated over different time periods are not strictly comparable. I therefore also use 5-year APR as an alternative cost summary with a consistent time period for all borrowers. For borrowers whose deal periods are shorter than 5 years, I assume that they stay on their reversion rate for the remaining years. While it is a less plausible baseline description for the market, the 5-year approach also has an added benefit of reflecting some of the potential effect of reversion rates after the introduction period expires for borrowers who do not remortgage immediately.

Under both of these methods, the variation in APR within the sample is high. For instance, the difference between being at the 75th percentile and the 25th percentile of APRs is 0.97 percentage points (c.£1,750 per year for an average loan) on the deal period basis and 0.6 percentage points (c.£1,080 per year) on a 5-year basis. This interquartile difference is also highly economically significant, representing 2.5-4% of post-tax household income on average.

Not all of this price dispersion is due to search frictions, however. At least some is likely to be attributable to product differentiation and borrower risk. First, average APRs vary considerably depending on the features the product has. For instance, Table A.1a shows that the cost of a mortgage, at least in terms of deal period APR, increases with the length of time over which the interest rate is fixed. This is not surprising, as bearing the risk from potential monetary policy variation has additional costs for lenders, but nonetheless a risk averse borrower way be willing to pay lenders an additional premium to insure against rate variation. Similarly, borrowers may be willing to pay extra for flexible features of the mortgage or, for instance, reducing up-front fees if they are already close to their liquidity constraint due to expenses associated with a house purchase and move.

Differences in borrowers' credit risk are another reason why borrowing costs might vary even without any frictions in the market. Loan-to-value ratio (LTV), for instance, is a big determinant of lender's expected costs arising from credit risk. Most obviously, the LTV de-

[13] In line with standard practice and Financial Conduct Authority rules, the APR for borrower i with mortgage j over a period of T months a rate that sets the present value of all mortgage cash flows to zero, i.e. the solution to:

$$L_i = P_{ij0} + \sum_{t=1}^{T} \frac{P_{ijt}}{(1 + APR_{ijT})^{\frac{t}{12}}} + \frac{B_{ijT}}{(1 + APR_{ijT})^{\frac{T}{12}}}$$

where P_{ijt} is the total payment contractually required by mortgage j in month t, comprising any fees, interest on the outstanding balance and the required proportion of capital repayment, P_{ij0} is the total up-front fee, L_i is i's loan amount and B_{ijT} is the balance outstanding on i's mortgage at after T months if product j is chosen.

termines how shielded the lender is from loss in case the borrower defaults by the property they can repossess as collateral. With an LTV of 50% (and in the absence of transaction costs), for example, property prices need to halve for the lender to be unable to cover their loan advance with the repossessed collateral, whereas with LTV of 100% any fall in property prices could expose the lender to a loss. Furthermore, in line with literature on screening (Bester, 1985; Freixas and Rochet, 2008), loans with lower LTV are likely to attract borrowers with a lower probability of defaulting in the first place. Figure A.1b shows the borrowing cost variation across the quartiles of LTV distribution for the same product type (2-year fixed rate). As predicted by theory, the average borrowing costs are significantly higher at higher LTV ratios.

4 Introduction to dominance analysis

To summarise, borrowing costs vary a lot between households in the sample. But at least a part of this variation may be due to heterogeneous preferences for mortgage types, features, and the timing of payments. Borrowers qualifying only for a subset of observed loans is also a contributing factor. Understanding whether variation in costs arises from these two legitimate confounders rather than market frictions that hamper searching or choices is of considerable policy interest. It is also of academic interest as mortgage choices are an high-stakes environment with large implications for household budget. Therefore, there are very strong incentives for households to invest effort in this decision.

Because of the heterogeneity of the borrowers and the products, APR does not provide a sufficient metric for ranking products as it implicitly assumes a constant discount rate or no constraints on intertemporal budget transfers. Neither do APRs, by construction, take into account situation-contingent charges such as the early repayment penalty.

Instead of this, I develop a new 'dominance-based' approach to identify cases where households incur avoidable extra borrowing costs that cannot be fully explained by either the individual's limited eligibility or preferences for product features or timing of mortgage expenses. This allows me to explore the extent to which the observed variation in borrowing costs arises from market frictions such as search costs and begin to investigate what market or borrower characteristics might be at play.

The methodology developed in this paper builds on existing research that consider dominated choices in other household finance contexts, such as Sinaiko and Hirth (2011) and Bhargava, Loewenstein, and Sydnor (2017) for insurance and Elton, Gruber, and Busse (2004) and Egan (2016) for investment. One of the main novel contributions of this paper is that combination of granular data on both borrowers and products and the detailed individual-specific product ranking algorithm allows me to identify strictly dominated choices for all of conventional retail UK mortgage lending rather than in specific contexts. This approach to identifying dominated transactions is also flexible enough to handle heterogeneity in product characteristics, lenders' eligibility requirements, as well as borrower characteristics and preferences, which are all important features of the UK mortgage market and other household finance settings. Because this approach does not require evaluating trade-offs between product features or price components, it is particularly well-suited to cases with multidimensional pricing and complex product design.

Identification of dominated choices

Since the approach used in this paper is not standard, I first define the key concepts:

Definition 1 (Available product set) *Let C_i be the set of products available to a borrower i who applies for a mortgage at time t_i. Then for any mortgage product j, $j \in C_i \iff$*

1. *j was offered by the lender at t_i and was available for at least a month in total; and*
2. *i satisfies all of j's eligibility criteria, $e_i \in [min_e^j, max_e^j] \;\; \forall e \in E$.*

The eligibility criteria come from two sources: explicitly stated criteria in Moneyfacts (eg minimum and maximum age, property value or loan-to-value ratio) and additional rules

derived lender behaviour in the data (minimum credit score, and maximum loan-to-income and LTV ratios accepted for the given product).

The minimum time on the market requirement is added to avoid potential distortions from lenders' flash deals of loss-leader products, i.e. a heavily discounted mortgage offer that is offered only briefly or in limited numbers but allows the lender to advertise very attractive rates in a concurrent marketing campaign. Given the imperfections in the data, it would be possible (or even likely) that such a product is not available to the consumer making a choice at exactly the 'right time' even if they meet the eligibility criteria on paper.

Definition 2 (Strict comparability) *Let, $\vec{f_j}$ be the vector of non-price features of product j. Then, without any assumptions about borrower preferences beyond non-satiation, two mortgage products j and k are* strictly comparable *if and only if they have the same non-price product features, $\vec{f_j} = \vec{f_k}$,*

Definition 3 (Comparable set of available products) *Let $\tilde{C}_{ij}$ be the comparable set of alternatives to product j for borrower i. For any mortgage k, $k \in \tilde{C}_{ij} \iff$*

1. *k is available to borrower i, $k \in C_i$, and*
2. *k is strictly comparable to j, $\vec{f_j} = \vec{f_k}$.*

I restrict effective choice sets to only comparable product to ensure that mortgage products can be ranked for each borrower without any assumptions about the form of their utility function. Otherwise, the borrower's risk or time preferences, or the degree of uncertainty they face, could mean that they might be willing to forgo the lowest cost in present value terms to get (or avoid) certain product features. Taking borrowers' revealed preferences over features as given and restricting comparisons to mortgages that are substantively the same on main dimensions avoids this problem, albeit at the cost of also making it impossible to assess whether borrowers select product features that are suitable for their preferences.

The available data allow me to hold the following non-price product features constant: mortgage rate type (fixed or variable, and duration of the deal period), whether the mortgage allows repaying the principal faster than the contractual monthly payments even if with some restrictions (early repayments), whether the mortgage allows underpayments or payment holidays (typically subject to earlier overpayments). This list covers all product features that are explicitly listed in regulatory rules as affecting whether or not a particular mortgage product is appropriate to the borrower's needs and circumstances.[14] This level of detail also goes beyond what is normally observed by researchers in mortgage household finance literature, which — often for the reasons of data availability — tends to focus on borrower choice between fixed and variable contracts (eg Campbell and Cocco, 2003; Coulibaly and Li, 2009; Van Hemert, 2010).

Definition 4 (Dominance) *Mortgage $j \in C_i$ with a vector of price elements $\vec{p_j} \in \mathbb{R}^D$, is dominated by mortgage k for borrower $i \iff$*

1. *k is a comparable alternative to j, $k \in \tilde{C}_{ij}$, and*
2. *all elements of cost of j are, individually, at least as high as for k, and strictly higher for at least one price element, $\vec{p_j} \geq \vec{p_k}$ and $p_{jd} \neq p_{kd}$ $\exists d = 1, ..., D$.*

[14] The rule in question is MCOB 4.7A.6. of the FCA Handbook. The list provided in this section of the Handbook also contains several considerations regarding borrower's eligibility for the mortgage, which are covered by Definition 1, which identifies the set of available contracts.

I control for the price elements individually: initial interest rate, interest rate after teaser expiry (reversion rate), up-front fee (adjusted for cashback and other promotional discounts), fee at normal termination of the mortgage, and penalty charge for early repayment.

The dominance ranking between two products is defined only with reference to the specific borrower i for two reasons. First, for any comparison to be meaningful, it needs to be feasible for the borrower to take out either of the loans. This requires checking the specific combination of borrower characteristics against eligibility criteria for the two products. Second, some mortgages involve fixed fees and in others fees are expressed as a percentage on loan amount, so it is possible for the product ranking on fees to flip depending on the size of the mortgage. Calculating and ranking fees for each borrower for their observed loan amount solves this problem.

Definition 5 (Dominated choice) *Let $\mathcal{D}_i$ be the set of all mortgages that dominate the mortgage chosen by borrower i. Then, i's choice is* dominated *if and only if $\mathcal{D}_i \neq \varnothing$.*

I apply these concepts to assess the quality of borrowers' choices in this paper by going through the following steps for each borrower i in the sample. First, I identify the set of all products i is eligible for and which have the same features as i's chosen mortgage. Second, I calculate costs of each mortgage in C_i for i. Finally, I compare the costs of i's chosen product with every product on each price dimension separately to identify products that satisfy the conditions for dominance. This process is computationally intensive when, as in mortgages, there are many dimensions of product heterogeneity. It does, however, produce a conservative filter for product choices that are difficult to rationalise in a completely non-parametric way, which requires no assumptions about the utility function of the individual borrowers.

Cost of dominated choices

Measuring the cost of dominated product choices requires assumptions that go beyond the non-parametric method of purely detecting dominance. In particular, to calculate the cost of a mortgage one needs to, at the very least, decide on a time period for comparison, decide weights assigned to costs incurred in different time periods and, finally, make some assumptions about borrower behaviour over time (eg time of refinancing and whether any penalties are triggered). Comparisons of outcomes for different borrower types further require a consistent approach to be used for borrowers.

I calculate and compare total mortgage costs and the excess costs as a result of dominated choices by building on the standard cost metric for loans which were discussed in the preceding section, Annual Percentage Rates (APRs). As before, I calculate excess costs over two time horizons: the duration of the mortgage deal period (which abstracts from any additional costs borrowers might incur if they do not refinance their loans at the end of their incentive rate period) and a '5-year basis', which keeps the time horizon fixed across all borrowers at 5 years — the longest widespread deal period length. I revisit the sensitivity of the results to the selected comparison periods in the Robustness section.

For either of these cost measures, the annual mortgage cost over a period of T months for borrower i with mortgage j is simply $M_{ijT} = L_i \times (APR_{ijT})$, where L_i is i's loan amount and APR_{ijT} is the annual percentage rate calculated on the basis of i's loan size and the price structure of product j over T months using the standard formula in footnote 13.

Calculating the excess cost from a dominated product choice requires accounting for two additional complications: slower repayment of the underlying loan due to increased costs (which is not captured by the APR) and the possibility of multiple products with costs that are strictly superior to one's chosen mortgage.

Definition 6 (Excess mortgage cost) *Let $\mathcal{D}_i$ be the set of all mortgages that dominate the mortgage j taken out by borrower i. Also let B_{ijT} be the outstanding loan balance i still needs to repay after T months on a mortgage product j. Then, the excess cost i incurs per year as a result of making a dominated mortgage choice is the average of the differences in annual mortgage costs between the dominated product and all of thedominating alternatives in D_i:*

$$\Delta M_{iT} = \frac{1}{|D_i|} \sum_{k \in D_i} L_i(APR_{ijT} - APR_{ikT}) + \frac{12}{T}(B_{ijT} - B_{ikT})$$

This cost can also be expressed as a percentage reduction in actual mortgage costs the borrower forgoes due to their dominated product choice:

$$\%\Delta M_{iT} = \frac{\Delta M_{iT}}{M_{ijT}} * 100$$

The first term of the cost differential between two products is the higher interest payments captured by the APR difference. The second term is the adjustment for the difference in the remaining loan balance outstanding at the end of the comparison term. This difference arises because standard mortgage payments on the cheaper loan will include more principal repayment and thus yield a lower balance at the end of any given time horizon (other than the end of contract). This difference in balances needs to be compensated; otherwise the borrower would continue to incur 'excess cost' from the dominated choice relative to a superior alternative even after remortgaging as they would be paying interest of a larger loan than they would have had if they took out a superior mortgage.[15]

I calculate the excess cost of a dominated choice as an average across all dominating alternatives to ensure additional robustness. Analysis based purely on the 'best available mortgage' benchmark would make the estimated costs very sensitive to changes in the 'frontier' products over time and across groups, and might be less representative of the overall expected cost of not searching efficiently. Furthermore, comparing choices to the best possible alternative is less useful from the policy perspective, as even under well-functioning search it will often be unrealistic to expect that all borrowers choose the best possible mortgage product.

The definition of the excess cost from dominated choices is a financial measure, and it does not take account of borrowers' time preferences. Remaining agnostic about individual borrowers' (potentially heterogeneous) time and risk preferences is, after all, one of the main reasons for measuring search effectiveness using strict dominance between products rather than ranking them on a single aggregated price metric. However, the qualitative findings on the scale of the excess costs would not be materially affected by alternative time preference assumptions because of the short time periods for comparison. For instance, with the annual discount factor of 0.98 commonly used in the literature (see, for instance, Blundell et al. (2016)) would produce a 5-year cost only 5% less than the original

[15] To simplify the computational burden of calculating annual excess costs, I divide the total additional balance repayment needed equally between years in the comparison period.

approach. Most importantly, because dominance checks for product superiority (or at least equivalence) on all price dimensions individually and for each borrower, there is no set of time preference assumptions, even heterogeneous and time-varying, that would make the cost of holding a dominated product negative.

Finally, I combine the definitions of dominance and excess cost to specify a sub-group of consumers whose dominated choices clearly result in non-trivial costs to filter out marginally dominated choices that could be partly driven by measurement error and explore factors that might affect consumers experiencing material losses.

Definition 7 (Strongly dominated choice) *Let $\mathcal{D}_i$ be the set of all mortgages that dominate the mortgage j which borrower i takes out in the data. Then, i's mortgage choice is* strongly dominated $\iff$ *:*

1. *i's mortgage product is dominated by available alternatives, $\mathcal{D}_i \neq \varnothing$; and*
2. *as a result, over their deal period i incurs annual excess costs of at least:*
   - *£250 in absolute terms $(\Delta M_{iT} \geq £250)$ and*
   - *5% of their annual mortgage payment under product j $(\%\Delta M_{iT} \geq 5)$.*

As shown in Section 8 on robustness checks, the qualitative results in the rest of this paper are robust to using alternative excess cost cut-offs to define 'strong' dominance.

Headline results on frequency and costs of dominated mortgage choices

Applying the definitions of dominance above to the data reveals that 30% of mortgage borrowers chose products that were dominated by another available alternative. These choices have material economic consequences: buying a dominated product led borrowers to forgo mean savings of approximately £550 per year over the deal period (just over 3 years on average), equivalent to reducing mortgage costs by 12.7% or to losing 1.5% of their household income. The excess cost of dominated choices is only slightly lower at 10% of annual mortgage costs if calculated on a standardised 5-year basis, as shown in Table A.4. The lower annual cost is likely at least in part driven by the additional up-front fees paid on the dominated product being spread over a larger number of years.

Dominated choices have significant economic consequences for a large proportion of households. Mortgage products selected by 17.5% of the observed borrowers meet the criteria for strongly dominated choice given above and thus lead to considerably higher mortgage costs than were possible under available superior alternatives.

Figures A.2 and A.3 elaborate on the distribution of annual excess costs per year incurred as a result of dominated choices. Households that forgo relatively small amounts when they choose a dominated mortgage are quite common. Dominated choices 'cost' about a half of the borrowers £350 or less per year, and for 13% of the sample annual costs are as low as £100. The distribution also has a thick right tail, however, and large losses are also quite frequent. Around a third of borrowers with dominated products are foregoing the opportunity to reduce their mortgage costs by 10% or more, and a non-trivial minority of 15% face avoidable excess cost of over £1,000 per year over a period of 2 to 5 years.

Given that the total amounts 'left on the table' are in the hundreds of pounds even for households with relatively small excess costs, it appears unlikely that the dominated choices are fully rationalised by preferences over intangible product attributes such as the lender's brand. Some form of market friction appears to be needed to rationalise the large amounts of money so systematically forgone. For instance, the complexity of researching and assessing costs of different mortgage options, combined with low financial sophistication of some of the borrowers could result in prohibitively high search costs that result in narrow sets of products that are considered.

While there is clear and potentially interesting demographic variation, it appears likely that it occurs on many demographic and behavioural dimensions at once and therefore requires a more systematic regression approach before influences of individual factors can be meaningfully discussed.

5 Demographic variation

Univariate trends

The likelihood of a household choosing a dominated mortgage product, and the costs of doing so, vary with many standard demographic characteristics. While dominated choices are present to some degree across all types of borrowers, demographics appear to play a role in shaping propensity towards such choices.

Figure A.4 illustrates high-level univariate trends in dominance rates and costs with the three core borrower characteristics: credit score, income and age. The dominance approach already controls for these variables for all three these variables in assessing each borrower against product eligibility criteria and identifying whether their chosen mortgage is dominated by any other available options. Yet, there is clear residual variation in both the likelihood of making a dominated choice and its cost across all three of these factors. For instance, borrowers in the lowest score band (35-45) are nearly 10 percentage points more likely to make a dominated choice than those with the highest credit scores (75-85), and their excess cost as proportion of the overall mortgage payment is nearly twice as large.[16] The borrower's age and income are also associated with varying rates: for both, the difference in average dominance rates between the top and the bottom band is around 3 percentage points. Also notably, the variation in probability of dominated choice and costs follows a relatively common hump-shaped pattern in the quality of financial decision-making (see, eg Finke, Howe, and Huston (2017) for an overview), with middle-aged individuals performing best of all.

Overall, however, while there is clear and potentially interesting demographic variation, it appears likely that it occurs on many demographic and behavioural dimensions at once and therefore requires a more systematic regression approach before influences of individual factors can be meaningfully discussed.

Regression specification

I investigate how the dominated choice probabilities and costs vary with borrower demographic characteristics and financial circumstances using the following general model:

$$Y_i = X_i\beta + N_i\xi + Z_i\delta$$

where $Y_i = \log \frac{\pi_i}{1-\pi_i}$, is a transformation of π_i, is the probability that borrower i's chosen product is *strongly* dominated, or $Y_i = \%\Delta M_{i,Deal}$, the excess cost from a strongly dominated choice over the deal period expressed as a percentage of the borrower's total mortgage cost.[17]

[16] Individuals with the worst credit ratings (30 and below) have somewhat lower dominated choice probabilities. This is likely because individuals with such poor credit typically have much fewer mortgage options available to them, and so the likelihood of there being another strictly superior alternative to any given product is lower.

[17] Regressions focus on strongly dominated choices to minimise effects of small measurement errors on classification of transactions as dominated, and also to explore demographic effects on cases with large costs and greater economic significance. Qualitative findings are robust to alternative specifications of strong dominance. See section 8.4

On the right-hand side, X_i is a vector of i's demographic and credit risk characteristics, including to household income, loan amount, credit score, borrower type, employment status, number of children, unsecured debt volume, etc, as well as some quadratic forms of continuous demographic variables (see the full list of demographic covariates in Table A.8), N_i is a quadratic polynomial of the total number of products in the choice set C_i, Z_i is a vector of indicator variables for search channel (and, in some specifications, region of house purchase). β, ξ and δ are vectors of parameters, which are estimated with MLE for specifications that focus on the probability of dominated choice and OLS for the excess cost specifications. I calculate cluster-adjusted standard errors using the approach in Cameron, Gelbach, and Miller (2011) to account for possible geographical and temporal correlation between behaviour of households (eg due to regional advertising campaign by a specific lender). Clusters are based on a combination of year and quarter of transaction with postcode area of the mortgage property, producing 624 groups in total.

Fixed effects in Z_i include indicators for whether the transaction was advised and also for the borrower searching lenders directly or through an intermediary. These factors are important to control for since borrowers are highly likely to select into search channels depending on characteristics related to other variables in the model, such as complexity of their credit situation, their financial capability or income. The potential for selection bias also means that the estimates of advice or intermediation fixed effects in this model are very unlikely to have a meaningful interpretation. Because of this, I restrict my attention to demographic covariates in this section, and Iscenko and Nieboer (2018), a companion paper, focuses on identifying meaningful causal effects of mortgage advice and intermediation on dominance rates (as well as other borrower outcomes) by applying DiD matching estimation to the recent UK policy changes. The results reported later in the paper are also robust to using more detailed channel fixed effects, with additional fixed effects for the size and type of the intermediary firm (if any) used by the borrower.

Marginal demographic effects

To simplify interpretation of quadratic terms and logistic regression coefficients, Tables A.6 and A.7 present average marginal effects obtained from regressions above, for the probability and excess cost. The tables also report marginal effects which are scaled by one standard deviation of the explanatory variable to illustrate relative economic significance.[18]

Overall, the estimated demographic marginal effects are consistent with the idea of heterogeneous search costs changing the likelihood that a borrower discovers (or selects) better alternatives during their search. There are broadly two channels through which the variables in the model can affect search costs: the borrower's ability to search (eg financial capability and available time) and the complexity of choice they are facing (eg number of options and application complications).

Borrower ability to search effectively

Credit history The process of identifying dominated choices explicitly ensures that the borrower exceeds the minimum credit scores that have been accepted for a potentially dominating alternative in a similar LTV band. Yet, regression results show that after controlling for a range of demographics, borrowers with high credit scores are still better at

[18] Regression results before transformations are provided in Table A.8.

avoiding dominated choices and pay less if they make one. One possible mechanism behind this effect could be financial capability: it is likely that borrowers who are good at managing their money will both have better credit histories and make better high-stakes financial decisions. A borrower with lower capability is likely to have more difficulty identifying relevant products and comparing their costs, which in effect could be thought of as higher (perhaps sometimes prohibitive) search costs. Figure A.5 provides additional detail on the variation in the estimated average marginal effect of credit score on the probability of making a strongly dominated at different credit score levels. The estimated marginal effect is most beneficial for borrowers with the worst credit histories, and grows smaller as credit score improves. This pattern is consistent with borrowers who are worst at managing their finances in general having most to gain from slightly better financial sophistication and ability to search for products. The ratio of unsecured debt to income — another potential proxy for ability to manage one's finances – has very similar effects to credit scores, with higher levels of debt being associated with higher likelihood of dominated choices and larger excess cost.

Credit history could also influence the effectiveness of borrowers' search through the second channel mentioned above — complexity of the mortgage choice (conditional on any level of sophistication). Lenders' minimum acceptable credit scores and tolerance for minor credit history 'blemishes' is often not public or can be searched in standardised way, unlike many other eligibility criteria. This can mean that (despite satisfying the criteria in the data), households with less-than-perfect credit scores may be more uncertain about being accepted for the best mortgages or at the very least may need to spend more time and effort investigating their eligibility for any given mortgages. I explore the potential borrower uncertainty about some of the eligibility standards further in section 6.2.

Income Household income continues to have a strong negative association with the likelihood of a strongly dominated choice after controlling for other demographic factors. The average marginal effect of income is, in fact, considerably larger than univariate trends might suggest, with a £1,000 increase in household income being associated with 0.2pp reduction in dominance rate on average. As Figure A.5 shows, however, the size of this effect varies a lot with the level of household income. Among the poorest households in the sample with post-tax incomes around £25k, an additional thousand in household income reduces the likelihood of a dominated choice by 0.4 percentage points or more. In fact the marginal effect of increasing household income from £30k to £40k alone lowers the risk of making a strongly dominated choice by around 15%. However, the beneficial marginal effect of income on the quality of product choices declines sharply as income increases and disappears entirely when household income reaches £100k or more.

Income should not have a direct effect on probability of dominated choices as it is explicitly considered when identifying sets for which of products the borrower is eligible. There is, however, extensive evidence that household income and wealth are closely correlated with education and financial capability (summarised, for example, in Lusardi and Mitchell, 2014; Calvet, Campbell, and Sodini, 2009). More affluent households are thus more likely to have the numerical skills and experience to compare mortgage products effectively, resulting in lower search costs other things equal. This explanation is also consistent with the pattern of change in the estimated effects with the level of income. For low income households, additional earnings are likely to be associated with more years of education or a more technical occupation. High-earning households, however, are already likely to have

graduate-level education and be employed in a quantitative industry. In this case, a small additional increase in income is unlikely to proxy for a change in factors that are relevant for better financial decisions.

There may also be a counter-veiling effect of income on search costs, as the opportunity cost of time is often linked to earnings in the context of consumer search (Marmorstein, Grewal, and Fishe, 1992) and the broader literature starting with, for example, Becker (1965). The higher cost of time for high-earners could be a contributing factor to the declining beneficial effects of income on the likelihood of a strongly dominated choice and also slight detrimental effects of income on the size of the excess cost from dominated mortgages (see Table A.7).

Age Age is another factor that is commonly linked to financial sophistication, and found to have an effect on the quality of financial decisions more generally (Agarwal et al., 2009; Lusardi and Mitchell, 2014). The original research by Finke, Howe, and Huston (2017), as well as their overview of the other literature on the topic, suggest that both financial capability test scores and observed quality of financial decisions tend to peak between the age of 45 and 55, and then decline sharply into old age. Broadly in line with earlier findings, age has a small positive average marginal effect of increasing age on the probability of a strongly dominated choice and the size of the resulting loss. The variation in the marginal effect with main borrower age in A.5 reveals that the relationship between average quality of decisions and age follows the standard inverse U-shape. Increasing age having a positive effect for younger borrowers, but turning progressively detrimental later in life. One distinction prior studies is that, other things equal, ability to avoid inferior mortgage products peaks around age 30, earlier than estimates in the existing literature. This could be because the model already includes some controls the quality of financial decisions such as credit scores, which would also capture some of the age-related changes financial ability to the extent that it affects credit use.

Borrower type The estimated differences in dominated choice probability between borrower types are large and have directions that are consistent with the theory. Remortgagors have the lowest probability of making dominated choices, followed by first-time buyers whose probability is more than 3 percentage points higher, and then borrowers moving house who, with another 3.5 percentage point increase in dominance rates, are nearly 50% more likely to choose an inferior product than remortgagors.

One potential explanation for this pattern is the difference between the three groups in the extent to which their time is likely to be occupied by a concurrent housing transaction instead. Home movers are very likely to be both buying and selling a property at the same time on top of needing to make a mortgage choice. Involvement in these demanding and high-stakes housing transactions is likely to significantly increase the opportunity cost of shopping around for a mortgage instead, and thus reduce search. First-time buyers are only dealing with one housing transaction while choosing a mortgage, and so are likely to have more time to think about the mortgage relative to movers, which could account for better dominance rates. Finally, remortgagors typically have no ’distractions’ from housing, and are therefore likely to have the most time to dedicate to mortgage search.

It may be initially surprising that first-time buyers — households who by definition have the least experience with mortgage products — do not perform worse than movers. It is

likely that selection effects also contribute to this finding. The UK housing market in the past years has been extremely challenging for first-time buyers, as the house price growth and stagnation in real incomes make it very difficult for those not already on the housing ladder to save enough for a deposit. It seems plausible that households who have managed to purchase their first property in these difficult conditions are considerably better than average in terms of their financial nous, and also may be able only make the purchase if they keep costs down as much as possible. The only first-time buyers in my 2015/16 sample are households who satisfy these requirements. Moving and remortgaging, in contrast, are easier and open to a wider range of households regardless of financial competence since they can rely on the realised equity in their existing property.

Other Additional factors that can contribute to ability to search are education and demands other than the concurrent housing transaction that affect the borrower's free time. Regression coefficients on relevant variables also appear plausible. First, the detrimental marginal effects of the number of children on both probability of strongly dominated choice and the resulting excess cost may be another example of alternative demands on the borrowers' time increasing the opportunity cost of mortgage search. Second, the borrower being located in an area with higher proportion of residents in low-skill occupations according to the UK Census — an indirect proxy for borrower's own education and thus financial capability to shop around effectively — is also associated with slightly higher dominance rates.

Complexity of the mortgage choice

The regression results also suggest that factors that increase the complexity of the choice (for a borrower of any given capability) have expected effects.

The number of available products has a small estimated marginal effect for one extra product in the choice set. However, increasing the number of products by one standard deviation has a more material effect. This is, in part, a reflection of the large variation of 'complete' choice sets. Regression estimates suggest that the effects decline very rapidly after choice set size reaches 10-15. Once there are dozens of products to wade through, an extra 10 does not matter much as the borrower is unlikely to consider all of them.

There are also large estimated effects for some of the standard factors that can increase the complexity of application and uncertainty about acceptance (even if on paper the loan satisfies the requirements). Those characteristics include being self-employed, buying a newly built home or a property that is otherwise unusual (as reflected by it being sold at a discount relative to house prices in the area). While the dominance calculation ensures that borrowers satisfy the standardised posted criteria, there can be additional uncertainty about banks' idiosyncratic treatment of unusual cases, which could make search and comparing lenders more complicated.[19]

[19] As discussed in section 8.1, the general findings on dominance and excess cost are robust to excluding borrowers in these unusual circumstances.

6 Dominating products

In this section I explore the possible role of different product features in dominated choices. I start with the different components of the total cost of the mortgage and later consider the differences between dominated and dominating products in their explicit and implied product eligibility requirements.

Salience of price features

A hypothesis often seen in behavioural literature is that consumers often focus on a small subset of all relevant factors when making complex decisions. For instance, rather than calculating total costs that balance all elements, households could focus on most immediate or prominent elements of the price and choose products that are cheapest on that basis, even if this product also has large 'shrouded' costs (Bar-Gill, 2012; Bordalo, Gennaioli, and Shleifer, 2016; Gabaix and Laibson, 2006). In principle, mortgage markets, including in the UK, are a candidate for such errors as well since advertising and consumers' own discussions often focus on the introductory interest rate (Gurun, Matvos, and Seru, 2016). If true, this behaviour could result in high observed dominance rate as households that focus on the introductory rate alone might be in effect indifferent between two mortgages that have the same introductory rate but might differ a lot on other dimensions, and might pick the more expensive one as long as the headline rate was the best they've seen. I test whether this holds in practice by calculating the proportion of dominated choices where the headline where the household chose a mortgage with a strictly higher introductory interest rate rather than minimising that dimension as salience ideas might predict.

As I show in Figure A.6, far from being the only dimension that borrowers minimise effectively, the introductory rate is in fact the price element where a strictly positive difference with the dominating product is most common. Around 92% borrowers with dominated mortgages ended up with a strictly higher introductory rate relative to better available alternatives. In contrast, buyers of dominated mortgages pay strictly higher up-front fees in only 55% of the cases. Moreover, the differential between the initial rates on the dominated product and better alternatives on average accounts for over a half of the total excess cost that buyers of dominated products incur. (I use costs incurred over 5 years here to also see the potential impacts of reversion rates). However, if I use the deal period basis only and assume borrowers switch immediately after it expires, the contribution of initial rates to total cost is even higher. Buyers of dominated products do not therefore appear distracted by minimising introductory rate at the cost of overlooking differences in other cost elements.

Product eligibility requirements

The stakes in applying for a mortgage are very high: a rejection or closer investigation by a lender can cause delays that may jeopardise the whole underlying housing transaction. Households therefore have strong incentives to avoid products they have a high risk of being rejected from. The same is even more true for an intermediary searching the market on borrowers' behalf, as they only receive commission from the lender if the deal goes

through, and do not share in any borrowing cost savings a household might get on a cheaper mortgage. In this section, I investigate the possibility that the observed dominated choices are at least in part driven by borrowers' risk aversion under uncertainty (or confusion) about the possibility of rejection by lenders offering superior alternatives. Unfortunately, there are no administrative UK data on rejected applications. Instead I explore how product eligibility criteria (both publicly posted and derived from the data on observed acceptances) differ for chosen dominated products and the dominating alternatives that were available to the borrower.

To do this, I calculate the difference in the minimum standards for the chosen product and each dominating alternative on a range of eligibility criteria: maximum loan amount, maximum term of the mortgage, maximum LTV ratio (posted by a lender and estimated from accepted borrowers with the product), maximum borrower age, maximum loan-to-income (LTI) ratio (both posted and estimated) and minimum credit score (estimated only) accepted for the product in the borrower's LTV band. A dominating product with a positive (negative) difference is classified as having stricter (easier) requirements. The requirement is defined as much stricter/easier, if the absolute value of the difference is 50% or more of the total 'headroom' available until breaching the minimum standard on their chosen (dominated) product. (This difference cannot be greater than 100% because the borrower meets all the minimum criteria for the 'dominating' products by definition.)

Figure A.7 presents the results of this comparison. The first surprising finding is that although the superior alternatives are strictly cheaper, they do not necessarily have more demanding eligibility requirements. In fact, the median difference between the minimum eligibility requirements on the chosen product and the 'dominating' alternative is zero for all criteria that are explicitly posted by lenders. Superior alternatives have stricter acceptance requirements for only around a quarter of cases for most of these explicit criteria, and in most cases the borrower is able to clear those stricter requirements with reasonable 'headroom'. This would suggest that borrowers are not trying to 'play it safe' when it comes to simple and explicitly advertised criteria such as LTV or maximum term of the mortgage, and in fact get quite close to the 'frontier' of products they are eligible for.

The situation is rather different for criteria that are inferred from the data on other accepted borrowers, and, to a lesser extent, for the explicit LTI eligibility criteria, which often tend to be multi-tiered depending on borrower circumstances (eg income). Around two thirds of dominating alternatives have stricter (estimated) acceptance standards for credit score and loan-to-income ratios and nearly a half do so for stated LTI maxima as well. Borrowers are not nearly as likely to select alternatives which they only just qualify for when it comes to these less transparent criteria compared to simpler standards described above.

This difference could plausibly contribute some of the observed dominated choices. Borrowers (or agents acting on their behalf) may expect (often correctly) that another strictly cheaper alternative involves somewhat higher lending standards, but be uncertain about the exact requirements and thus their own eligibility. Even though they do, in fact, qualify, borrowers may be unwilling to obtain more information by applying, given the housing transaction at stake.

7 Familiarity

Having outlined the borrower and product characteristics that are associated with dominated choices, I also consider the search process may be contributing to the observed outcomes. In particular, I focus on whether the strength of a pre-existing relationship with a lender plays a role in product choice, given that there is extensive evidence that household and professional choices are biased towards the familiar options in other contexts such as equity investments (Coval and Moskowitz, 1999; Huberman, 2001; Pool, Stoffman, and Yonker, 2012, and many others).

Intermediary search

A natural hypothesis for why dominated choices arise is that the set of mortgage products the borrower considers (consideration set) is smaller than the available choice set. Given the structure of the UK market, there are some ex ante reasons to think that consideration sets may be constrained by the way in which the borrower searches the market. A borrower can either search by themselves and apply directly to lenders, or use the services of an intermediary. There are many different firms of intermediaries, which vary greatly in their size and the nature of their commercial agreements with lenders. Some of the mortgage products may, for example, be available exclusively to a subset of intermediaries in the market. Some intermediaries may only cover products from a subset of existing lenders ('a panel'). Alternatively, a lender may only make their most attractive offers available to direct (non-intermediated) applicants to reduce commission costs.

Purchases made through intermediaries offer particularly useful grounds for exploring the possible links between dominated choices and the options actively considered as UK market structure offers a sequence of identifiable restrictions on intermediary's consideration set. In Table A.9, I show how the frequency of choices that are dominated by *alternatives within the a restricted consideration set* changes with these restrictions.[20]

The first feature of the UK market that can restrict products intermediaries consider in practice is that some products are only available to borrowers who go to the lender directly. Often, this is a product-specific restriction, for example a lender may accept business from intermediaries in general, but have a promotional product that is available only to direct applicants. An intermediary can, in principle, know about these direct-only deals and suggest that the borrower contact the lender directly. The intermediary cannot, however, help their client in arranging this application or earn any commission for the resulting sale, so it is highly likely that the direct-only deals are not considered by them closely or at all. Eliminating these products from the consideration set reduces the rate proportion of dominated intermediated transactions by around a third, from 30.8% to 21.3%.

The source of the second institutional restriction is that some large intermediaries have panels of preferred (or 'tried and tested') lenders, whom employees are recommended to use. While it is possible for an employee to send a customer to an off-panel lender, this often requires additional paperwork and is likely to be avoided other than in very unusual

[20] The difference between the baseline dominance rate in the consideration set that includes all available products the borrower qualifies for and the rate in a more restricted set gives the extent to which intermediaries' dominated choices are rationalised by limitations on consideration sets in practice.

circumstances. Data on panels are not public, but I have privately supplied panel data from 5 large intermediary 'networks', which jointly account for approximately a third of intermediated transactions. The proportion of transactions that are dominated by products from in-panel lenders is 23.5%, down from 30.1% unrestricted dominance rate for the same 5 intermediaries. Part of the reason for only a moderate reduction is that intermediaries' official panels are usually very broad to cover a range of borrower circumstances, and may not necessarily reflect lenders that are used a lot in practice.

A natural extension of the idea of panels is to investigate whether in most normal circumstances intermediaries tend to consider a smaller subset of lenders with whom they are most familiar. Repeated interactions allow intermediaries to learn about some of the uncertain eligibility requirements discussed in the previous section. They may believe that their existing relationship with some of the lenders may ensure a smoother application process for the customers. It is plausible that search costs and uncertainty are lower when using familiar lenders, and that intermediaries consider those in the first instance.

To proxy for familiarity, I calculate each lenders' share in business of any given intermediary. I then identify the most frequently used lenders which jointly account for 75 or 90% of each intermediary's mortgage transactions and restrict that intermediary's consideration set only to products from these preferred lenders. For most intermediaries, the subgroups of familiar lenders are small: the median number of lenders that make up 75% of intermediary mortgage transactions is 5, rising to 8 for 90% of transactions. Larger intermediary firms have more diverse sets of regular suppliers, but even those still remain around 10 on average and smaller than a typical panel.

As can be seen in Table A.9, restricting intermediary's consideration set only to lenders that supply 75% of their business has by far the largest effect on its own, nearly halving the rate of choices that are dominated within the consideration set to under 16%. When familiarity is combined with all other restrictions in this section, only 13% of intermediary transactions continue to be dominated by a product in the resulting consideration set compared to over 30% when using the whole available choice set. Almost 60% of dominated transactions through intermediaries can therefore be explained by plausible (and perhaps practically justified) restrictions on consideration sets that lead borrowers' agents to overlook the strictly superior alternatives.

Direct search

Next, I consider whether individual borrowers are drawn to lenders they already know, and the extent to which it might contribute to dominated choices. Mortgages are infrequent transactions for most households, but most mortgage lenders in the UK also offer other financial products which are more common, for instance, current accounts, personal loans and credit cards. Using the information in the credit files, it is possible to identify lenders with whom borrowing household already had one of these financial products at the time of their mortgage search. On average, around 13% of all mortgage products in a household's choice set came from lenders with an existing banking relationship.

Figure A.8 shows that the proportion of products from familiar lenders varies between whether the the borrower used an intermediary and whether they made a dominated choice. For each of the four sub-groups it also compares this figure to the proportion of borrowers who chose a familiar lender for their mortgage. The differences between direct and interme-

diated transactions are striking. Among households who were assisted by an intermediary, just over 20% ended up borrowing from a lender with whom they already had a financial product. While their probability of choosing a familiar lender is above the proportion of mortgage products from these lenders in their choice sets (10%), this could be because firms with good deals on other financial services might also offer more attractive mortgages and are therefore more likely to be both familiar to the borrower and to be selected out of a set of other mortgage products.

It seems highly unlikely that the same mechanism can explain the patterns for the borrowers who made their own choices, however. Both the dominated and non-dominated sub-groups of these borrowers had somewhat higher proportion of familiar lenders' products in their choice sets (14% and 19%, respectively) than in intermediated transactions, but they were overwhelmingly more likely to take out one of these familiar mortgages as a result. Over 55% of direct borrowers whose mortgages were not dominated selected a lender they knew, as did two thirds of borrowers with dominated products. Both groups were 3 times more likely to go to one of their existing banks for a mortgage than their intermediated counterparts. The underlying drivers that make borrowers so likely to choose their own banks are important to investigate in further work, but outside the scope of this paper. Overall, however, whoever is making decisions — the intermediary or the borrower themselves — seems to be particularly drawn to offers where there is some degree of familiarity from past experiences.

To test whether selecting a familiar bank also increases the likelihood of making a dominated choice, I re-run the model described in section 5.2 with an additional indicator variable that takes the value of 1 if the borrower's mortgage is from a familiar financial institution. The coefficient on this variable is highly statistically significant (Table A.13). Even after all the controls discussed above, selecting a familiar bank is associated with a 3.7 percentage point increase in the likelihood of making a strongly dominated choice, a large effect compared to the unconditional probability of 17.5%. Given the mean cost of a strongly dominated choice of £810, this suggests that, other things equal, a borrower staying with the familiar lender faces a 'familiarity premium' of £30 per year in expected additional excess costs.[21]

[21] For borrowers with dominated products, the size of their excess cost does not not materially different between those who used a lender they knew or a mortgage supplier without an existing relationship. This also assumes that all borrowers without dominated products have excess cost of zero.

8 Robustness

The findings reported above are robust to a wide range of alternative assumptions. In this section I summarise some of the key robustness checks: restricting the sample to very prime borrowers, the lower bound on costs of dominated choices, sample representativeness weights and alternative definitions of strong dominance. I also compare the headline findings on the rate of dominated choices with a benchmark of borrowers choosing products at random from their available choice set.

Prime borrowers only

The regression results that show large economic and statistical significance of some of the factors that complicate the mortgage application can raise the concern that the proposed model does not adequately deal with eligibility in non-routine circumstances. If true, this would mean that filtering products based on standard eligibility criteria may overestimate the range of products a borrower would in reality qualify for and over-estimate dominance rates as a result.

I explore whether this concern is likely to be material by investigating a restricted sub-sample of very low risk borrowers in routine circumstances that are highly unlikely to involve lender discretion on top of assessing applicants against the posted criteria. This 'ultra-prime' sub-sample contains borrowers who satisfy all of the following criteria: LTV$\leq$65%, loan-to-income ratio $\leq$2.5, mortgaged property that is not a new-build or sold at a discount that could indicate other irregularities, credit score in the top 25% of all borrowers, and the borrower is not self-employed or due to repay the loan after reaching retirement age.[22] This is a very conservative set of assumptions, which restricts the sample to approximately 63,000 observations.

Table A.10 compares the average dominance rates and costs incurred in this restricted prime sample to the full-sample results reported earlier. In general, although there is some improvement in most of the metrics among the vanilla sub-sample, it is very small. For instance, the proportion of borrowers who buy dominated products is still at 26.1%, a marginal decrease from the original 29.9%. The absolute value of excess borrowing costs incurred due to dominated choices appears more affected as it falls to approximately £284 from £550 on a deal period basis. However, part of the effect is due to a much smaller size of the loans in the ultra-prime sample due to the selection rule that caps LTV at 65%. In fact, excess costs relative to the annual mortgage payment are higher in among very prime borrowers, at 13.1% compared to 12.7% average for the whole sample. Overall, the fairly small effects of restricting the analysis to borrowers where lender discretion and heterogeneity in internal standards are extremely unlikely to play a role suggests that those factors are not a major underlying driver of the findings.

[22] There are no official categories of prime borrowers. However, the criteria above have been discussed with the FCA specialists on firms' business models, who confirmed that this was a plausible way of capturing borrowers at the top of the range of prime lending.

Cost specification

I undertake 2 checks to determine whether the choice of time horizons for calculations is causing me to overestimate the size and economic significance of excess costs from dominated choices (eg Table A.4).

I calculate average annual excess cost and cumulative excess cost separately under all time horizons from 1 to 25 years for each borrower with a dominated mortgage.[23] As before, the borrowers are assumed to incur no cost from their dominated mortgage in years after remortgaging. For each borrower, I obtain the minimised excess cost from the dominated choice under two approaches: (a) the average annual cost difference between the dominated and superior products over the time period that minimises the *cumulative* excess costs and (b) the smallest possible annual excess cost. In effect, the first measure is the cost of a dominated choice conditional on the household realising their error immediately and responding optimally to minimise its consequences. The second measure is simply a lower bound on the standard excess cost statistics reported in this paper.

Table A.11 summarises the results of this comparison. The most surprising finding is that the annual excess cost over the time horizon that minimises total costs of a dominated choice is actually higher than under the baseline assumptions — with the median of £350 and a mean of £708 per year, compared to £330 and £550 per year, respectively, on a deal period basis. This is because for many households the cost-minimising response to making a dominated choice is to absorb the early redemption penalty and remortgage very quickly, incurring a higher additional cost but over a shorter period. In fact, 1 year is the cost-minimising time on the dominated mortgage for over 56% of the sample. While interesting as a sensitivity check, this behaviour is clearly out of line with empirical facts as practically no consumers remortgage before their deal period expires (FCA, 2018), and the vast majority of borrowers have deal periods of 2 years and longer.

The sensitivity check using the time horizon that produces the smallest possible annual costs predictably reduces excess costs relative to the baseline assumptions. On that basis, the mean estimated annual cost falls to £229 and the median to £109 (see Figure A.10 for the full distribution). However, this scenario again involves implausible borrower behaviour as it implies that nearly all borrowers stay on their mortgage products for years or even decades after their introductory rate expires. For nearly a third of the sample it means not remortgaging for 25 years or more, in contrast with the relatively prompt responses to expiry of the short UK deal periods actually observed in the market (see section 2.1). Furthermore, these low annual costs are produced by behaviour that is the opposite minimisation of total excess costs which are what matters for household overall utility.

The sensitivity checks with more conservative cost measures do not undermine the earlier findings that a large number of households incur additional borrowing costs of hundreds of pounds (over multiple years) as a result of choosing dominated mortgage products.

[23]In cases where remortgaging after the given number of years (eg 1 year) triggers any contingent penalties for early repayment, the calculation takes into account the difference in any such penalties between dominated mortgage and the average for all dominating products.

Sample representativeness

As discussed in section 3.2, the composition of this paper's primary sample deviates from the population on a small number of dimensions: for instance, the proportion of loans with a fixed introductory rate, or of borrowers buying their first property. Since the probability of dominated choices, and costs incurred as a result vary with borrower demographics, I explore whether the sample findings are likely to apply in the overall population by re-weighting the sample to be closer to population distribution by using multivariate probability weights based on indicators for a fixed-rate loan and borrower type (first-time buyer and remortgager) and for quintiles of income, borrower age, LTV and loan amount.

Table A.12 presents the results of the re-weighting. The demographic means in the weighted sample are almost identical to those in the population, including in variables that previously diverged. The headline results on dominance and excess cost are very close in the un-weighted and weighted sample. The proportion of dominated choices declines slightly from 29.9% to 29.5% while the strong dominance rate increases by 0.2 percentage points to 17.5%. Excess costs are also very similar. For instance, the annual excess cost as a % of the borrower's mortgage payment is 12.4% before re-weighting and 12.7% after. Correcting for divergences in demographics results in only small changes but has non-trivial computational costs, so the baseline findings in the paper are presented without re-weighting.

Definition of strong dominance

Regression specified in section 5.2 uses the indicator for borrower's choice being strongly dominated (excess cost above £250 and 5% of annual mortgage cost) as an outcome variable. To demonstrate that the qualitative results of interest are robust to other definitions of strongly dominated choice, Table A.13 shows regression results under several alternative cut-offs: any dominated choice, excess cost above £250 or 5%, excess cost above £250 or 10%, and excess cost above £500 or 10%. The signs, significance and scale of coefficients for the main variables of interest discussed in section 5.3 and 7 are similar across a range of alternative definitions. The model with any dominated choice as the dependent variable has the most deviations from baseline and lower significance of some of the covariates. This is surprising, however, as the primary motivation for using strong dominance is the concern about the possibility of noise introduced into some of the 'lightly' dominated choices by possible measurement error.

Are choices better than random?

It could be easy to interpret the earlier sections as focusing too much on the flaws and not recognising the extent of search that does carried out effectively. I check this by investigating whether borrowers would do materially worse than their observed outcomes if they had just picked a mortgage at random out of products they qualified for.

I go through their comparable set of available products (as defined in subsection 4.1) for each borrower and identify all alternatives, whether chosen or now, that are dominated for that particular borrower. I then use the ratio of dominated products to the total number

of alternatives in the individual's comparable choice set as the probability of this borrower making a dominated choice if their choices were random.[24] The mean probability of making a dominated choice at random in the sample is approximately 70%, a lot higher than the realised rate of dominated choices of 30%.

Figure A.9 elaborates on this by comparing probabilities and realised rates of dominated choices in small consumer clusters defined by permutations of the range of demographic factors (represented by the individual grey dots). If borrowers were choosing as well as random, one would expect probabilities and realised dominance rates to be broadly equal in many cases. That is, the results for individual groups of borrower would be concentrated around the dotted 45 degree line. Instead, it is clear that the vast majority of the consumer 'clusters' have performed better than random chance. In fact, the cubic spline fitted to summarise the (univariate) relationship between the average proportion of dominated products in the borrower cluster and its realised dominance rates has the slope of 0.5. On average, 10 percentage point increase in the proportion of dominated products in the choice set, the frequency of dominated mortgage choices increases by at most 5 percentage points.

This exercise provides reassurance that households invest effort into search and make better decisions that chance would suggest, regardless of their demographic profile and use of specialist support (eg an independent intermediary).

[24]Given the extreme computational demands of these pairwise comparisons, I cannot derive probabilities by simulating random behaviour.

9 Discussion and conclusion

This paper proposes a dominance-based approach to identifying search frictions in markets with heterogeneous consumers, varying product features and complex pricing. Applying this methodology to choices of borrowers in the UK mortgage market between January 2015 and July 2016 reveals the following key conclusions:

- Around 30% of UK households choose mortgage products that are strictly dominated by apparently available alternatives and incur economically significant costs as a result.
- The likelihood of choosing a dominated mortgage and the size of the resulting loss vary with the demographics that are often associated with financial capability — income, age and credit score (ability to manage other finances) — and also the complexity of the choice the borrower faces.
- Uncertainty about some of the lending criteria could also be contributing to dominated choices, especially among borrowers who are not at the top of the credit score or income distributions and face some (at least perceived) risk of falling short of lenders' requirements.
- Perhaps due to the uncertainty about product eligibility criteria and the overwhelming numbers of alternatives, intermediaries and borrowers appear to focus more on familiar lenders.

These findings are robust to a wide range of robustness checks. Any specific policy implications of these empirical facts, however, need to be interpreted in the broader context of the market, as discussed in the FCA Mortgages Market Study (FCA, 2018).

The dominance methodology developed in this paper offers a way of detecting poor consumer outcomes on cost that does not require making trade-offs between different product features and cost elements. This makes it more robust to uncertainty about borrower preferences than traditional cost comparison metrics such as APR benchmarking. This methodology is particularly well-suited to measuring frictions in consumer search in markets with multidimensional pricing, large degree of product differentiation and heterogeneous consumers needs — circumstances that apply to many retail financial markets and beyond.

This robustness comes at the price of only being able to identify choices that *cannot* be explained on cost grounds under any assumptions about the consumer's future behaviour or preferences for product features. There may be cases where a borrower has alternatives that are not strictly superior on all price characteristics to their chosen product, but would still have been better for that individual consumer given their preferences.[25] Extending the proposed dominance methodology to 'rank' products on factors other than cost could be an interesting area of further research, whether in household finance or other retail product choice applications.

The level of dominated choices could be lower than reported in this paper if lenders impose additional eligibility criteria that are not reported in the data and cannot be inferred

[25] For instance, the alternative might have slightly higher fixed fees but a lower interest rate than the chosen product, and the specific borrower has no strong need to minimise up-front costs and their loan is large enough loan for interest rate savings to offset the higher fees. The chosen product might have been a better option for another household, however, with a higher preference to minimise borrowing costs early on or a much smaller loan.

from observed transactions (in the way that for example minimum acceptable credit scores can be). If that were the case, some of the superior mortgage products might not have been available to the consumer when they appeared to be.

As discussed in the Robustness section, this concern is unlikely to be material for the findings in this paper in practice. The rate of dominated choices is 26%, close to the general sample, even among the very prime borrowers—those with relatively high incomes, low LTVs, and no factors that could typically complicate a mortgage application (self-employment, unusual property, etc). Those types of borrowers would be less likely to be investigated closely by the lenders and are much less likely to have their options restricted by idiosyncratic criteria as well. The high rate of dominated choices even among this very prime set of borrowers suggests that any additional unobservable lending criteria are unlikely to be a major driver of the economically significant results of this research.

This research also provides initial evidence about the specific frictions in the search and product choices that consumers make in the mortgage market. Further research, potentially in a more structural economic context, may be able to shed further light on how different factors interact in driving choices of particular products over others, albeit at the cost of making more assumptions about borrowers' needs and preferences than are required in this paper.

Bibliography

Agarwal, S., Driscoll, J. C., Gabaix, X., and Laibson, D. (2009) The age of reason: financial decisions over the life cycle and implications for regulation, *Brookings Papers on Economic Activity* **2009**, 51–117.

Agarwal, S., Rosen, R. J., and Yao, V. (2015) Why do borrowers make mortgage refinancing mistakes?, *Management Science* **62**, 3494–3509.

Allen, J., Clark, R., and Houde, J.-F. (2014a) Price dispersion in mortgage markets, *The Journal of Industrial Economics* **62**, 377–416.

Allen, J., Clark, R., and Houde, J.-F. (2014b) The effect of mergers in search markets: evidence from the canadian mortgage industry, *The American Economic Review* **104**, 3365–3396.

Andersen, S., Campbell, J. Y., Nielsen, K. M., and Ramadorai, T. (2015) Inattention and inertia in household finance: evidence from the danish mortgage market.

Bar-Gill, O. (2012) *Seduction by contract: law, economics, and psychology in consumer markets*. Oxford University Press.

Becker, G. S. (1965) A theory of the allocation of time, *The Economic Journal* **75**, 493–517.

Belgibayeva, A. and Majer, T. (2018) Tbc, *Financial Conduct Authority Occasional Paper* **35**.

Benetton, M. (2017) Leverage regulation and market structure:an empirical model of the uk mortgage market.

Best, M. C., Cloyne, J., Ilzetzki, E., and Kleven, H. J. (2015) Interest rates, debt and intertemporal allocation: evidence from notched mortgage contracts in the united kingdom, *Bank of England Staff Working Paper* **543**.

Bester, H. (1985) Screening vs. rationing in credit markets with imperfect information, *The American Economic Review* **75**, 850–855.

Bhargava, S., Loewenstein, G., and Sydnor, J. (2017) Choose to lose: health plan choices from a menu with dominated option, *The Quarterly Journal of Economics* **132**, 1319–1372.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016) Female labor supply, human capital, and welfare reform, *Econometrica* **84**, 1705–1753.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2016) Competition for attention, *The Review of Economic Studies* **83**, 481–513.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009) Measuring the financial sophistication of households, *American Economic Review* **99**, 393–98.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011) Robust inference with multiway clustering, *Journal of Business & Economic Statistics* **29**, 238–249.

Campbell, J. Y. and Cocco, J. F. (2003) Household risk management and optimal mortgage choice, *The Quarterly Journal of Economics* **118**, 1449–1494.

Cloyne, J., Huber, K., Ilzetzki, E., and Kleven, H. (Sept. 2017) *The effect of house prices on household borrowing: a new approach*. Working Paper 23861. National Bureau of Economic Research.

Coulibaly, B. and Li, G. (2009) Choice of mortgage contracts: evidence from the survey of consumer finances, *Real Estate Economics* **37**, 659–673.

Coval, J. D. and Moskowitz, T. J. (1999) Home bias at home: local equity preference in domestic portfolios, *The Journal of Finance* **54**, 2045–2073.

Egan, M. (2016) Brokers vs. retail investors: conflicting interests and dominated products.

Elton, E. J., Gruber, M. J., and Busse, J. A. (2004) Are investors rational? choices among index funds, *The Journal of Finance* **59**, 261–288.

Financial Conduct Authority (2018) *Mortgages market study: initial findings*. eprint: https://www.fca.org.uk/publications/market-studies/mortgages-market-study.

Finke, M. S., Howe, J. S., and Huston, S. J. (2017) Old age and the decline in financial literacy, *Management Science* **63**, 213–230.

Freixas, X. and Rochet, J.-C. (2008) *Microeconomics of banking*.

Gabaix, X. and Laibson, D. (2006) Shrouded attributes, consumer myopia, and information suppression in competitive markets, *The Quarterly Journal of Economics* **121**, 505–540.

Gurun, U. G., Matvos, G., and Seru, A. (2016) Advertising expensive mortgages, *The Journal of Finance* **71**, 2371–2416.

Huberman, G. (2001) Familiarity breeds investment, *The Review of Financial Studies* **14**, 659–680.

Iscenko, Z. and Nieboer, J. (2018) The value of advice and intermediation in the UK mortgage market, *Financial Conduct Authority Occasional Paper* **34**.

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

Lusardi, A. and Mitchell, O. S. (2014) The economic importance of financial literacy: theory and evidence, *Journal of Economic Literature* **52**, 5–44.

Marmorstein, H., Grewal, D., and Fishe, R. P. H. (1992) The value of time spent in price-comparison shopping: survey and experimental evidence, *Journal of Consumer Research* **19**, 52–61.

Pool, V. K., Stoffman, N., and Yonker, S. E. (2012) No place like home: familiarity in mutual fund manager portfolio choice, *The Review of Financial Studies* **25**, 2563–2599.

Sinaiko, A. D. and Hirth, R. A. (2011) Consumers, health insurance and dominated choices, *Journal of Health Economics* **30**, 450–457.

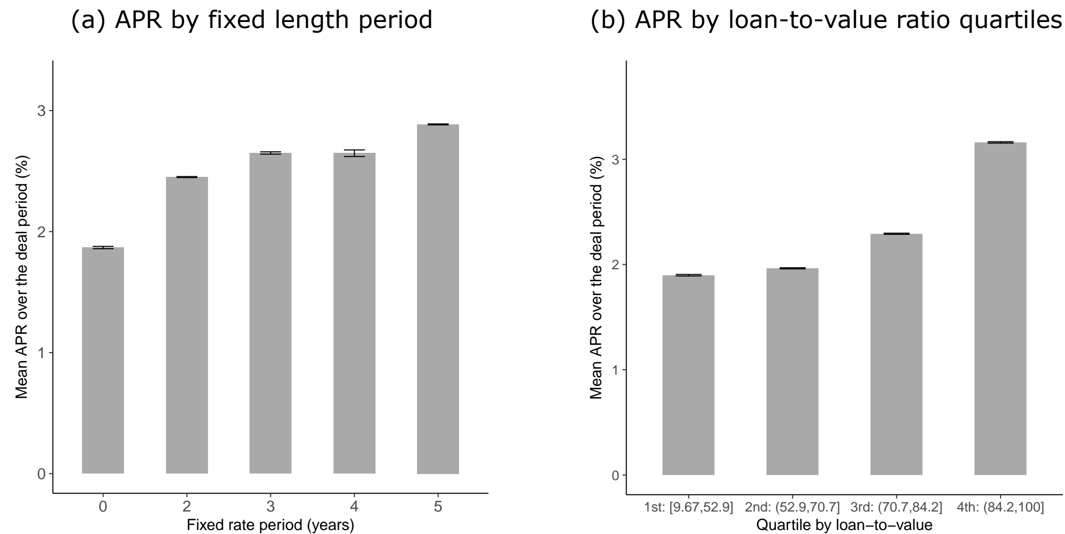
Van Hemert, O. (2010) Household interest rate risk management, *Real Estate Economics* **38**, 467–505.

Woodward, S. E. and Hall, R. E. (2012) Diagnosing consumer confusion and sub-optimal shopping effort: theory and mortgage-market evidence, *The American Economic Review* **102**, 3249–3276.

A Figures and tables

Figure A.1: Variation in borrowing costs by product type

(a) APR by fixed length period (b) APR by loan-to-value ratio quartiles



Note: Black error bars show the 99.9% confidence interval for conditional mean estimates.

Table A.1: Selected summary statistics: Population-sample comparison

Variable	Sample			Population[a]		
	Mean	σ	Median	Mean	σ	Median
Income[b](£1000)	43.74	26.95	37.65	43.55	34.55	37.00
Loan value (£1000)	179.69	128.60	146.79	184.45	134.33	150.53
Main borrower age (years)	38.17	9.12	37.00	37.67	8.98	37.00
Loan-to-value (LTV, %)	66.14	20.50	70.73	67.73	19.03	72.63
Initial interest rate (%)	2.56	0.738	2.44	2.58	0.791	2.44
Mortgage term (years)	23.33	7.50	25.00	23.93	7.40	25.00
=1 if self-employed	0.104			0.104		
=1 if first-time buyer	0.26			0.275		
=1 if remortgager	0.406			0.36		
=1 if fixed rate mortgage	0.921			0.90		
Observations		695,849			1,087,766	

Notes: (a) Due to mandatory nature of regulatory reporting by mortgage lenders, full PSD data are the population of residential mortgage lending in the UK. To focus on the potential selection arising from imperfect merge with other data sources, PSD population data are filtered down to mortgages relevant for this research using the criteria in section 3.2, e.g. LTV$\geq$20%. (b) Income is post-tax household earnings added across all borrowers named on the loan.

Table A.2: Sample descriptive statistics: Demographic

	Mean	σ	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
Income[a] (£1000)	43.99	31.83	27.38	37.66	52.20
Property value (£1000)	290.26	221.82	156.00	230.00	350.00
Loan value (£1000)	179.97	130.07	99.95	146.87	221.00
Main borrower age (years)	38.18	9.12	31.00	37.00	45.00
Loan-to-value (LTV, %)	66.14	20.50	52.90	70.73	84.16
Credit score[b]	63.19	8.49	59.74	64.97	68.89
Balance on unsecured debt[c] (% income)	19.31	29.55	0.169	4.87	28.30
Postcode: % in low-skill occupations	29.03	12.78	19.58	27.20	36.74
Mortgage term (years)	23.33	7.50	18.00	25.00	30.00
=1 if impaired credit history	0.00298				
=1 if first-time buyer	0.26				
=1 if home mover	0.334				
=1 if remortgager	0.406				
=1 if self-employed	0.104				
Observations			695,849		

Note: (a) Income is post-tax household earnings added across all borrowers named on the loan. (b) Overall borrower credit score as reported to mortgage lenders by one of the three major credit bureaus in the UK, normalised to range from 0 to 100. (c) Unsecured debt is all borrowing on the household credit file excluding mortgages, car loans and leasing agreements.

Table A.3: Sample descriptive statistics: Borrowing costs

	Mean	σ	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
APR over deal period (%)	2.57	0.759	2.02	2.44	2.99
APR over 5 years (%)					
if stay on reversion rate	3.25	0.575	2.89	3.21	3.59
if switch to same loan after deal period	2.59	0.797	2.01	2.50	3.06
Initial interest rate (%)	2.70	0.761	2.14	2.54	3.14
Initial payment$a$ (£per month)	889.22	632.57	515.59	726.76	1070.85
Initial payment (% income)	24.71	7.98	19.61	24.09	28.91
Up-front fee (£)	580.24	530.64	0.00	760.00	999.00
Early repayment penalty (% of loan)	2.63	0.97	1.50	2.60	3.00
Deal period length$b$ (years)	3.15	1.68	2.00	2.00	5.00
=1 if fixed rate mortgage	0.921				

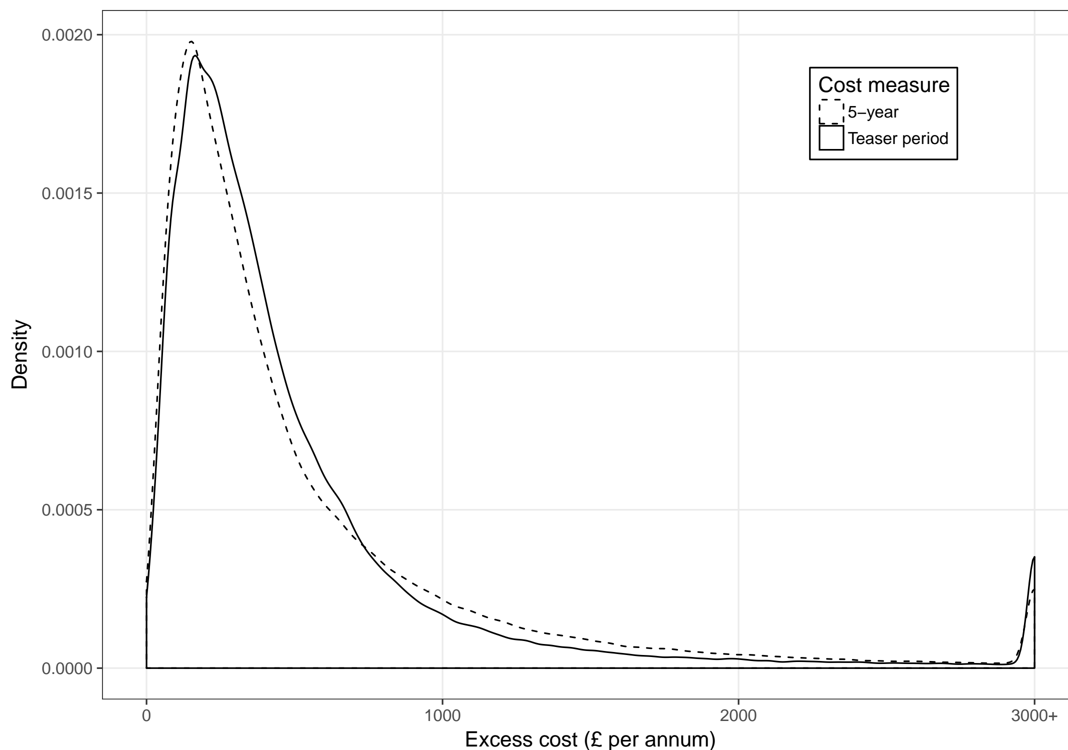
Note: (a) Initial mortgage payment is the minimum monthly interest and principal repayment the borrower has to pay during their introductory deal period. (b) Length of the fixed rate or the introductory discounted variable rate discount.

Table A.4: Dominated choices: headline results

	Mean	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
Households with dominated products (%)	29.9			
Households with strongly dominated products (%) [a]	17.5			
Excess cost from dominated choices (deal period)[b]				
£ per year	549.70	184.63	330.15	583.73
% of annual mortgage payment	12.73	5.94	10.19	15.88
% of household income	1.47	0.531	0.955	1.64
Excess cost from dominated choices (5-year)[c]				
£ per year	565.56	170.67	327.71	659.31
% of annual mortgage payment	10.13	4.35	7.77	13.97

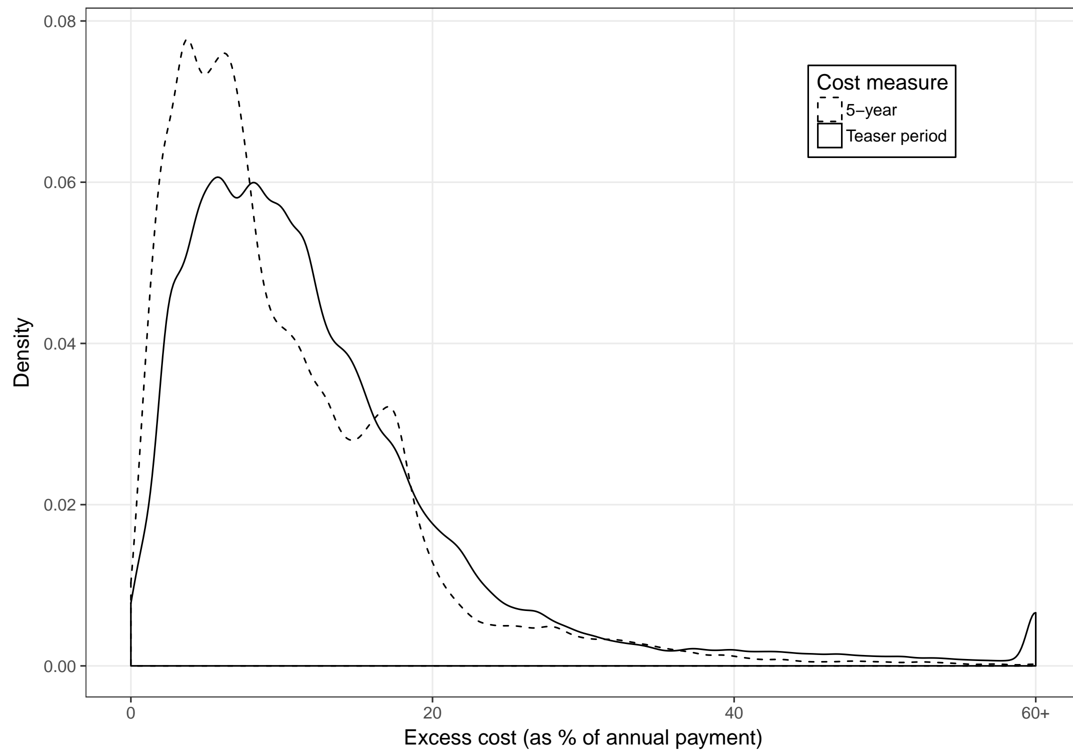
Notes: (a) strong dominance is defined as a choice which results in annual excess cost greater than or equal to £250 per annum and 5% of borrower's annual mortgage cost. (b) All distributions of excess costs are for the subset of borrowers with dominated products. Excess cost (as per definition 6 in section 4.2) is the mean difference between household's borrowing costs with the chosen (dominated) mortgage product and borrowing costs (for the same loan amount, etc.) under each of the strictly superior mortgage products that dominate the observed choice. (c) Five-year excess cost assumes that the borrower remains on their contractual reversion rate until the end of the five-year period if their fixed rate or discounted rate deal expires earlier.

Figure A.2: Distribution of absolute excess costs from dominated choices



Note: All distributions of excess costs are for the subset of borrowers with dominated products. Where excess costs are calculated on the basis of remortgaging after five years, the borrower remains on the reversion rate until the end of the five-year period if their fixed rate or discounted rate deal expires earlier.

Figure A.3: Distribution of relative excess costs from dominated choices



Note: All distributions of excess costs are for the subset of borrowers with dominated products. Where excess costs are calculated on the basis of remortgaging after five years, the borrower remains on the reversion rate until the end of the five-year period if their fixed rate or discounted rate deal expires earlier.

Table A.5: Demographic comparison of borrowers with and without dominated mortgages

Variable	Sub-sample mean		
	Not dominated	All dominated	Strongly dominated
Income (£1000)	45.06	41.50	43.05
Property value (£1000)	298.69	270.55	276.91
Loan value (£1000)	185.68	166.59	181.79
Borrower age (years)	37.75	39.16	38.12
Loan-to-value (LTV, %)	66.47	65.39	69.99
Loan-to-income ratio	3.02	2.97	3.12
Mortgage term (years)	23.50	22.95	24.19
Postcode: % in low-skill occupations	28.53	30.18	29.91
Credit score	63.45	62.57	61.94
Balance on unsecured debt (£)	7960.04	8537.65	9236.82
Balance on unsecured debt (% income)	18.41	21.42	22.27
=1 if impaired credit history	0.00231	0.00177	0.00287
=1 if self-employed	0.093	0.131	0.134
=1 if first-time buyer	0.265	0.246	0.276
=1 if home mover	0.314	0.381	0.409
=1 if remortgager	0.42	0.373	0.315
Observations	487,687	208,373	121,561

Table A.6: Marginal effects on probability of strongly dominated choice (in pp)

	Marginal effect[a]	Effect of 1σ change
Income (in £1000)	−0.19	−5.17
Loan value (£1000)	0.05	6.32
=1 if joint mortgage[b]	−0.15	
Number of dependent children	1.23	1.18
Borrower age	0.14	1.30
=1 if home mover	3.63	
=1 if remortgager	−3.25	
Postcode: % in low-skill occupations	0.03	0.40
Number of available products[c]	0.01	3.45
Application complications:		
=1 if new-build property	3.61	
=1 if self-employed	4.78	
=1 if non-standard property[d]	2.38	
Credit score	−0.28	−2.36
Balance on unsecured debt (% income)[e]	0.08	2.28
Loan-to-value ratio (LTV, %)	0.24	4.84
Postcode: % unemployed	0.11	0.22
Observations	647,758	

Notes: (a) Marginal effects calculated using regression specification (2) in Table A.8. All reported effects are statistically significant at 1% level unless stated otherwise. Standard errors are computed using the delta method and clustered on year and postcode area. (b) Not significant. (c) Number of products in borrower's comparable set of available products as defined in definition 3. (d) Proxy for the mortgaged property having non-standard characteristics takes the value of 1 if the purchased property is in the bottom price decile among properties with the same number of rooms mortgaged in the same postcode and in the quarter of transaction. This reflects that properties with any lending restrictions are typically sold at a considerable discount. (e) Total amount of borrowers' non-mortgage and non-car debt (i.e. credit card balances, personal loans, used overdrafts and other forms of short-term credit) at the point of mortgage application, as percent of post-tax income.

Table A.7: Marginal effects on excess cost (in pp of annual mortgage cost)

	Marginal effect[a]	Effect of 1σ change
Income (in £1000)	0.0910	2.45
Loan value (£1000))[b]	−0.0831	−10.69
=1 if joint mortgage	−0.2540	
Number of dependent children	0.2150	0.21
Borrower age	0.0135	0.12
=1 if home mover[c]	0.0262	
=1 if remortgager[d]	−0.1995	
Postcode: % in low-skill occupations[d]	−0.0044	−0.06
Number of available products[d,e]	−0.0002	−0.06
Application complications:		
=1 if new-build property[d]	0.5724	
=1 if self-employed	1.9019	
=1 if non-standard property[d,f]	−0.2353	
Credit score	−0.1744	−1.48
Balance on unsecured debt (% income)[g]	0.0138	0.41
Loan-to-value ratio (LTV, %)	−0.0776	−1.59
Postcode: % unemployed	−0.1356	−0.27
Observations	112,363	

Notes: (a) Marginal effects calculated using regression specification (4) in Table A.8. All reported effects are statistically significant at 1% level unless stated otherwise. Standard errors are computed using the delta method and clustered on year and postcode area. (b) Not significant. (c) Significant at 5% level.
(d) Since the outcome variable is measured in percent of the overall mortgage payment, loan value has a strong mechanical effect on the measure by increasing the denominator and the estimated marginal effect does not have a meaningful interpretation. (e) Number of products in borrower's comparable set of available products as defined in definition 3. (f) Proxy for the mortgaged property having non-standard characteristics takes the value of 1 if the purchased property is in the bottom price decile among properties with the same number of rooms mortgaged in the same postcode and in the quarter of transaction. This reflects that properties with any lending restrictions are typically sold at a considerable discount. (g) Total amount of borrowers' non-mortgage and non-car debt (i.e. credit card balances, personal loans, used overdrafts and other forms of short-term credit) at the point of mortgage application, as percent of post-tax income.

Table A.8: Regression results for main borrower characteristics

	=1 if strongly dominated, logistic			Average excess cost [a], OLS	
	(1)	(2)	(3)	(4)	(5)
Borrower age (years)	0.013***[b]	−0.042***	−0.043***	0.172***	0.175***
Borrower age (years) ^{2}		0.001***	0.001***	−0.002***	−0.002***
Loan-to-value ratio (LTV, %)	0.021***	0.017***	0.018***	−0.078***	−0.083***
log(Loan value (£1000))	−0.106***	6.150***	6.324***	−105.104***	−108.041***
log(Loan value (£1000))^2		−0.245***	−0.254***	4.050***	4.180***
log(Income (£1000))	0.034	−4.023***	−4.206***	34.366***	35.555***
log(Income (£1000))^2		0.181***	0.190***	−1.555***	−1.608***
=1 if joint mortgage	−0.019	12.043***	12.002***	−82.421***	−76.184***
=1 if joint$\times log$(Income (£1000))		−1.942***	−1.934***	14.160***	12.991***
=1 if joint$\times log$(Income (£1000))^2		0.076***	0.075***	−0.602***	−0.548**
Number of dependent children	0.076***	0.089***	0.092***	0.215***	0.189***
=1 if remortgagers	−0.216***	−0.252***	−0.251***	−0.199	−0.211
=1 if home movers	0.235***	0.242***	0.250***	0.026	−0.036
Credit score	−0.022***	−0.028***	−0.029***	−1.552***	−1.543***
Credit score^2		0.0001*	0.0001*	0.011***	0.011***
=1 if self-employed	0.342***	0.320***	0.322***	1.902***	1.857***
=1 if non-standard property[b]	0.142***	0.165***	0.156***	−0.235	−0.201
=1 if new-build property	0.265***	0.243***	0.252***	0.572*	0.569*
Postcode: % in low-skill occupations	0.002***	0.002***	0.002***	−0.004	−0.005
Postcode: % unemployed	0.004	0.008**	0.005*	−0.136***	−0.093***
Balance on unsecured debt (% income) [c]	0.004***	0.006***	0.006***	0.016***	0.015***
Balance on unsecured debt (% income)^2 [c]		−0.00002***	−0.00002***	−0.00005***	−0.00004**
Number of available products [d]	0.001***	0.002***	0.002***	0.004***	0.004***
Number of available products^2 [d]		0.00000***	0.00000***	0.00000***	0.00000***
Channel FEs	Yes	Yes	Yes	Yes	Yes
Regional FEs	No	No	Yes	No	Yes
Observations	647,758	647,758	647,758	112,363	112,363
R^2				0.233	0.234

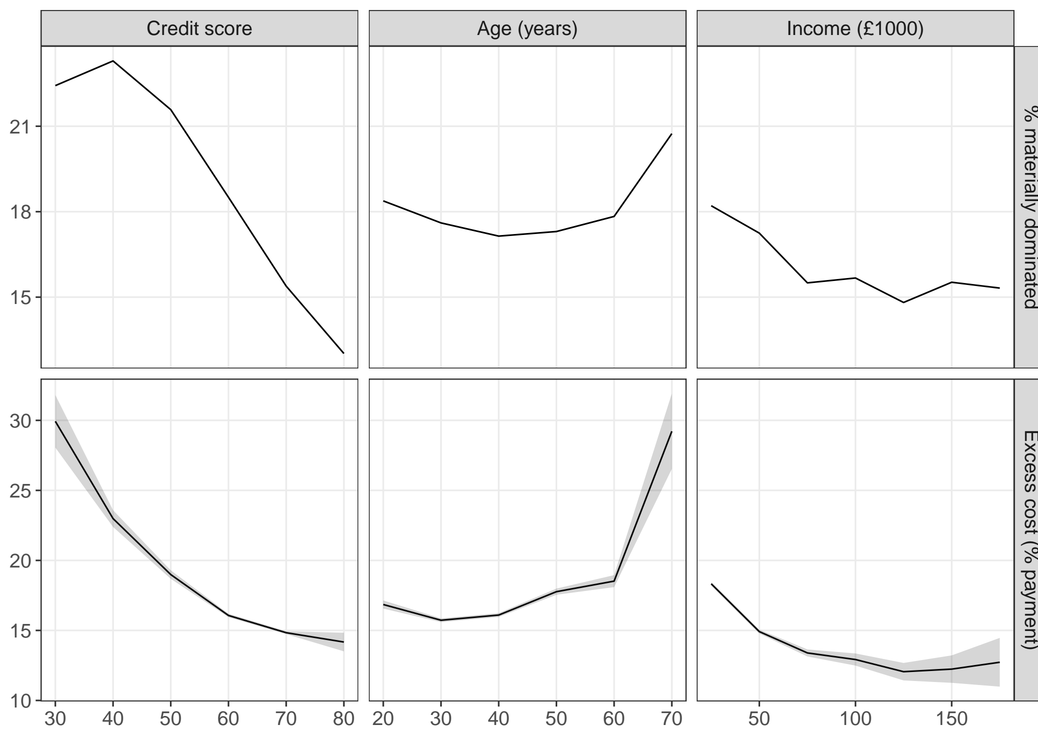
Notes: (a) Average annual excess cost expressed as percentage of the total annual mortgage payment calculated over the duration of the initial deal period. (b) Proxy for the mortgaged property having non-standard characteristics takes the value of 1 if the purchased property is in the bottom price decile among properties with the same number of rooms mortgaged in the same postcode and in the quarter of transaction. This reflects that properties with any lending restrictions are typically sold at a considerable discount. (c) Total amount of borrowers' non-mortgage and non-auto debt (i.e. credit card balances, personal loans, used overdrafts and other forms of short-term credit) at the point of mortgage application, as percent of post-tax income. (d) Number of products in borrower's comparable set of available products as defined in definition 3.
*p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered on year and postcode area.

Table A.9: Effects of changing consideration sets on in dominance rates of intermediated transactions

		% transactions dominated within consideration set	
	Assumed consideration set	All intermediaries	Intermediaries with known panels
	All available products	30.8	30.1
A	No direct-only products	21.3	20.2
B	Products from panel lenders		23.5
	Products from familiar lenders:		
C	jointly accounting for 90% business	18.4	19.4
D	jointly accounting for 75% business	15.4	15.7
	All restrictions (A+B+D)	12.8	12.9
	Observations	133,818	432,291

Notes: The percentages of business are calculated on the basis of the number of borrowers directed by the intermediary to a given lender in 2014 as a proportion of the total number of borrowers who took mortgages through this intermediary in the same year.

Figure A.4: Dominance rate and excess cost split by demographics



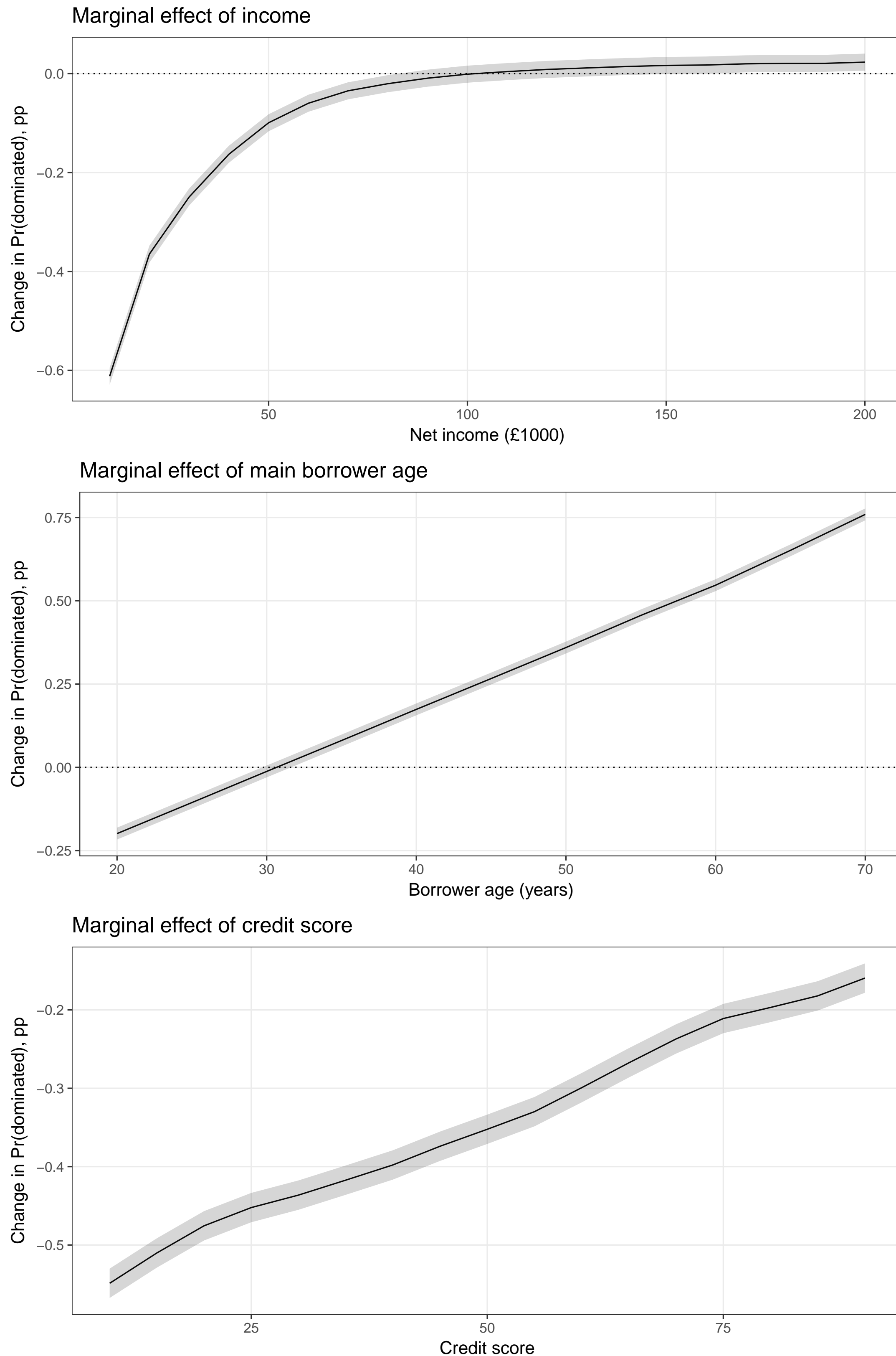
Note: Average rates and costs calculated in 10-point groups for credit score and age, and in £25k buckets for income. Borrowers in top and bottom 0.5% on each of the demographics are classified as outliers and omitted from the graphs. Grey bands show the 99.9% confidence interval for the conditional means.

Table A.10: Comparison of mean results in the very prime and full samples

	Very prime	All
Households with dominated products (%)	26.1	29.9
Mean excess borrowing cost (deal period)		
£ per year	284.66	549.70
% of annual mortgage payment	13.12	12.73
Mean excess borrowing cost (5-year)		
£ per year	360.63	565.56
% of annual mortgage payment	11.45	10.13
Observations	62,979	695,831

Note: The 'very prime' sub-sample contains borrowers who satisfy all of the following criteria: LTV$\leq$65%, loan-to-income ratio $\leq$2.5, mortgaged property that is not a new-build or sold at a discount that could indicate other irregularities, credit score in the top 25% of all borrowers, and the borrower is not self-employed or due to repay the loan after reaching retirement age.

Figure A.5: Variation in marginal effects of income, age and credit score on dominated choices



Note: Individual marginal effects calculated using regression specification (2) in Table A.8 and aggregated into bins of £10 for income and of 5 for age and credit score to obtain conditional averages. Standard errors obtained using delta method. Grey bands show the 99.9% confidence interval for the conditional marginal effect estimate.

Figure A.6: Role of mortgage price components(PC) in excess costs

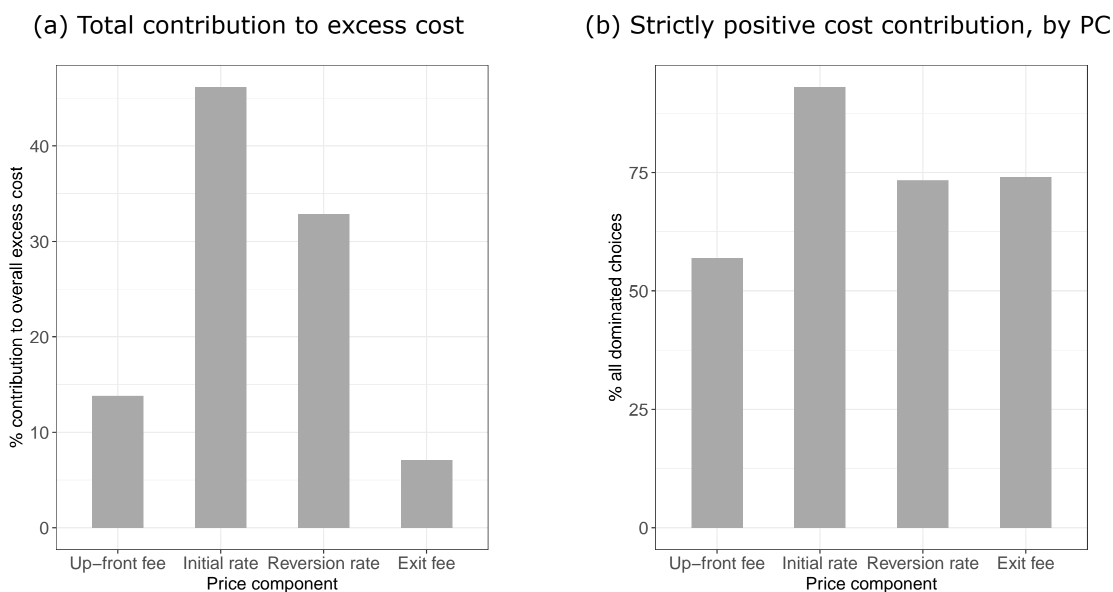
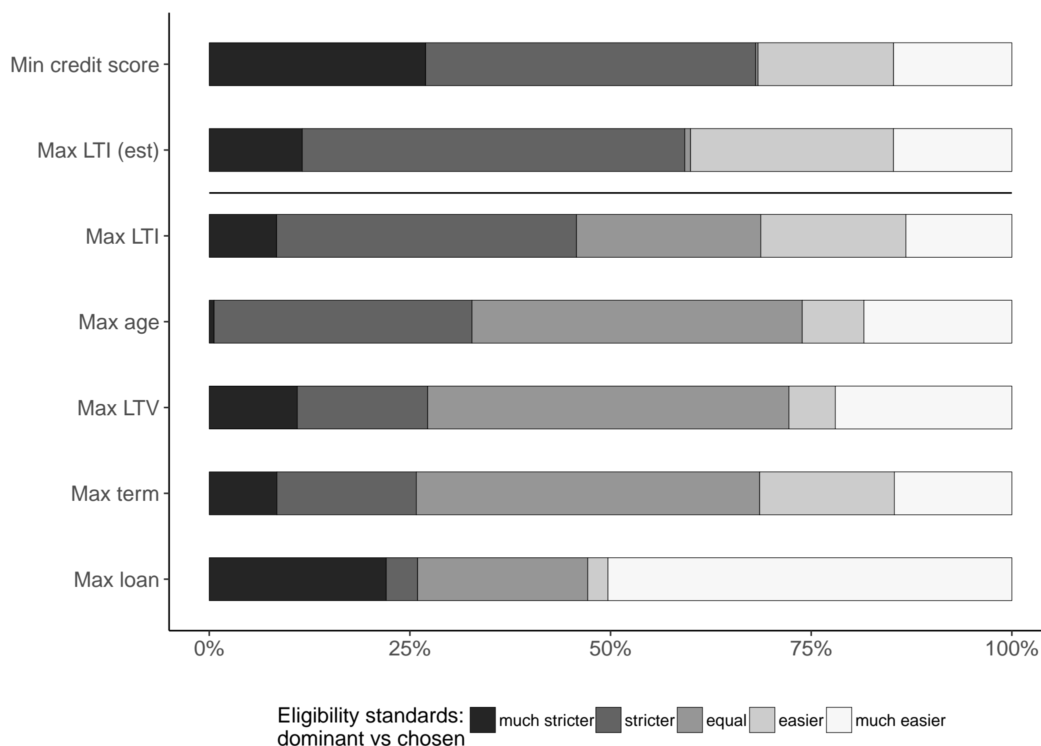
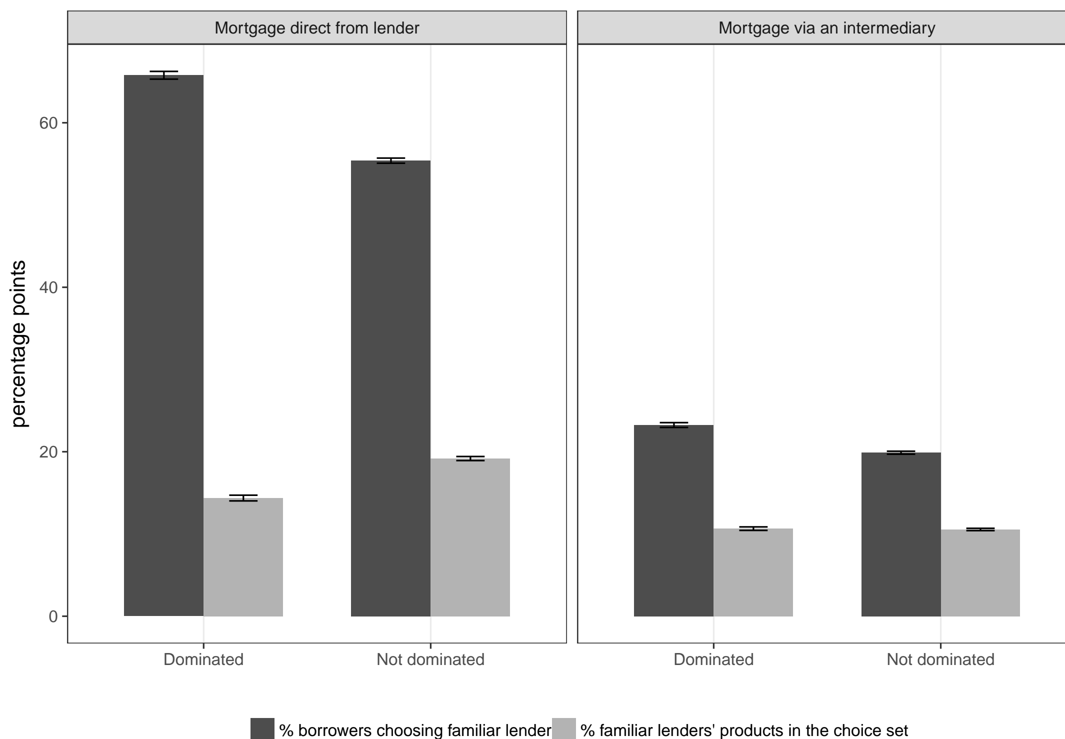


Figure A.7: Eligibility criteria of better products compared to chosen dominated option



Note: A dominating alternative with a higher eligibility standard than the chosen product is classified as having stricter (easier) requirements. The requirement is defined as much stricter/easier, if the absolute value of the difference between in the minimum requirements of the dominating and the chosen products is 50% or more of the total 'headroom' the borrower has over breaching the minimum standard on their chosen (dominated) product. (This difference cannot be greater than 100% because the borrower meets all the minimum criteria for the 'dominating' products by definition.)

Figure A.8: Propensity to choose lenders with an existing relationship

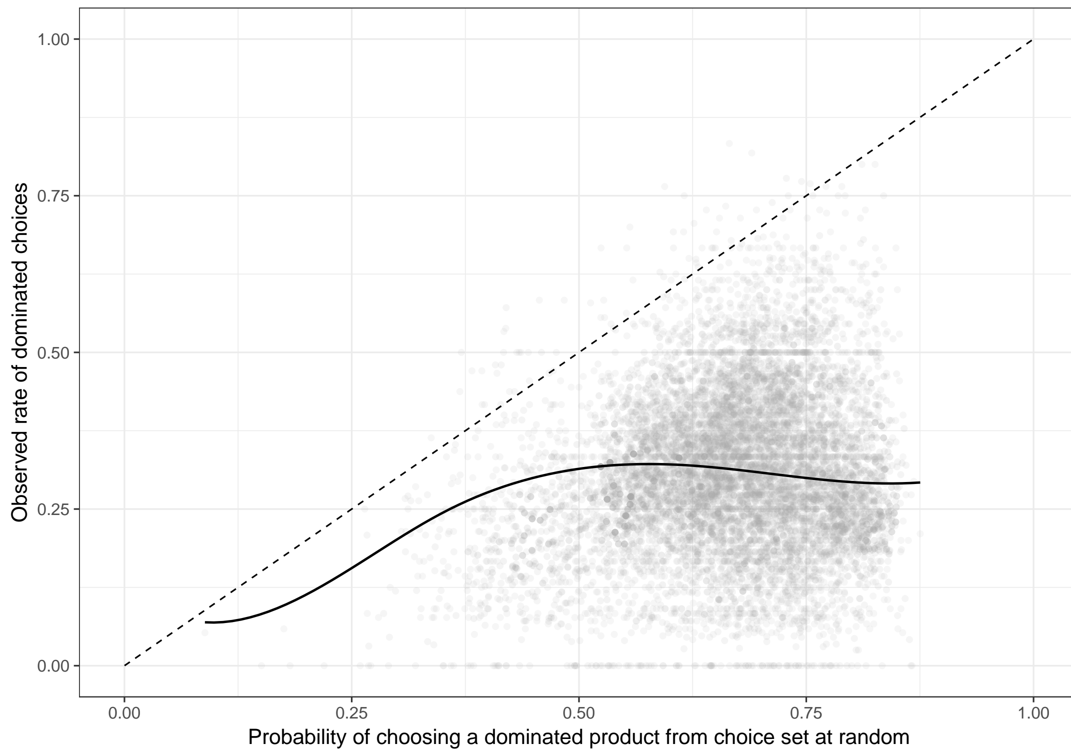


Note: A borrower is considered to have an existing relationship with the lender if they had at least one of the following products with this lender at or before the time of their mortgage application: a current account, an overdraft, a credit card or a personal loan (mortgages are excluded due to data limitations).

Table A.11: Distributions of annual excess cost (£pa) under different time horizons

			Smallest possible	
	Deal period	5 years	Total cost	Annual cost
Mean	549.70	565.56	708.49	229.02
$Q_{0.25}$	184.63	170.67	156.34	47.50
$Q_{0.5}$	330.15	327.71	350.32	108.75
$Q_{0.75}$	583.73	659.31	745.92	256.47
Years (median)	2	5	1	17

Figure A.9: Observed dominated choice rates vs predicted under random mortgage choices



Note: Individual dots represent each of the 15,800 clusters of borrowers defined by permutation of quintiles of key demographic variables (LTV, credit score, age, etc) and additional discrete borrower categories (e.g. self-employed vs not). If households chose products from their full comparable set of available products (definition 3) at random, one would expect observations to lie along the dotted 45 degree line. The solid line summarises the actual relationship between probabilities of dominated choices and the realised dominated rates using a cubic spline with 4 degrees of freedom.

Figure A.10: Distribution of excess costs incurred due to dominated mortgage choice

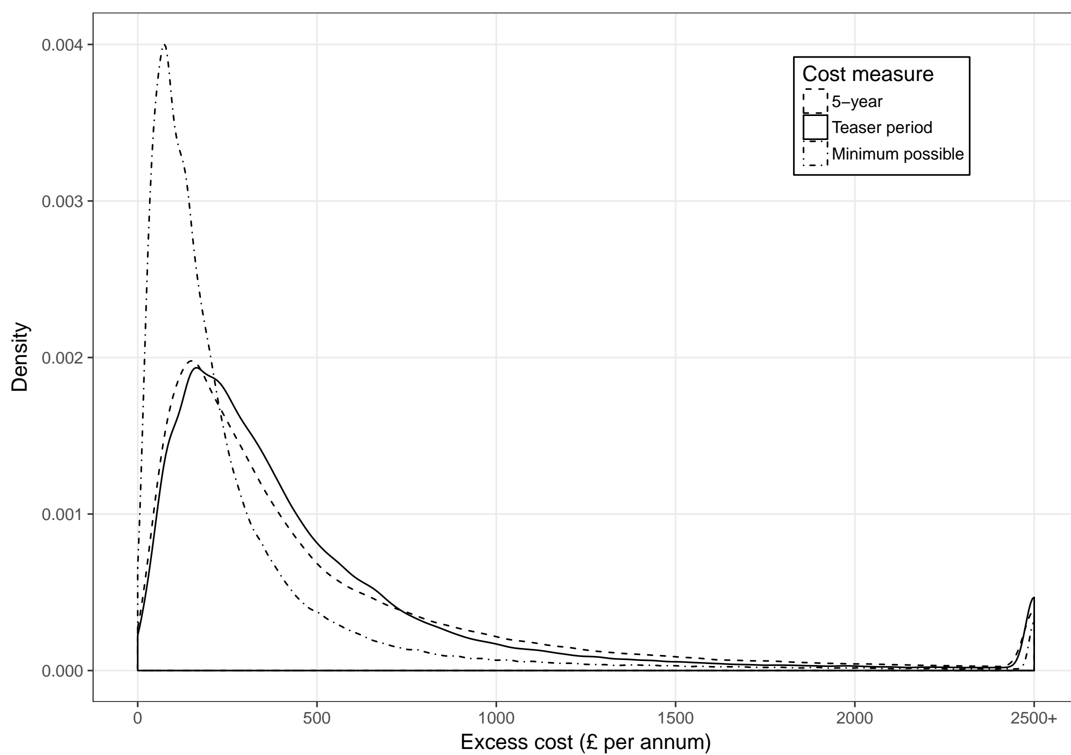


Table A.12: Effects of sample re-weighting on selected variables and results

	Mean values		
Variable	Sample	Weighted sample	Population
Gross income (£1000)	63.16	63.32	63.42
Net income (£1000)	43.78	43.70	43.64
Loan value (£1000)	179.75	185.15	186.63
Borrower age (years)	38.17	37.66	37.63
Loan-to-value (LTV, %)	66.14	67.48	67.89
Initial interest rate (%)	2.56	2.56	2.57
Mortgage term (years)	23.33	23.91	24.01
=1 if self-employed	0.104	0.105	0.104
=1 if first-time buyer	0.26	0.275	0.275
=1 if home mover	0.334	0.365	0.365
=1 if remortgager	0.406	0.36	0.36
=1 if fixed rate mortgage	0.921	0.899	0.899
Results:			
=1 if dominated	0.299	0.295	
=1 if strongly dominated	0.175	0.177	
Excess cost (% annual payment)	12.73	12.43	
Observations	695,849	695,849	1,042,204

Table A.13: Robustness of effects to alternative definitions of strong dominance

	=1 if dominated choice and excess cost ≥:				
	Baseline (£250 & 5%)	Any (£0 & 0%)	£250 or 5%	£250 & 10%	£500 & 10%
Borrower age (years)	−0.044***	−0.053***	−0.048***	−0.031***	−0.031***
Borrower age (years)^2	0.001***	0.001***	0.001***	0.001***	0.001***
Loan-to-value ratio (LTV, %)	0.018***	0.014***	0.016***	0.016***	0.019***
$log($Loan value (£1000)$)$	6.168***	−3.636***	−6.756***	8.045***	10.559***
$log($Loan value (£1000)$)^2$	−0.246***	0.146***	0.274***	−0.345***	−0.422***
$log($Income (£1000)$)$	−4.078***	−3.428***	−2.689***	−5.498***	−3.327***
$log($Income (£1000)$)^2$	0.185***	0.153***	0.115***	0.252***	0.158***
=1 if joint mortgage	12.724***	−0.085	−0.470	21.956***	16.304***
=1 if joint$\times log($Income (£1000)$)$	−2.058***	0.207	0.292	−3.777***	−2.670***
=1 if joint$\times log($Income (£1000)$)^2$	0.080***	−0.019	−0.024	0.160***	0.106***
Number of dependent children	0.089***	0.075***	0.084***	0.090***	0.083***
=1 if home mover	0.251***	0.271***	0.269***	0.274***	0.255***
=1 if remortgager	−0.244***	−0.117***	−0.156***	−0.188***	−0.357***
Credit score	−0.013***	0.003	0.005	−0.022***	−0.043***
Credit score^2	−0.0001	−0.0002***	−0.0002***	−0.00001	0.0002***
=1 if self-employed	0.305***	0.349***	0.326***	0.304***	0.326***
=1 if non-standard property	0.163***	0.204***	0.181***	0.149***	0.158***
=1 if new-build property	0.249***	0.188***	0.242***	0.308***	0.206***
Postcode: % in low-skill occupations	0.002***	0.001***	0.001***	0.002***	0.002***
Postcode: % unemployed	0.009**	0.005*	0.004	0.010***	0.003
Balance on unsecured debt (% income)	0.005***	0.003***	0.004***	0.006***	0.007***
Balance on unsecured debt (% income)^2	−0.00001***	−0.00001***	−0.00001***	−0.00002***	−0.00002***
Number of available products	0.002***	0.002***	0.002***	0.002***	0.001***
Number of available products^2	0.000***	0.000***	0.000***	0.000***	0.000***
=1 if chose familiar lender	0.270***	0.261***	0.270***	0.213***	0.143***
Channel FEs	Yes	Yes	Yes	Yes	Yes
Observations	639,509	639,509	639,509	639,509	639,509

*p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered on year and postcode area.

Additional figures and tables

Figure A.11: Likelihood of at least one dominating alternative in consideration set if products drawn at random, by consideration set size

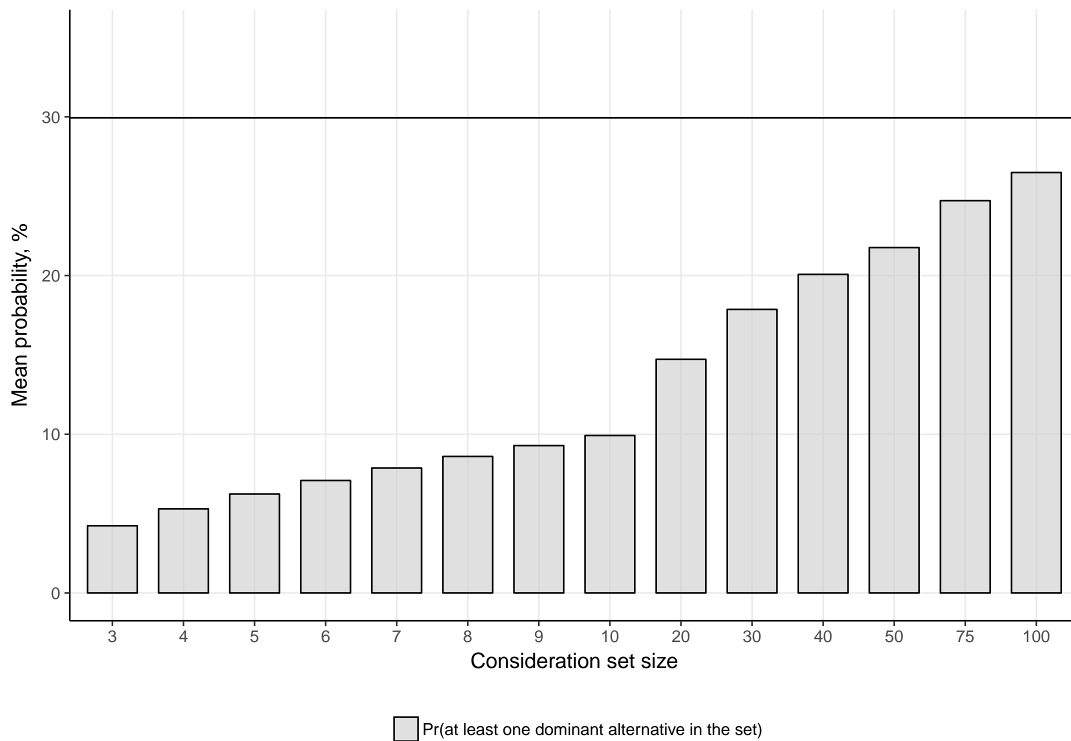


Table A.14: Dominated choices: mean headline results by borrower type

	All	FTBs	Movers	Remortgagers
Households with dominated products (%)	29.9	28.4	34.1	27.5
Households with strongly dominated products (%)	17.5	18.6	21.4	13.5
Excess cost from dominated choices (deal period)				
£ per year	549.70	591.58	641.65	426.33
% of annual mortgage payment	12.73	13.30	12.53	12.55
% of household income	1.47	1.89	1.54	1.13
Excess cost from dominated choices (5-year)				
£ per year	565.56	537.51	648.78	499.09
% of annual mortgage payment	10.13	9.94	9.89	10.49

Table A.15: Variation in dominance rates in intermediated transactions, by lender familiarity

	% dominated	% strongly dominated
lenders accounting for the top 75% of intermediary's sales	27.4	16.9
other lenders	31.7	18.2