University College London

Submission for the degree of Ph.D. in Economics

Methods for analysing consumer choice and their applications to the UK mortgage market

Zanna Iscenko

January 2020

Declaration

I, Zanna Iscenko, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis uses a unique combination of administrative and commercial data on the UK mortgage market to study how consumers make decisions in complex settings that involve multi-dimensional pricing, restricted product availability (from eligibility criteria) and potential brand loyalty towards some of the suppliers.

Chapter 2 develops a method for detecting apparently inefficient consumer decisions by identifying cases when the chosen product was strictly dominated by another available alternative. It finds that 30% of UK borrowers in 2015-16 chose mortgages that were strictly dominated and incurred substantial avoidable costs as a result. It also describes how the propensity to make dominated choices varies with borrower demographics, product characteristics, and presence of any existing relationship between the borrower and product provider.

Chapter 3 uses a structural approach to investigate these findings further. It applies a limited attention discrete choice model with two latent classes of borrowers to study which characteristics of suppliers and products affect (a) the likelihood of the alternative being considered and (b) preferences for the considered alternatives. It documents substantial inattention towards the available alternatives, which tends to be worse among the less financially sophisticated borrowers and finds that a borrower already holding another product with a lender significantly boosts that lender's probability of being considered. Even conditional on paying attention, all borrower types show a strong preference for lenders they are familiar with. The chapter also contains a counterfactual policy simulation of making all borrowers pay full attention and reports the effects in the new market equilibrium.

Chapter 4 (joint work with Jeroen Nieboer at the Financial Conduct Authority), uses difference-in-difference matching to measure the effects of the introduction of mandatory mortgage advice in the UK in 2014 on the borrowing costs and chosen product characteristics of the previously non-advised borrowers.

Impact Statement

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. This thesis advances the policy and academic understanding of consumer decisions in this important market in several ways.

On the regulatory side, this research has directly helped inform the view of the UK Financial Conduct Authority (FCA) on how well the UK mortgage market was working for consumers and the appropriateness of further policy interventions. In 2018, the FCA published earlier versions of chapters 2 and 4 of this thesis in its Occasional Paper series. The findings in these papers were cited extensively by the FCA Mortgages Market Study and the press coverage about it.[1] Chapter 4 also provided UK regulators with evidence on the (partial) impacts of a past policy intervention, namely the FSA Mortgage Market Review advice requirement introduced in 2014. In addition to the impact of my findings, as part of this research, I also helped the FCA create, clean and combine datasets that have subsequently been used in other regulatory analysis.

In terms of academic impact, this thesis makes substantive contributions in the field of household finance. I create and use new, uniquely granular combinations of datasets that have not been previously used in academic research. This allows me to provide new insights on live questions in the literature, such as: measurement and drivers of price dispersion (chapter 2), the effects of limited attention and brand loyalty on product choices (chapter 3), and the impacts of mandatory advice (chapter 4). The value of these contributions is further evidenced by Chapter 2 being among the three runners-up for the best student paper at the CEPR Household Finance Workshop 2019.

Furthermore, chapters 2 and 3 combine existing techniques in novel ways to develop new methods for analysing consumer decision-making and measuring frictions in a regulated market with a proliferation of differentiated products, multi-dimensional pricing, and potential pre-existing links between consumers and some of the suppliers. While I apply these methods to mortgages in this thesis, they can also help the study of other markets with similar characteristics, such as other financial products, utilities, and telecommunication services.

1 See, for instance, the FCA's Mortgages Market Study Interim Report, Chapter 5 and the related BBC, Financial Times, and The Guardian articles, among many others.

Acknowledgements

I am deeply indebted to my supervisor, Richard Blundell, for his insightful comments and unwavering encouragement. After every meeting, he left me with more confidence about existing work, more ideas about future improvements and more enthusiasm about pursuing those. This was exactly what I needed. I really cannot thank him enough.

I am also profoundly grateful to my second supervisor, Ian Preston, whose advice and concern about my welfare made it possible to deal with obstacles that seemed insurmountable at the time. Thank you, Ian, for reaching out when you did.

I owe a debt of gratitude to Matthew Rabin, David Laibson, and Walter Beckert for their encouraging me to pursue a PhD in the first place. And to many of the students and staff at UCL, especially Tom Hoe, for their advice, humour and perspective that have helped me finish it.

My research could not have happened without the support of the staff of the Financial Conduct Authority. I am immeasurably grateful to Peter Andrews, Carmen Suarez, Robin Finer, Karen Croxson, Nick Cook, and the many members of the Mortgages Market Study Team and the Household Finance and Economic Data Science Team, for their help over the years. A special mention, of course, also goes to my great co-author, Jeroen Nieboer.

On a more personal level, I am thankful for my partner, who has been brilliant throughout and has patiently endured more conversations about mortgages than anyone should ever be subjected to without getting a degree for it.

I am also grateful to my family and friends, who did not disown me despite my occasional PhD-related abandonment. I look forward to being able to spend more time with them again...

Contents

Abstract 5

Impact Statement 6

Acknowledgements 7

1 Preface 19

2 Strongly dominated mortgage choices 23
- 2.1 Introduction . . . 23
- 2.2 Background on the UK mortgage market . . . 27
  - 2.2.1 Product design . . . 27
  - 2.2.2 Price posting and search . . . 28
  - 2.2.3 Tools and policy context . . . 29
- 2.3 Data . . . 30
  - 2.3.1 Datasets . . . 30
  - 2.3.2 Sample construction and representativeness . . . 31
  - 2.3.3 Descriptive statistics . . . 32
- 2.4 Introduction to dominance analysis . . . 37
  - 2.4.1 Identification of dominated choices . . . 38
  - 2.4.2 Cost of dominated choices . . . 40
  - 2.4.3 Results on the frequency and costs of dominated mortgage choices . . . 43
- 2.5 Demographic variation . . . 44
  - 2.5.1 Univariate trends . . . 44
  - 2.5.2 Regression specification . . . 46
  - 2.5.3 Marginal demographic effects . . . 47
- 2.6 Familiarity . . . 54
  - 2.6.1 Intermediary search . . . 54
  - 2.6.2 Direct search . . . 56
- 2.7 Product and supplier characteristics . . . 58
  - 2.7.1 Supply of dominated products . . . 58
  - 2.7.2 Salience of price features . . . 61
  - 2.7.3 Contributions of additional supplier characteristics . . . 62
- 2.8 Robustness . . . 65
  - 2.8.1 Prime borrowers only . . . 65
  - 2.8.2 Definition of strong dominance . . . 66
- 2.9 Discussion and conclusion . . . 67

3 Limited attention and lender familiarity in mortgage choices 69
- 3.1 Introduction . . . 69

3.2 Data . . . 73
3.2.1 Datasets . . . 73
3.2.2 Sample description . . . 75
3.3 Model . . . 78
3.3.1 Demand . . . 78
3.3.2 Supply . . . 80
3.4 Estimation . . . 82
3.4.1 Demand . . . 82
3.4.2 Supply . . . 88
3.5 Results . . . 90
3.5.1 Demographic variation . . . 90
3.5.2 Limited attention . . . 93
3.5.3 The role of familiarity . . . 94
3.6 Counterfactual simulation: forced full attention . . . 98
3.7 Robustness . . . 102
3.8 Conclusion . . . 103

4 The effects of mandatory mortgage advice 105
4.1 Introduction . . . 105
4.2 Background . . . 108
4.2.1 Product selection . . . 109
4.2.2 The role of advice . . . 110
4.2.3 Mortgage Market Review (MMR) . . . 111
4.3 Data . . . 112
4.3.1 Sources . . . 112
4.3.2 Estimation periods . . . 113
4.3.3 Summary statistics . . . 113
4.3.4 Outcome variables . . . 114
4.4 Method . . . 115
4.4.1 Difference-in-difference matching . . . 115
4.4.2 Propensity score model . . . 119
4.5 Results . . . 125
4.5.1 ATT estimates . . . 125
4.5.2 Robustness . . . 129
4.6 Conclusion . . . 129

5 Conclusion 131

Bibliography 131

A Appendix for Chapter 2 141
A.1 Additional robustness checks . . . 141
A.1.1 Cost specification . . . 141
A.1.2 Sample representativeness . . . 142
A.1.3 Are choices better than random? . . . 142
A.2 Additional figures and tables . . . 143

B Appendix for Chapter 3 151
B.1 Extra materials 151

C Appendix for Chapter 4 155
C.1 Parallel trends 157
C.2 Findings and robustness checks 160

List of Tables

Table 2.1 Sample descriptive statistics: Demographic 33
Table 2.2 Dominated choices: headline results 44
Table 2.3 Estimated marginal effects 48
Table 2.4 Effects of changing consideration sets on in dominance rates of intermediated transactions 55
Table 2.5 Comparison of mean results in the very prime and full samples 66
Table 3.1 Sample descriptive statistics 76
Table 3.2 Out-of-sample performance: mean product characteristics 89
Table 3.3 Estimated marginal effects on borrower type and on product demand across types, in pp 92
Table 3.5 Summary inattention measures, split by type 94
Table 3.6 Average marginal effects on P(Attention), by type 96
Table 3.7 Marginal effects and willingness to pay for familiar banks, split by type 98
Table 3.8 Summary of average changes under full attention relative to the baseline 100
Table 4.1 Summary statistics for the main demographic variables . . 114
Table 4.2 Summary statistics for the main outcome variables 116
Table 4.3 Predictive power of the propensity score model 124
Table 4.4 Effects of the MMR advice requirement on previously non-advised 126
Table A.1 Selected summary statistics: Population-sample comparison 144
Table A.2 Regression results for main borrower characteristics . . . 145
Table A.3 Distributions of annual excess cost (£pa) under different time horizons 146
Table A.4 Effects of sample re-weighting on selected variables and results 146
Table A.5 Robustness of effects to alternative definitions of strong dominance 149
Table B.1 Estimated parameters from the baseline model: Full . . . 152
Table B.2 Estimated marginal costs and demand elasticities 153
Table B.3 Alternative specifications in a single type model 154
Table C.1 Covariates for the propensity score, Pr(Advice) 156
Table C.2 Effects of the MMR advice requirement on previously non-advised, ATT 160
Table C.3 Sensitivity of the ATT to different bandwidths 161
Table C.4 Sensitivity of the ATT to alternative lengths of the MMR implementation period (excluded from estimation) 161

List of Figures

Figure 2.1	Variation in borrowing costs by product type	36
Figure 2.3	Distribution of excess costs from dominated choices	45
Figure 2.5	Dominance rate and excess cost split by demographics	46
Figure 2.6	Variation in marginal effects of income, age and credit score on dominated choices	51
Figure 2.7	Propensity to choose lenders with an existing relationship	57
Figure 2.8	Proportion of dominated mortgage transactions, by lender size category	59
Figure 2.9	Frequency of dominating other lenders, by size category	61
Figure 2.10	Role of mortgage price components(PC) in excess costs	63
Figure 2.12	Maximum proportion of dominated choices explainable by tastes for lender characteristics	63
Figure 3.1	Out-of-sample performance: Lender market share prediction accuracy	89
Figure 3.2	Distribution of predicted probabilities of Type 2, (σ_{i2})	91
Figure 3.3	Distribution of the predicted P(Attention)$_{il}$	93
Figure 3.5	Distribution of P(Attention)$_{il}$, by familiarity	95
Figure 3.6	Distributions of predicted probabilities, by lender familiarity and type	97
Figure 3.8	Impacts of the counterfactual simulation with full attention	99
Figure 4.1	Use of regulated advice over time	112
Figure 4.3	Propensity score densities, by group	125
Figure 4.4	Effect of advice requirement on fixed rate periods	127
Figure A.1	Distribution of excess costs incurred due to dominated mortgage choice	144
Figure A.2	Observed dominated choice rates vs predicted under random mortgage choices	147
Figure B.1	Distribution of predicted type by demographic characteristics	151
Figure C.1	Parallel trends in cost outcomes, by quartile of propensity score	158
Figure C.3	Parallel trends in product outcomes, by quartile of propensity score	159

1 Preface

Mortgages have been central to household finance research ever since the seminal publication by Campbell (2006) on mortgage refinancing inertia among US households, which arguably founded the field. The importance of understanding mortgage choices in the literature is hardly surprising, given that they are by far the largest liability that most households have in their lifetime and can have significant consequences for households' decisions about other financial products, housing and consumption more broadly.

There has been a lot of research on mortgage decisions in recent years, focusing especially on questions of mortgage pricing (Allen, Clark and Houde, 2014a; Allen, Clark and Houde, 2014c; Alexandrov and Koulayev, 2018; Bhutta, Fuster and Hizmo, 2019; Liu, 2019), intermediation and advice (Woodward and Hall, 2012; Guiso et al., 2018; Mysliwski and Rostom, 2018; Belgibayeva and Majer, 2018; Robles-Garcia, 2019; Foà et al., 2019) and the timing of remortgaging (Agarwal, Driscoll and Laibson, 2013; Agarwal, Rosen and Yao, 2015; Keys, Pope and Pope, 2016; Andersen et al., 2018).

This thesis makes multiple contributions to this literature. Chapter 2 provides new evidence on price dispersion and proposes a new way to measure it in the context of horizontally differentiated products with multi-dimensional pricing like mortgages, by focusing on strictly dominated choices. Chapter 3 delves deeper into understanding borrower choices that lead to price dispersion, such as limited attention and brand loyalty. Finally, Chapter 4 provides a novel perspective on the effects of advice in the mortgage market by focusing on the introduction of (effectively) compulsory pre-transaction advice for this market in the UK. All three chapters use a uniquely granular combination of datasets about borrowers, lenders, and products in the UK mortgage market that has not been previously used for research in economics or finance.

Furthermore, the methods I develop in this thesis — such as detecting strictly dominated choices, modelling limited attention and measuring the effects of brand familiarity — have wider applicability for understanding consumer behaviour in other markets beyond mortgages.

The rest of this chapter provides a brief summary of the methods and findings of each of the substantive chapters.

CHAPTER 2 develops a nonparametric methodology for assessing the quality of households' product search in the UK mortgage market without any assumptions about preferences over horizontally differentiated product characteristics. Using a uniquely detailed combination of lending transaction reports, product information and credit bureau data for 700,000 UK households who took out a mortgage between January 2015 and July 2016, I conduct pairwise multidimensional comparisons between all products for which each borrower was eligible. This allows me to identify choices where the chosen mortgage was strictly dominated in terms of pricing by an available alternative with the same non-price characteristics. I find that 30% of UK households in the sample chose dominated mortgages, paying £550 more per year (equivalent to 12.7% of borrowers' annual mortgage costs) on average as a result.

I further find that borrowers with characteristics typically associated with lower financial capability (lower income and credit score, old age) and greater time constraints tend to make dominated mortgage choices more often. Going to a lender with whom a borrower has an existing relationship, such as a checking account or a credit card, is also associated with increased probability of making a dominated mortgage choice.

The source of dominated products is not concentrated in a small number of suppliers; most lenders in the market originate non-trivial proportions of dominated loans. This finding suggests that dominated choices are driven primarily by inefficient search of the available options by borrowers and their brokers rather than a small number of 'bad' lenders with a captive customer base.

CHAPTER 3 explores the channels through which limited attention and brand familiarity affect consumer behaviour in the UK mortgage market using a unique dataset comprising transaction-level data, linked credit files and extensive data about firms' products, advertising and locations. I focus particularly on the way in which existing relationships with mortgage lenders through existing personal current accounts affect borrowers' attention and preference for products from these lenders.

I develop a structural discrete choice model with limited attention, in which the probability of considering each alternative is determined by its characteristics, and the borrowers make a choice out of the resulting consideration set based on their preferences. I include factors related to lender familiarity (such as existing banking relationship, branch locations and advertising expenditure) in both the attention and preference parts of the model, which allows me to explore through which channel (if any) these factors influence choices. I allow there to be two latent borrower types that may have different attention and preference parameters. A particular borrower's probability of belonging to each type depends on his or her demographic characteristics.

I identify two distinct types of borrowers. The first one, Type 1, has characteristics commonly associated with lower financial sophistication (e.g. lower income or worse credit history) and tends to be less price-sensitive and more inattentive. Borrowers in this type are also much less likely to consider unfamiliar lenders than familiar ones. The second type of borrower, characterised by higher income, creditworthiness and education, also displays inattention but a lesser extent, and their attention is less influenced by existing relationships with lenders. Both types, however, show a strong preference for familiar lenders when choosing amongst the options they consider. Their behaviour is consistent with being willing to trade off an equivalent of 5% of their post-tax annual income for the benefit of taking out a mortgage with a lender with which they already have a current account.

I also simulate the effects of a policy that makes all borrowers pay attention to all available options. It has only a small effect on market shares and prices, although borrowers' welfare (primarily for Type 1) does increase on average as a result.

CHAPTER 4 is co-authored with Jeroen Nieboer at the FCA and London School of Economics. We exploit a change in regulation due to the UK Mortgage Market Review (MMR) in 2014 to estimate the effects of mandatory advice on mortgage choices. The proportion of advised mortgage transactions for home buyers (as opposed to those refinancing an existing property) increased from three quarters pre-MMR to effectively the whole market after MMR came into force. To estimate the effects on the borrowers 'treated' by the policy change (i.e. those 'brought' into getting advice) on various consumer outcomes, we construct a unique dataset that links the entire population of residential mortgage originations before and after the MMR with credit reference agency data and detailed mortgage product data. We use difference-in-difference matching on the repeated cross-sections (Blundell and Costa Dias, 2009) to estimate the counterfactual outcome for mortgage borrowers that did not use mortgage advice before the MMR, but were effectively forced to do so afterwards.

We find that the introduction of mandatory advice led to a significantly increased concentration of borrowers in one product type, 2-year fixed mortgages. It also made borrowers more likely to choose longer mortgage terms and use independent brokers. However, we document little to no effect of mandatory advice on various metrics of borrowing cost, despite the substantial rise in the use of brokers.

2 Strongly dominated mortgage choices[1]

2.1 Introduction

Mortgage debt is the largest single liability for households in many developed countries.[2] Choosing a mortgage is therefore one of the most important financial decisions households need to make. It is hardly a surprise that policy makers around the world are interested in how effectively households shop around for their mortgages and what borrowing costs they incur as a result.[3] Academic research on price dispersion for comparable mortgage products and its drivers is relatively scarce, however, and the existing papers are often constrained by data limitations about borrower preferences, product characteristics and the alternatives realistically available to borrowers of different credit risk (Woodward and Hall, 2012; Allen, Clark and Houde, 2014a; Allen, Clark and Houde, 2014c; Alexandrov and Koulayev, 2018; Bhutta, Fuster and Hizmo, 2019).

In this paper I contribute new detailed evidence on households' ability to search for good deals in the high-stakes setting of the mortgage market. To do this, exploiting uniquely granular combination of lending transaction reports, detailed product listings with lenders' eligibility criteria and credit files for nearly 700,000 UK households that took out a mortgage between January 2015 and July 2016. I also develop and implement a new methodology for measuring the effectiveness of households' product choices in the mortgage market or other circumstances with multidimensional product prices and heterogeneous household preferences. The proposed meth-

1 I would like to thank Richard Blundell, João Cocco, John Campbell, Ian Preston, as well as the participants in the 2018 CFPB Research Conference, 2019 CEPR Household Finance Workshop and 2019 Royal Economic Society 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 at the FCA as this research would not have been possible without their support.
All views (and any errors) in this paper are my own and do not represent the position of the FCA.

2 For instance, mortgage debt accounts for over 80% of household liabilities in the UK and around 70% in the USA and Canada. Data from the Bank of England *Money and Credit Statistics January 2018*, UK Student Loans Company *Student Loans Balance FY16/17*, Federal Reserve Bank of New York *Household Debt and Credit Report Q1 2018* and Bank of Canada *Credit Conditions June 2018*.

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

odology — dominance analysis — identifies 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 approach makes it possible to detect consumers who are likely to not be searching the market effectively without making any trade-offs between different price components or needing additional assumptions about household preferences (for instance, their degree of risk aversion or discount rate). First, using extensive data about individual borrowers and eligibility criteria for different mortgages enables me to directly identify the alternatives for which the borrower satisfied observable lending requirements (on factors including loan amount, loan-to- value ratio, property type and value, borrower income, employment status, age, history of credit impairments and credit score.) Second, only comparing products with identical substantive non-cost features ensures that households are not just paying extra to get a particular type of mortgage they prefer (e.g. a longer fixed rate). 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 a high relocation risk.

I find that nearly one in three UK consumers (29.9%) in 2015/16 chose mortgage products that were dominated by apparently available alternatives. The average cost savings forgone as a result of a dominated choice ('excess cost') are £550 ($750) per year, equivalent to 12.7% of borrowers' annual mortgage costs, over the course of their initial teaser period alone. For over 17% of all borrowers (around 60% of those with dominated products), choices were strongly dominated, that is they involved excess costs above both £250 ($340) per year and 5% of the borrower's total annual mortgage costs. The frequency of dominated choices are similar among borrowers who use mortgage brokers and those who apply directly to the lender.

Even among the very prime borrowers—those with relatively high incomes, low loan-to-value ratios, and no factors that could typically complicate a mortgage application (self-employment, unusual property, etc)—the rate of dominated choices is 26%. Those borrowers would be less likely to be investigated closely by the lenders and are much less likely to have their options restricted by idiosyncratic lending criteria (if any) not recorded in the standardised data used in the dominance analysis. 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 overall results.

I do, however, find substantial heterogeneity in the frequency of dominated choices more generally. Consumers are more likely to choose dominated mortgages 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 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 related to alternative demands on borrowers' time (such as the number of children or the complexity of the housing transaction) are also associated with more frequent dominated choices.

Although some behavioural models suggest that borrowers might make poor choices because they focus on headline prices and overlook other fees and charges (Bar-Gill, 2012; Bordalo, Gennaioli and Shleifer, 2016), this is not a major driving force for dominated mortgage choices. 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.

I find that decision-makers tend to gravitate towards familiar alternatives. Brokers 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.[4] 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 found 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 consumers who did not use a broker took out a mortgage with one of the lenders with whom they already had financial product (current account, loan or credit card). This was despite those lenders accounting, on average, for less than 20% of products for which the relevant consumer was eligible. Even 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%).

RELATED LITERATURE This paper contributes to several strands of literature in household finance. It most directly fits with behavioural finance research that investigates dominated choices in other household finance contexts, such as Sinaiko and Hirth (2011) and Bhargava, Loewenstein and Sydnor (2017) for insurance, and

4 It is not the case that all brokers used the same 5 lenders, however.

Elton, Gruber and Busse (2004) and Egan (2019) for investment products. To the best of my knowledge, there has been no comparable work on the mortgage market prior to this paper. This is an important extension as ex ante one might expect households to search more and make better decisions in the high-stakes setting of managing the cost of their largest financial liability. Other novel contributions of my research to this literature include being able to identify dominated choices in the UK mortgage market as a whole rather than the more narrow contexts that involve choices between pairs of products. The dominance approach I develop also provides a flexible way to handle heterogeneity in product features, lenders' acceptance criteria and borrower characteristics, which are important features of many other household finance settings.

My uniquely granular dataset allows me to explore borrower and product characteristics that are associated with dominated choices in greater detail than has been possible previously. These findings on factors linked to lower-quality decisions also contribute to the broader research in behavioural finance and economics, by providing additional insights on the roles of salience (Bar-Gill, 2012; Bordalo, Gennaioli and Shleifer, 2016; Gabaix and Laibson, 2006; Gurun, Matvos and Seru, 2016) , familiarity bias (Coval and Moskowitz, 1999; Huberman, 2001; Pool, Stoffman and Yonker, 2012) and financial literacy (Agarwal et al., 2009; Lusardi and Mitchell, 2014). I find several potential important influences on households' attention and search costs that can be investigated in future research.

More broadly, borrowers' mortgage decisions are an important topic in household finance. There is already extensive research on some aspects of the mortgage market, such as the timing of refinancing (Andersen et al., 2018; Agarwal, Driscoll and Laibson, 2013; Agarwal, Rosen and Yao, 2015; Campbell, 2006; Keys, Pope and Pope, 2016, ,and many others) and the choice between fixed and variable rate contracts (e.g. Campbell and Cocco, 2003; Coulibaly and Li, 2009; Van Hemert, 2010). It has been much less common for papers to examine how borrowers make choices between different lenders or contracts of different costs, to a large extent due to the challenges posed by adverse selection and limited availability of data on mortgage offer prices. There are already some papers that consider these questions, such as Woodward and Hall (2012), Allen, Clark and Houde (2014a), Allen, Clark and Houde (2014c), Alexandrov and Koulayev (2018), and Bhutta, Fuster and Hizmo (2019). However, I extend the existing work in a number of ways. First, I construct a unique dataset that has not been used for mortgage research to date, and makes it possible to investigate the UK market with a different set of institutional features. Second, I document significant price dispersion and ineffective borrower search in both intermediated and non-intermediated transactions, whereas the existing literature largely focuses just on intermediated mortgages and the role of broker price discrimination as a driver of price dispersion. Third, I am able to construct individual choice sets that take account of borrowers' risk and mortgage product availability at the time of their choice, reducing any likely distortions from adverse selection or unavailable products. Finally, by applying the high standard of strictly dominated choices to detect ineffective search by households, I produce a non-parametric measure of market frictions that is robust to different assumptions about household preferences.

The rest of the chapter is structured as follows. Section 2.2 provides institutional background on the UK mortgage market. Section 2.3 describes the datasets and their representativeness. Section 2.4 sets out the methodology for identifying dominated choices and outlines the headline results. Section 2.5 investigates how the probability of making a dominated choice and the associated excess costs vary with borrower circumstances and demographic characteristics. Section 2.6 presents the initial findings on the role of lender familiarity in dominated choices of households and their brokers. Section 2.7 describes the characteristics of dominated mortgage products and of their suppliers. Section 2.8 summarises the main robustness checks and, finally, section 2.9 concludes. Additional robustness checks and more detailed tables and figures can be found in Appendix A.

2.2 Background on the UK mortgage market

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.

2.2.1 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%[5]) involve fully variable rate mortgages that typically track the Bank of England base rate for the contractual term. Variable contracts also offer attractive terms in the first 2-5 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 are commonly referred to as 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 (e.g. Agarwal, Rosen and Yao, 2015; Keys, Pope and Pope, 2016) and Denmark (Andersen et al., 2019). In fact, according to the FCA (2018a), 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, 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.

2.2.2 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 (e.g. borrower loan-to-value ratio (LTV)), features (e.g. length of the fixed-rate period and the degree of flexibility in repayments) and the corresponding set of prices, which normally include an

5 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.

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.[6]

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 from 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.

2.2.3 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 likely 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 intermediary rather than going to a lender directly. Intermediaries provide advice as

6 The sample mean for the size of the 'available product set' as defined in definition 1 in section 2.4.1.

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 is 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.

2.3 Data

2.3.1 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 (e.g. date, identity of the lender, loan amount, interest rate type), of the property used as collateral (postcode, purchase price) and of the borrower (e.g. age, date of birth, gross income, adverse credit events in the past).[7]

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

7 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.

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. Credit files are also a source of information about the existing relationships (e.g. through personal current accounts) the borrower has with the different lenders at the time of the application.

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 is 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. Finally, I use the quarterly bank branch location data from Experian GOAD and the borrower's historical postcode data from their credit files to calculate the distance between borrower's homes at the time of the application and each lender's closest branch.[8]

2.3.2 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).[9]

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

8 This calculation is discussed in more detail in Chapter 3, where location data are more central to the analysis.

9 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.

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 A.1.2.

The final sample used for all results in this paper unless otherwise stated contains 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 (Best et al., 2015; Cloyne et al., 2019; Benetton, 2019; Robles-Garcia, 2019; Liu, 2019), 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.

2.3.3 Descriptive statistics

2.3.3.1 *Demographic overview*

Panel A in Table 2.1 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

Table 2.1: Sample descriptive statistics: Demographic

	Mean	σ	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
PANEL A: Demographics					
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				
PANEL B: Borrowing costs					
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$d$ (£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$e$ (years)	3.15	1.68	2.00	2.00	5.00
=1 if fixed rate mortgage	0.921				
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. (d) Initial mortgage payment is the minimum monthly interest and principal repayment the borrower has to pay during their introductory deal period. (e) Length of the fixed rate or the introductory discounted variable rate discount.

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.[10] 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.

2.3.3.2 *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).

Panel B in Table 2.1 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

10 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.

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 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).[11]

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, 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

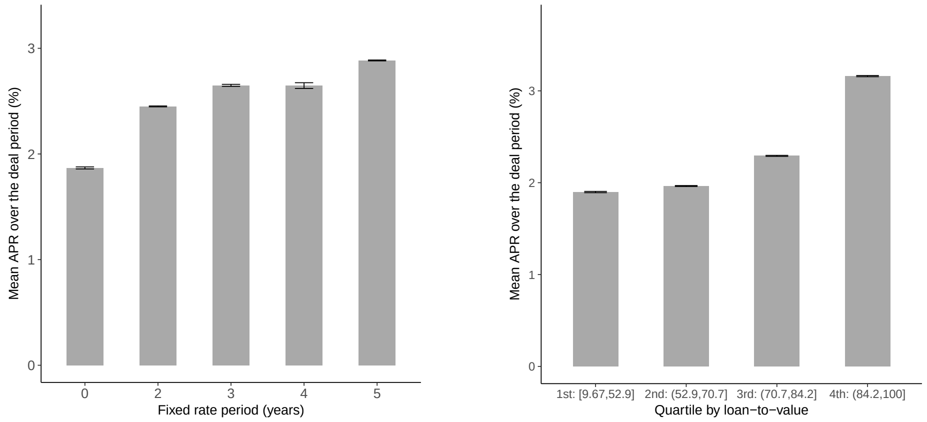
11 In line with standard practice and Financial Conduct Authority rules, the APR for borrower i with mortgage j over a period of T months is a rate that sets the present value of all mortgage cash flows to zero, i.e. the solution to:

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

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.

Figure 2.1: 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 based on unclustered standard errors.

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, Figure 2.2a 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 determines 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 2.2b 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.

2.4 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 is of considerable policy interest. It is also of academic interest as mortgage choices are a high-stakes environment with large implications for household budgets. 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 (2019) 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.

2.4.1 IDENTIFICATION OF DOMINATED CHOICES

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

Definition 2.1 (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.*

The eligibility criteria come from two sources: the explicitly stated criteria in Moneyfacts (relating to borrower status —e.g. a first-time buyer or not self-employed — or minimum and maximum values of characteristics such as age, household income, property value, loan amount, or loan-to-value ratio) and additional rules derived from lender behaviour in the data (minimum credit score, and maximum loan-to-income and LTV ratios accepted for each product).

The minimum time on the market requirement is added to avoid potential distortions from lenders' flash deals: heavily discounted mortgage offers that are offered only briefly or in limited numbers but allow 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 2.2 (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 2.3 (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.[12] 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 (e.g. Campbell and Cocco, 2003; Coulibaly and Li, 2009; Van Hemert, 2010).

Definition 2.4 (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 $p_{jd} \neq p_{kd}$ $\exists d = 1, ..., D$.*

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 promo-

12 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.

tional 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 2.5 (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.

2.4.2 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 (e.g. 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 11.

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 2.6 (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.[13]

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.[14]

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 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 subgroup 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.

13 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.

14 A potential conceptual downside of using the average rather than the best possible product is the risk of niche cases where making a better choice increases the excess cost. Consider, for instance, a borrower choosing the worst product out of a set with prices {11,9,1} (average excess cost of 6) compared to choosing the second-worst product (average excess cost of 8). However, this does not occur much in practice: the excess cost obtained with the proposed method are very highly correlated those obtained with the frontier method (correlation coefficient 0.968).

Definition 2.7 (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 2.8 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.

2.4.3 Results on the 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 2.2. 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 2.4a and 2.4b 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

Table 2.2: 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 2.6 in section 2.4.2) 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.

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.

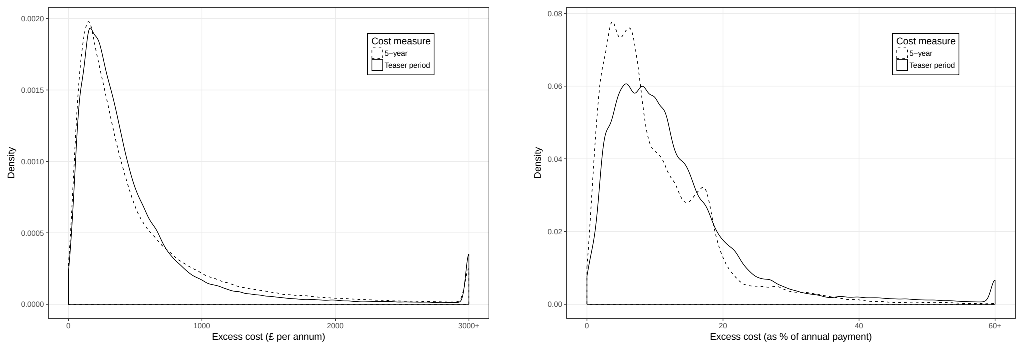
2.5 Demographic variation

2.5.1 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 domin-

Figure 2.3: Distribution of excess costs from dominated choices

(a) Absolute (b) Relative



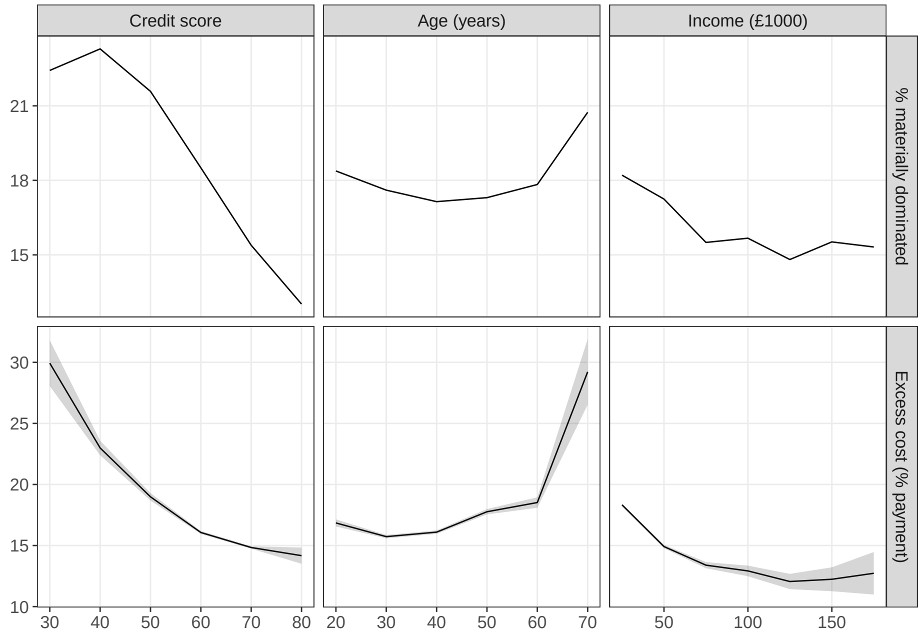
Note: All distributions of excess costs are for the subset of borrowers with dominated products. Where excess costs are calculated on the basis of remortgaging after five years, the borrower remains on the reversion rate until the end of the five-year period if their fixed rate or discounted rate deal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

ated choices are present to some degree across all types of borrowers, demographics appear to play a role in shaping propensity towards such choices.

Figure 2.5 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 three 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 a proportion of the overall mortgage payment is nearly twice as large.[15] 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, e.g. Finke, Howe and Huston (2017) for an overview), with middle-aged individuals performing best of all.

15 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.

Figure 2.5: 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.

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.

2.5.2 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.[16]

16 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

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.2), 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 (e.g. 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 difference-in-differences matching estimation to a recent UK policy change. The results reported later in the paper are robust to using more detailed fixed effects, e.g. additional fixed effects for the size and type of the intermediary firm (if any) used by the borrower.

2.5.3 Marginal demographic effects

To simplify interpretation of quadratic terms and logistic regression coefficients, Table 2.3 presents average marginal effects obtained from regressions above, for the probability and excess cost. The table also report marginal effects which are scaled by one standard deviation of the explanatory variable to illustrate relative economic significance.[17]

with large costs and greater economic significance. Qualitative findings are robust to alternative specifications of strong dominance. See section 2.8.2

17 Regression results before transformations are provided in Table A.2.

Table 2.3: Estimated marginal effects

	Marginal effect[a]	Effect of 1σ change
PANEL A: Effects on P(Strongly dominated choice) (in pp)		
Income (in £1000)	−0.19	−5.17
Loan value (£1000)	0.05	6.32
=1 if joint mortgage	−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[b]	0.01	3.45
=1 if new-build property	3.61	
=1 if self-employed	4.78	
=1 if non-standard property[c]	2.38	
Credit score	−0.28	−2.36
Balance on unsecured debt (% income)[d]	0.08	2.28
Loan-to-value ratio (LTV, %)	0.24	4.84
Observations	647,758	
PANEL B: Effects on excess cost (in pp of annual mortgage cost)		
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
Number of available products[b]	−0.0002	−0.06
=1 if self-employed	1.9019	
Credit score	−0.1744	−1.48
Balance on unsecured debt (% income)[d]	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 specifications (2) and (4) in Table A.2, respectively. All reported effects are statistically significant at 1% level. Standard errors are computed using the delta method and clustered on year and postcode area. (b) Number of products in borrower's comparable set of available products as defined in Definition 2.3. (c) 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. (d) 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. (e) 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.

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 (e.g. financial capability and available time) and the complexity of choice they are facing (e.g. number of options and application complications).

2.5.3.1 *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 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 2.6 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' are 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.

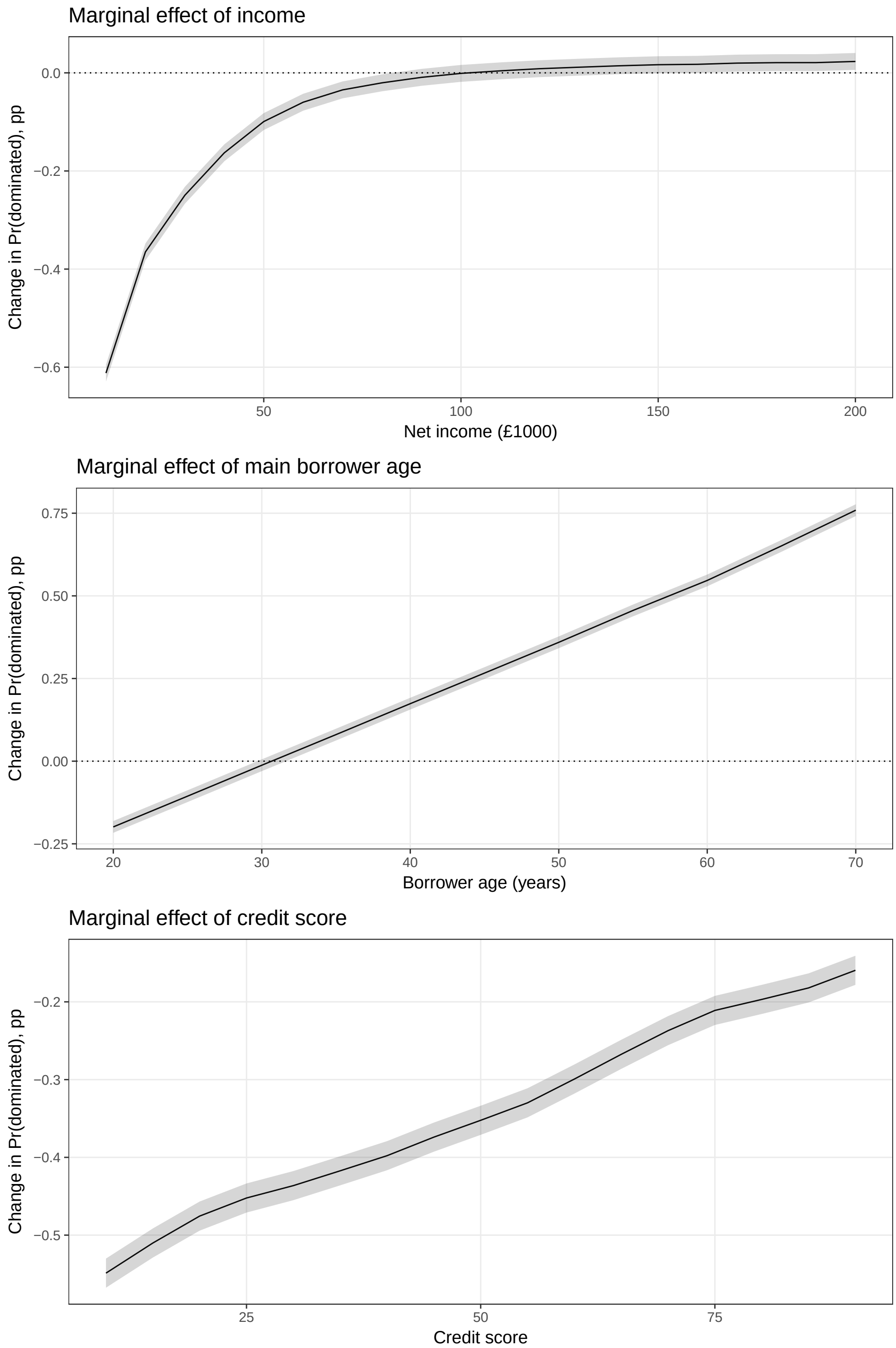
income Household income continues to have a strong negative association with the likelihood of a strongly dominated choice after controlling for other demographic factors. The average marginal effect of income is, in fact, considerably larger than univariate trends might suggest, with a £1,000 increase in household income being associated with 0.2pp reduction in dominance rate on average. As Figure 2.6 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 of products for which 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 countervailing 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 2.3).

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

Figure 2.6: 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.2 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.

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 2.6 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 for the quality of financial decisions such as credit scores, which would also capture some of the age-related changes in 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 sophistication, and also may only be able to 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.

2.5.3.2 *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.[18]

2.6 Familiarity

Having outlined the borrower and product characteristics that are associated with dominated choices, I also consider whether 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).

2.6.1 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 2.4, I show how the frequency of choices that are dominated by *alternatives within the restricted consideration set* changes with these restrictions.[19]

18 As discussed in section 2.8.1, the general findings on dominance and excess cost are robust to excluding borrowers in these unusual circumstances.

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

Table 2.4: 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.

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 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 the 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 with a lender allow intermediaries to learn about their eligibility requirements. The intermediary may also believe that their existing relationship with a lender 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.

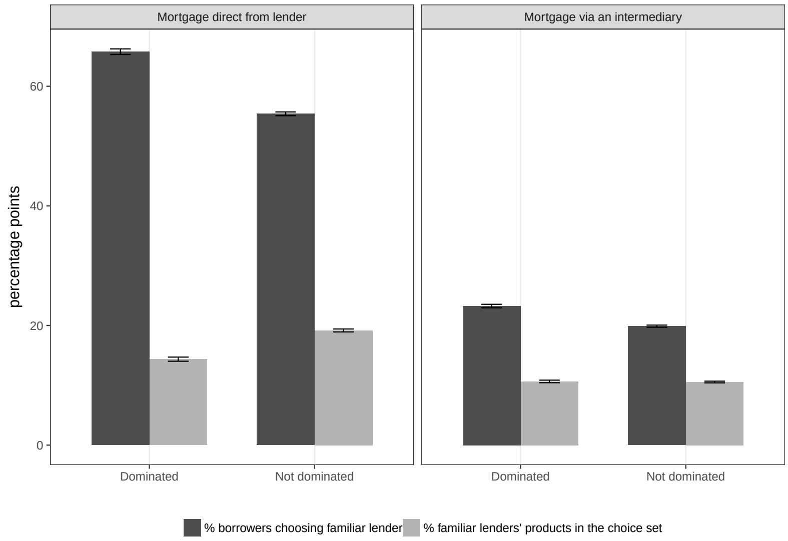
As can be seen in Table 2.4, restricting intermediary's consideration set only to lenders that supply 75% of their business has by far the largest effect on its own, nearly halving the rate of choices that are dominated within the consideration set to under 16%. When familiarity is combined with all 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.

2.6.2 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 2.7 shows that the proportion of products from familiar lenders varies between whether the the borrower used an intermediary and whether they made a dominated

Figure 2.7: 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).

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 intermediated 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 2.5.2 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.5). 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 per year, 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.[20]

2.7 Product and supplier characteristics

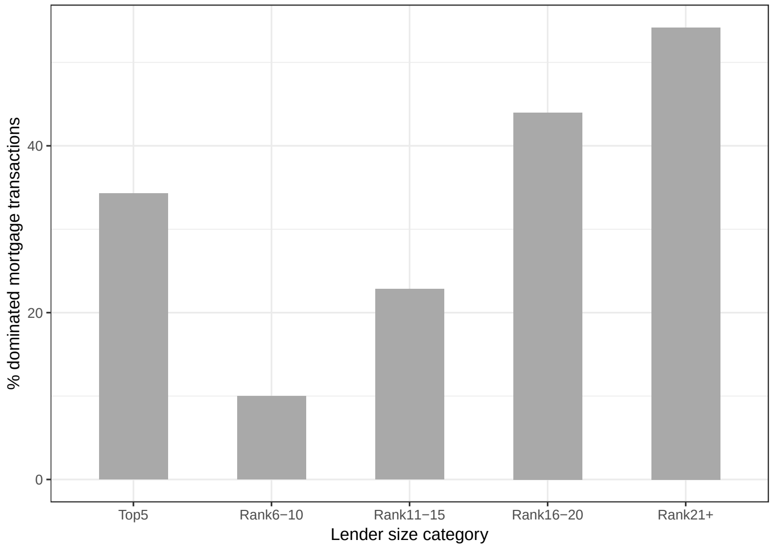
Finally, in this section I provide further descriptive evidence on dominated choices, focusing on the characteristics of products and suppliers that are associated with more frequent dominated choices by borrowers. First, I describe the patterns in lenders that supply more frequently dominated products, or products that frequently 'dominate' others. Second, I explore and fail to find evidence for the hypothesis that the relative salience of different price elements of mortgage product is a material driver of dominated choices. Finally, I investigate the extent to which the observed dominated choices can be explained by borrowers' strong preferences for characteristics of mortgage suppliers (e.g. an existing relationship through other products or closer branches) rather than characteristics of individual products alone.

2.7.1 Supply of dominated products

The dominance measure used in Section 2.4 is based on individuals and alternatives they face. Hence, the same mortgage product may be dominated for some borrowers who choose it, but not others. For discussing the broader market dynamics of who

20 For borrowers with dominated products, the size of their excess cost does not materially differ between those who used a lender they knew or one without an existing relationship. This simple calculation also assumes that all borrowers without dominated products have excess costs of zero, and the only source of the 'familiarity premium' is the increase in the likelihood of dominated choices.

Figure 2.8: Proportion of dominated mortgage transactions, by lender size category



supplies dominated products and where the strictly superior alternatives come from, we need to aggregate those borrower-based measures to firm level.

Suppliers of dominated products For each lender (or a category of lenders) j, I define j's proportion of dominated mortgage transactions as the ratio of the number of instances where a borrower chooses j's mortgage product which is dominated by another alternative to the total number of times j's mortgage products are chosen.

This ratio varies considerably between suppliers. For one half of the lenders, the proportion of dominated mortgage transactions lies between 18% and 62%, and the average is 43%. Extremes are very uncommon. Every lender, big or small, had at least some mortgage transactions where their product was dominated by other alternatives available to the borrower. There are also only a couple of very small suppliers whose products were dominated for every borrower who chose them.

Figure 2.8 shows that there are systematic differences in the proportion of dominated mortgage transactions depending on lender size.[21] Perhaps predictably, suppliers whose mortgage loans are dominated most often tend to be small, potentially reflecting both aversion to more expensive products among at least some borrowers and also suppliers' higher costs due to lack of access to lending economies of scale.

The proportion of dominated mortgage transactions does not decline monotonically as lender size increases which is what one would expect if only those two factors

21 Here and henceforth, lender size bands are based on lender's rank by total mortgage lending volume between January 2015 and June 2016.

were at play. The lowest proportion of dominated mortgages of below 10% is seen among the second-tier lenders ranked between 6 and 10 by size, with smaller banks ranked 10-15 not too far behind. The top 5 largest lenders buck the trend with over 30% of their mortgage transactions dominated by other alternatives. This might suggest that a getting a mortgage from a Top 5 bank offers additional benefits that borrowers value and trade-off against the product specific characteristics captured by the dominance measure. Alternatively, it is possible that the largest lenders are a natural focal point for borrowers who are less able or willing to search (and thus to identify the best priced deals).

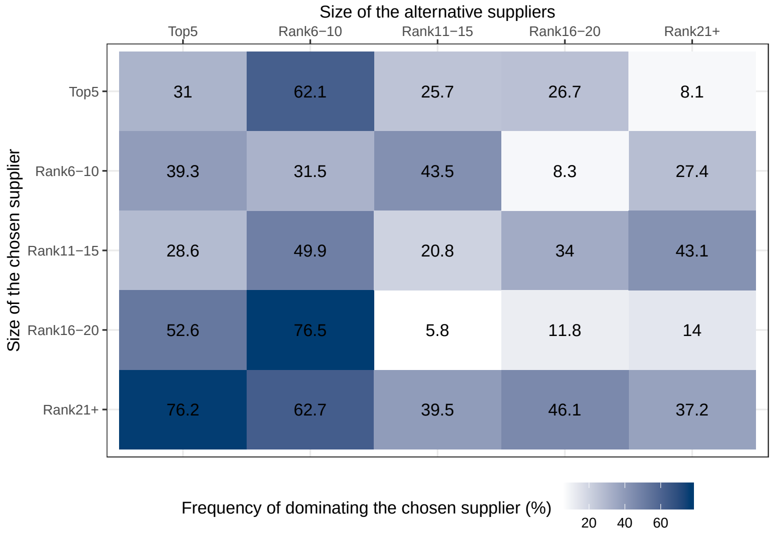
suppliers of dominating products It is also useful to consider which supply the *dominating* products (i.e. the available alternatives that were strictly superior on price and product characteristics to the chosen dominated product). This is also not a straightforward question, as there were two or more dominating alternative mortgages for more than a half of the borrowers who made dominated choices.

To compare lenders' relative importance as a source of dominating alternatives, I calculate two measures. First, lender k's 'frequency of dominating other products' is a simple ratio of the number of dominated mortgage choices in which lender k's mortgages are among the dominating alternatives to the number of all dominated choices in the sample. Second, k's frequency of dominating lender (or a category of lenders) j is the ratio of dominated choices where j's mortgage was chosen and it was strictly dominated by at least one product from k to the total number of dominated choices of j's mortgages.

If the majority of dominated choices were driven by dominating alternatives from the same one or two lenders, one would expect the frequencies of dominating other products to be very high (approaching 1) for those suppliers, and be near zero otherwise. What we see in practice is very different. On average, each supplier of a dominating alternative appears in fewer than 5% of dominating choices. Even for the 5 most frequent suppliers of dominating alternatives, the proportion of dominated choices in which they appear is below 24%. All this strongly suggests that suppliers of dominating products vary from one borrower's case to the next, and no single 'superior' lender is responsible for driving the majority of the dominated choices.

Figure 2.9 illustrates the complex nature of dominance relationships between lenders further. Out of all dominated mortgages originated by the largest 5 lenders, at least one dominating product came from another Top 5 firm in 31% of the cases, from lenders ranked between 6 and 10 in 63% of the cases, and from smaller lenders ranked 11 and below in around 40% of the cases.

Figure 2.9: Frequency of dominating other lenders, by size category



For the choices where a mortgage from a lender in the 6-10 bank was dominated, one of the dominating alternatives also came from a Top 5 lender in more than a third of the cases even though those lenders have much higher proportion of dominated mortgage originations on average. In general, no single group of lenders is solely responsible for dominating the other lenders' mortgages. Instead, when a borrower's choice of a particular lender is dominated, it is due to a mix of better products being available from different lenders at the time of that particular choice. In fact, it is not uncommon for the same 2 lenders to dominate each other's products depending the borrower's circumstances (e.g. the LTV band).

These stylised facts strongly suggest that the supply of dominated products is more complicated than some 'bad' lenders that are consistently dominated by 'better' maverick ones. Dominated choices, instead, are a product of a complex match between the alternatives on offer by different suppliers and the individual borrower's circumstances and abilities to find them.

2.7.2 Salience of price features

A hypothesis often seen in behavioural literature is that consumers 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) and borrowers appear to be more cost-sensitive to products with fees (Liu, 2019). If true, this behaviour could result in a 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 in which 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 2.10, 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% of 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). 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.

2.7.3 Contributions of additional supplier characteristics

I also explore to what extent dominated choices could, at least in principle, be explained by extremely strong preferences for other choice dimensions other than price or product characteristics. I rank the borrower's chosen product and all of its dominating alternatives on four additional attributes of the choice situation and the lenders: (a) whether the borrower has an existing relationship (current account) with the lender, (b) the distance to the lender's closest branch from the borrower's old address, (c) the maximum loan-to-value ratio the lender accepts for the product, and the lender's customer satisfaction score.[22]

22 Based on the results of a survey by Which? reported in *Advice Guides: Compare mortgage lenders* in May 2016.

Figure 2.10: Role of mortgage price components(PC) in excess costs

(a) Total contribution to excess cost

(b) Strictly positive cost contribution, by PC

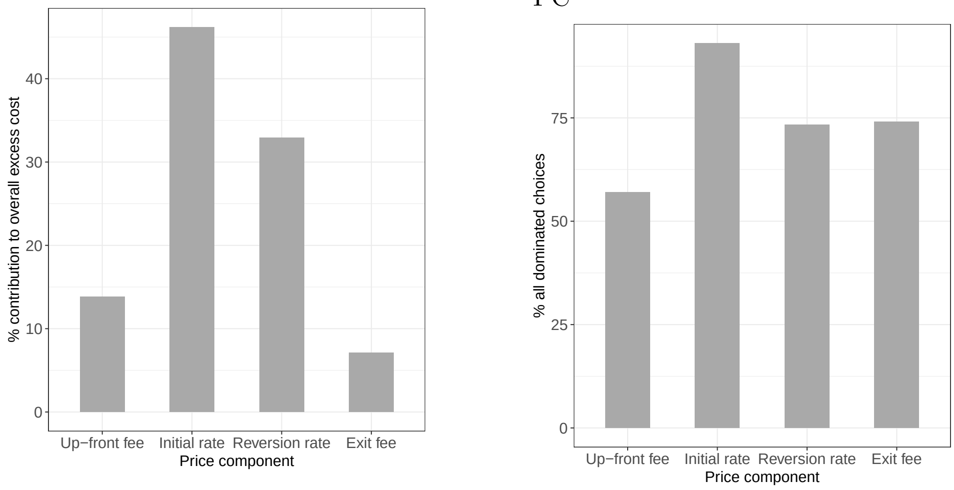
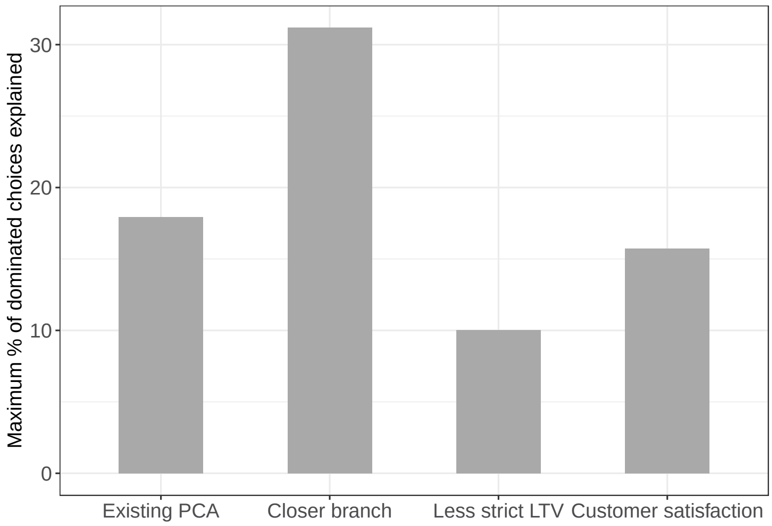


Figure 2.12: Maximum proportion of dominated choices explainable by tastes for lender characteristics



A borrower's dominated choice is considered 'potentially explainable' by one of four criteria if the chosen product is even slightly better on this criterion than all of the dominating alternatives. In this case, it is possible that strong preferences for the comparative advantage offered by chosen product (e.g. for dealing with the lender with which one already has an account) might have outweighed the strictly lower costs of the available dominating alternatives.[23]

For some borrowers, the 'explanations' offered by such advantages will not be very plausible – it is extremely unlikely, for instance, that someone would choose to pay hundreds of pounds for being with a lender whose branch is 0.1mi closer. This very generous definition of 'potential explainability', however, makes it possible to obtain an upper bound on the proportion of dominated choices that could be, at least in principle, attributed to the additional measurable lender characteristics.

Figure 2.12 shows the maximum proportion of the observed dominated choices attributable to each new measure. Potential preference for closer branches could explain at most a third of dominated choices. The other individual criteria fare even worse - only up to 18% of the dominated choices are attributable to preference for using one's current account provider for a mortgage, up to 16% to preference for higher customer satisfaction ratings and up to 10% to perceived lender strictness as proxied by the maximum LTV ratio they accept for the product. To put it differently, lenders who offer dominating alternatives are not systematically worse than providers of the chosen dominated products on other non-price dimensions borrowers might care about.

Overall, in just over 50% of the dominated mortgage choices the chosen lender is slightly better than all dominating alternatives on *any* of the four criteria. This is the maximum proportion of dominated choices that could be explained by preferences for these additional characteristics. Nearly a half of the observed dominated choices cannot be attributed to borrower preferences, however extreme, for lender familiarity, proximity, customer service and the (observable) lower lending standards.

Therefore, additional explanations, such as the difficulty some borrowers have in identifying the available alternatives or comparing them, are needed to account for many of the observed dominated mortgage choices.

23 This can also be thought of as extending the definition of dominance in Definition 4 to include the four criteria above.

2.8 Robustness

The findings reported above are robust to a wide range of alternative assumptions. In this section I summarise two main robustness checks for the results in the paper: (a) calculating the dominated choice rates and excess costs for the subset of 'very prime' borrowers, who are least likely to face any eligibility restrictions, and (b) varying the definition of strong dominance for the demographic regressions in section 2.5. Additional robustness checks can also be found in section A.1 in this chapter's appendix.

2.8.1 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.[24] This is a very conservative set of assumptions, which restricts the sample to approximately 63,000 observations.

Table 2.5 compares the average dominance rates and costs incurred in this restricted prime sample to the full-sample results reported earlier. In general, the improvement in most of the metrics in the ultra-prime sub-sample is very small. For instance, the proportion of borrowers who buy dominated products is still at 26.1%, only 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

24 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.

Table 2.5: 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≤65%, loan-to-income ratio ≤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.

higher 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.

2.8.2 Definition of strong dominance

Regression specified in section 2.5.2 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.5 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 2.5.3 and 2.6 are largely unchanged across a range of alternative definitions. The model with any dominated choice as the dependent variable has the most deviations from baseline approach and lower significance of some of the covariates. This is unsurprising, 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.

2.9 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 reveals that around 30% of UK households between January 2015 and July 2016 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. There is also a clear tendency for both intermediaries and borrowers to focus more on familiar lenders. The findings on the high rate of dominated choices are robust to a wide range of robustness checks, including, importantly looking at the very prime consumer sub-group, whose options are least likely to be restricted by lenders' idiosyncratic lending criteria.

The dominance methodology developed in this paper offers a way of detecting poor consumer outcomes 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]

This research also provides initial evidence about the specific frictions in the search and product choices that consumers make in the mortgage market. Research in a more structural setting (see, e.g. Chapter 3) should be able to shed further

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 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.

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.

3 Better the lender you know? Limited attention and lender familiarity in UK mortgage choices[1]

3.1 Introduction

Price dispersion in the UK market for residential mortgages is considerable, and the vast majority of mortgage products are not exclusive to the lender's existing customers. Consequently, the returns to shopping around can be significant. Despite this, over 30% of UK customers with a personal current account and a mortgage, took out their mortgage with the lender that already provided their current account(FCA, 2018c). Moreover, among mortgage borrowers who did not use a broker, over a half chose a lender with whom they already had another financial product even though those 'familiar' lenders on average accounted for less than 20% of mortgage options the borrower could choose from (Chapter 2[2]).

The marketing literature has long been aware of the importance of brand loyalty in understanding consumer choices (Brown, 1953; Tucker, 1964; Oliver, 1999; Palmatier et al., 2006, and many others)[3]. There is also a more recent but growing body of empirical research in economics that documents brand inertia in health insurance (Handel, 2013; Heiss et al., 2016; Ho, Hogan and Scott Morton, 2017), banking (Honka, Hortaçsu and Vitorino, 2017), energy markets (Hortaçsu, Madanizadeh and Puller, 2017) and many others. However, with the notable exception of Allen, Clark and Houde (2019), the existing economic research has focused primarily on repeated choices of the same product rather than on cases where a relationship

1 I am very grateful to Richard Blundell, Ian Preston, David Laibson, Abi Adams, Peter Andrews, Thomas Hoe, Claudia Robles-Garcia and Patrick Coen for their thoughts, comments and suggestions. I am also thankful to the seminar participants at the UK Financial Conduct Authority for their feedback and ideas.
The code for this project partly draws on the *alogit* package for Stata by Jason Abaluck and Mauricio Caceres Bravo, and the *pylogit* Python package by Timothy Brathwaite, for which I am very grateful to their authors.
All views (and any errors) in this paper are my own and do not represent the position of the FCA.

2 Hereafter cited in its earlier public version: Iscenko (2018).

3 See also Bronnenberg, Dubé and Gentzkow (2012) and Bronnenberg, Dubé and Moorthy (2019) for comprehensive overviews of the literature.

with the supplier exists in one market, but influences the consumer's choices in others, as is the case with mortgages and personal current accounts.

There is an important unanswered question as to the mechanism through which brand familiarity affects behaviour. One option from the literature on limited attention is that existing providers are chosen more often because consumers consider them by default, and only pay attention to other alternatives if the default option is sufficiently bad to warrant it (Heiss et al., 2016; Abaluck and Adams, 2017; Ho, Hogan and Scott Morton, 2017; Hortaçsu, Madanizadeh and Puller, 2017). An alternative possibility is habit formation, where using a provider increases the consumer's taste for their products through habit (Dubé, Hitsch and Rossi, 2010) or reduced cognitive effort of operating a familiar environment (Murray and Häubl, 2007).

In this paper I explore these questions in the context of the UK mortgage market with a unique combination of transaction-level data, linked credit files and extensive datasets about firms' products, advertising and locations. I specifically investigate: (a) whether and how an existing current account relationship with a lender affects consumers' mortgage choices after controlling for other brand awareness factors, (b) the extent to which the attention and preference channels separately contribute to these effects, and (c) whether the effects are materially different depending on borrower demographics.

I develop a limited attention structural model based on the tradition of discrete choice models with unobserved consideration sets. In this class of models, the probability of each alternative being considered is determined by the characteristics of the alternative in question, and the borrowers then make choices out of the resulting consideration sets according to their preferences (Swait and Ben-Akiva, 1987; Goeree, 2008; Van Nierop et al., 2010; Abaluck and Adams, 2017; Crawford, Griffith and Iaria, 2019). I extend this alternative-specific consideration model by adding latent class model (Ben-Akiva et al., 1997; Greene and Hensher, 2003) elements, allowing for two types of borrowers that differ in their attention and preferences. Including factors associated with lender familiarity—existing banking relationships, proximity of branches, advertising—in both attention and preference parts of the model (alongside with the conventional product characteristics, where appropriate) gives me a unique opportunity to test what their effects are and which channel dominates.

This modelling approach allows me to distinguish two distinct types of borrowers. Type 1, which has demographic characteristics commonly associated with lower financial sophistication (lower income, worse credit history, lower education), is less price sensitive and somewhat prone to inattention. Borrowers in this type are also a lot less likely to consider lenders with whom they have no existing products (with consideration probability of 0.56) than familiar alternatives (with probability of 0.97). The average welfare gain that a Type 1 forgoes on average due to their limited attention is equivalent to reducing annual mortgage costs by 1.2% of their post-tax income. Although inattention is also present to some extent among Type 2 borrowers, who tend to be richer, more credit-worthy and more price-sensitive, their lapses in attention are much less linked to having an existing relationship with the lender, and around half as costly as for Type 1 (at 0.6% of their post-tax income). After controlling for existing relationships with lenders, other factors, such as advertising expenditure or branch presence near a borrower's home, have only a small effect on attention for both borrower types.

Both borrower types are similar, however, in exhibiting very strong preferences for lenders with whom they have an existing current account *among the options they consider*. The implied "own-lender" interest rate premium that borrowers are willing to trade off for going to their current account provider is equivalent to over 5% of post-tax annual income for both types.

I also simulate the effects of a hypothetical intervention that enforces full attention across all borrowers. I find that it has a fairly small effect on market shares and prices paid. Overall consumer surplus improves on average, by an equivalent of reducing interest rates by 17 basis points, but there are notable distributional differences.

The intervention improves welfare for most of the more inattentive Type 1 borrowers, whose average annual mortgage payments fall by £130 and average utility increases considerably. For Type 2 borrowers, however, the situation is less clear-cut. Most of them incur significantly higher borrowing costs after the intervention (an average increase of £200 per year) because of the positive demand shock to the best-priced mortgages on the market. The increased ability to find more suitable products on other preference dimensions (rate type, fees, location, etc) is just sufficient to compensate Type 2 for these price increases, leading to only a very small increase in average consumer surplus. Welfare declines in the new equilibrium for over a third of all borrowers (28% of Type 1 and 47% of Type 2). Overall, the benefits of making borrowers pay full attention (even if it were feasible) are surprisingly muted relative to the scale of the regulatory intervention it would require.

related literature My work contributes to several broad strands of literature.

First, I contribute to the household finance literature on mortgage decisions. Due to their importance as the largest household financial liability, as well as data availability, mortgages have attracted a lot of research interest recently, especially in the area of price dispersion (Allen, Clark and Houde, 2014a; Iscenko, 2018; Bhutta, Fuster and Hizmo, 2019), intermediation and advice (Woodward and Hall, 2012; Mysliwski and Rostom, 2018; Iscenko and Nieboer, 2018; Robles-Garcia, 2019; Foà et al., 2019), and, extensively, on promptness of remortgaging decisions (e.g. Campbell, 2006; Agarwal, Driscoll and Laibson, 2013; Andersen et al., 2019, , and others). Out of the recent literature in this field, my work is closest to Allen, Clark and Houde (2019), who find that brand loyalty has a significant effect on lenders' market power in the Canadian mortgage market. Due to the central role of price negotiation for Canadian mortgages, and its absence in the UK price-posting mortgage market, however, Allen, Clark and Houde (2019) and I model very different market structures. As a result, we provide complementary insights on the channels through which brand loyalty can affect consumer behaviour and equilibrium outcomes. Beyond mortgages but still within household finance, this paper fits in with the relatively new strand that applies methods from structural industrial organisation to study behaviour in retail financial markets (Handel, 2013; Heiss et al., 2016; Ho, Hogan and Scott Morton, 2017; Honka, Hortaçsu and Vitorino, 2017; Nelson, 2018; Benetton, 2019; Robles-Garcia, 2019). My findings about the differences in attention between the two borrower types, the types' demographic characteristics, and the distributional effects of interventions to help them also echo the recurring themes about the naive and sophisticated consumers in the behavioural industrial organisation literature (Gabaix and Laibson, 2006; Eliaz and Spiegler, 2006; Grubb, 2015; Armstrong, 2015, etc).

Second, this research is also part of the broader applied work on limited attention in consumer choice (e.g. Swait and Ben-Akiva, 1987; Goeree, 2008; Van Nierop et al., 2010; Crawford, Griffith and Iaria, 2019).[4] The model I develop is one of the first to exploit the recent identification results in Abaluck and Adams (2017), who show that attention and preference parameters in limited attention multinomial logit models are identified under significantly less strict restrictions than used in earlier papers. This allows me to avoid excluding advertising and other familiarity variables from the preference part of the model (unlike e.g. Goeree (2008) and Honka, Hortaçsu and Vitorino (2017)) and to provide additional evidence on the (small) effect advertising has on preference formation. More importantly, I am also able to explore and compare both potential channels through which lender familiarity affects behaviour in ways that are new to the literature.

4 A recent survey of this literature is available in Honka, Hortaçsu and Wildenbeest (2019).

I also find that accounting for the role of existing links with suppliers, especially in the preference channel, implies a much lower effectiveness of information remedies than has been suggested in earlier literature on limited attention, for instance the counterfactual simulations in Goeree (2008) or Hortaçsu, Madanizadeh and Puller (2017). The results of my counterfactural simulation help bring recommendations from structural models closer to the recent evidence from randomised controlled trials of interventions to encourage switching in the UK energy (e.g. Tyers, Sweeney and Moon, 2019) and financial (e.g. Adams et al., 2019) sectors. This regulatory testing has often found that even very transparent, simple and timely provision of information about alternatives and gains from switching had very modest success in inducing consumers to switch providers, resulting in switching rate (percentage point) improvements in single digits.

The rest of the paper is structured as follows. Section 3.2 provides an overview of the data.[5] Section 3.3 outlines the theoretical model, followed by a more detailed discussion of the estimation approach and practicalities in Section 3.4. Section 3.5 describes the estimation results and predictions of the model. Section 3.6 discusses a policy counterfactual where full attention is enforced. Section 3.7 covers robustness checks for the main findings. Section 3.8 concludes. Additional tables and figures are provided in Appendix B.

3.2 Data

3.2.1 Datasets

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

FCA PRODUCT SALES DATA (PSD) This transaction-level dataset of all UK residential mortgage lending, collected by the UK Financial Conduct Authority, has been increasingly used in household finance and industrial organisation papers on the mortgage market in recent years (Benetton, 2019; Benetton et al., 2019; Iscenko, 2018; Liu, 2019; Robles-Garcia, 2019). It contains extensive information about loan, collateral property, and borrower characteristics for each new mortgage in the UK.[6]

This paper covers mortgages issued for the purchase of a new property (no refinancing) during 18 months between January 2015 and July 2016. For comparabil-

5 For a more detailed discussion of the UK mortgage market structure more generally, see Iscenko (2018).

6 More detail about the PSD and variables it includes is availabe in the PSD001 Data Reference Guide.

ity of alternatives and individuals, borrowers with niche products (interest-only or government-subsidised schemes) are excluded from the sample.

The focus of this paper on the impacts of borrowers' existing links with lenders (or other forms of familiarity) on behaviour requires another significant restriction on the applicable transactions. Around 70% of mortgage loans taken out during the relevant period use mortgage brokers for their search. However, there is strong evidence that decisions of brokers can be affected by factors, such as commission, that are not fully aligned with their customers' preferences (Woodward and Hall, 2012; Egan, 2019; Robles-Garcia, 2019). Because of the difficulty in disentangling borrowers' attention and preferences from those of their brokers in intermediated transactions, this paper only uses the 30% of mortgage loans that borrowers take out by approaching the lender directly.

MONEYFACTS I merge the PSD with the daily commercial dataset of mortgage products on the market from Moneyfacts. This information enriches the dataset in two important ways: (a) it provides extensive additional information for each product (fees, any extra features, eligibility criteria, availability restrictions); and (b) it allows me to observe the potential choice sets for borrowers at any point in time without needing to infer it. In the UK mortgage market that is characterised by lenders posting menus of prices and eligibility criteria with no subsequent negotiation, Moneyfacts data about the price structure, features and criteria are a comprehensive and fixed characterisation of each mortgage product.

CREDIT BUREAU FILES Uniquely among the recent UK mortgage research (with the exception of Iscenko (2018) and Iscenko and Nieboer (2018)), I am able to incorporate borrowers' full credit bureau files for 6 years up to their mortgage application, obtained from one of the UK's top 3 credit reference agencies (called 'credit bureaus' in the US). These data cover over 90% of all mortgage transactions recorded in PSD. In addition to credit scores to gauge borrowers' riskiness, full credit files contain information about their personal current accounts (PCAs) and credit products at a given point in time, and about each borrower's location before moving to their mortgaged property. This information is essential for accurately identifying the existing relationships between borrowers and lenders, and the characteristics of their environment before the mortgage application which might shape attention to lenders (e.g. branch presence around the borrower's residence).

OTHER SOURCES OF DATA I draw on additional commercial datasets to obtain information about lenders' characteristics and their links with the borrowers.

Firstly, I use the quarterly bank branch location dataset from Experian GOAD. Combining this with borrowers' historical postcode data from the credit bureau makes it possible to explore each bank's branch presence around each individual borrower (e.g. the number of branches within 5 miles of where the borrower lives and the distance to the lender's closest branch).[7]

Second, I obtain data on monthly advertising expenditure by each lender from the Ebiquity Portfolio. This portal contains data on total advertising expenditure by financial institutions, split by topic, media format and coarse region (e.g. North East England), where relevant. I restrict the expenditure to campaigns related to each lender's general banking and mortgages (as opposed to, for instance, an advertising campaign focused solely on a lender's credit card business).

Finally, as described in more detail in Iscenko (2018), I also use public and quasi-public data sources such as the UK 2011 Census to obtain postcode-level proxies for missing demographic characteristics (e.g. educational attainment and socio-demographic characteristics) and HM Land Registry to help identify newly built properties for assessing borrowers' eligibility for specific products.

3.2.2 Sample description

After the sample restrictions described above and exclusion of missing data[8], the final sample comprises 86,288 borrowers and 3,071,550 person-product observations.[9] The products in the sample come from 12 lenders which represent over 80% of total lending in the relevant period and include all major UK banks.

I randomly split the available observations into the sample used for estimating the demand-side model in section 3.4.1.1 (75% of borrowers) and a hold-out sample used to assess the model fit in section 3.4.1.3 (25% of the borrowers). Due to the relatively small number of observations for some of the mortgage products, the supply-side estimation as set out in 3.4.2 uses all available data. All reported results in section 3.5 are based on the predictions obtained by applying the estimated model to the full sample.

Table 3.1 summarises some of the key characteristics of the borrowers in the sample, and the options they face. The sample is approximately equally split between the

7 All distances are calculated using the standard procedure of converting UK postcodes to latitude and longitude coordinates using the UK Office of National Statistics Postcode Directory and applying the Haversine distance formula.

8 Iscenko (2018) discusses in more detail data attrition in the PSD due to the merge with Moneyfacts and credit bureau data in more detail and provides evidence that this does not result in any material sample selection bias.

9 Section 3.4.1.1 describes the methodology for constructing these counterfactual choice sets.

Table 3.1: Sample descriptive statistics

	Mean	σ	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
PANEL A: Borrower characteristics					
Age (years)	37.13	9.76	29	36	44
=1 if first-time buyer	0.44	0.50	0	0	1
Income (£1000)[a]	41.64	24.44	25.48	35.71	50.30
=1 if joint loan	0.60	0.49	0	1	1
Loan value (£1000)	167.40	112.06	90	137	210
Property value (£1000)	269.62	189.16	143	215	334
Credit score[b]	62.89	8.35	59.48	64.58	68.50
Loan-to-value ratio (LTV, %)	66.53	21.91	51.02	72.86	85.00
Lenders in choice set (N)	6.58	3.32	3	6	10
of which: with PCAs relationship (N)	1.08	0.93	0	1	2
with any existing product (N)	1.35	1.15	1	1	2
=1 if chose a lender with PCA	0.54	0.50	0	1	1
Observations (individuals)			86,288		
PANEL B: Lender features (relative to each borrower)					
Branches within 5 mi (N)	2.71	2.45	1.00	2.00	4.00
Regional advertising spend (£per cap pcm)[c]	2.39	1.59	1.03	2.15	3.45
Observations (lender-borrower pairs)			567,461		
PANEL C: Product features					
Initial interest rate (%)	2.64	0.79	1.99	2.53	3.14
Upfront fee (£)	562	548	0	295	999
Early repayment penalty (% of loan)	2.34	1.20	1.54	2.50	3.11
Fixed period length (years)	3.22	2.40	2.00	2.00	5.00
Observations (options)			3,071,550		

Note: (a) Income is the sum of post-tax household earnings for all individuals named on the mortgage 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) Average advertising expenditure by a lender in the broad region of the borrower's residence over 6 months up to mortgage application.

first-time buyers (FTBs) (those taking out their first mortgage) and home movers borrowing to buy a new property. Despite the high proportion of FTBs, however, only just over 25% of the sample are in their 20s, reflecting the recent UK trend of households making their first property purchase later in life.

Credit scores tend to be high relative to the general UK population, which is not surprising within a sample of households which qualified for a mortgage. Both loan and collateral property values vary considerably in the sample, leading the loan-to-value (LTV) ratio to range from below 10% to 95% with an average of 66%. With lenders' mortgage product menus being very closely linked to borrower LTVs, this means that the different borrowers in the data will be facing very different choice sets. It is also clear from the product descriptive statistics in Panel C that there is substantial variation in costs that borrowers can incur across products.

Several statistics in Panel A, when taken together, caution against overlooking familiarity and assuming all lenders are on 'equal footing' with regards to borrower attention. On average, each borrower has around six lenders to choose from and has an existing current account (or even any product) with just one of them. Yet, over a half of borrowers choose lenders with whom they already have a product.[10]

The distribution of the number of lenders with whom each borrower has PCAs is also important for making modelling decisions. As summarised in Abaluck and Adams (2017), in addition to the alternative-specific consideration (ASC) approach used in this paper, there is an option of modelling inattention using default-specific consideration (DSC). DSC has been used in the past to study cases in energy markets (Hortaçsu, Madanizadeh and Puller, 2017) or healthcare (Ho, Hogan and Scott Morton, 2017) where the existing provider has consumer's attention by default and other alternatives get considered only if the default option is sufficiently unsatisfactory. This might seem to be a natural framework for analysis of existing links in banking as well. Unfortunately, instead of there being one natural 'default' for everyone, over 25% of mortgage borrowers have a current account with more than one lender, and approximately the same number have no links with any of the represented lenders. In this context, it is difficult to apply the DSC approach meaningfully even with some modifications.

10 As clarified in more detail later, to count as having an existing relationship with a lender at the time of application, the borrower has to have opened their account at least 6 months before applying for a mortgage. This significantly exceeds the length of a typical UK housing transaction and thus minimises the risk that borrowers open a personal current account purely because they intend to apply for a mortgage with the same lender.

Panel B highlights additional variables on lender characteristics as they relate to each borrower, which both suggest that borrower attention may be important for mortgage choice and are more consistent with an ASC-type approach where multiple options might be vying for a borrower's attention. Due to the concentration of property purchases in urban areas, it very common for borrowers to have a branch of a potential lender within 5 miles of their residence. The availability of branches means (a) that access to lenders is unlikely to be a major issue but also (b) that most borrowers are likely to be exposed to different brands regularly. Furthermore, there appears to be a lot of variation in lenders' advertising expenditure, which given the existing findings on links between advertising and consideration (e.g Goeree, 2008; Terui, Ban and Allenby, 2011; Honka, Hortaçsu and Vitorino, 2017) could mean that borrowers are prompted to consider lenders to a different extent.

3.3 Model

In this section I set out the basic model for borrower and lender decisions in the non-intermediated part of the UK mortgage market. As explained earlier, I focus on the subset of the market (approximately 30%) where borrowers choose their mortgage directly, without using the help of a broker to search for or apply for products. I assume that the choice of the direct as opposed to intermediated channel is determined exogenously.

3.3.1 Demand

There are I borrowers (households), indexed by i, who are choosing a mortgage. In line with recent work on mortgage choice (Iscenko, 2018; Allen, Clark and Houde, 2019; Robles-Garcia, 2019) but in a simplification from Benetton (2019), I assume that the loan amount and property value (and hence the LTV ratio) are pre-determined by borrower's demographics, circumstances and financial position (e.g. savings for a deposit). The focus of the mortgage choice in the model is, therefore, a discrete choice a mortgage product from a choice set of products for which a borrower qualifies given their exogenous demographics, loan amount and LTV.

3.3.1.1 *Attention*

Let C_i be a set of mortgage choices available to borrower i. In modelling inattention, I largely follow the limited attention multinomial approach set out in Goeree (2008) and the alternative-specific consideration approach in Abaluck and Adams (2017), where attention to an available option $j \in C_i$ is a random event that occurs with some probability ϕ_{ij} that is a function of j's characteristics. In a departure from the standard setting, however, I do not let the probabilities of considering each *option* be entirely independent.

In the context of mortgages, the nature of product listings in lenders' websites and marketing literature, mean that while looking up details of a specific product a borrower would be made aware of all of that lender's products for which they qualify.[11]Instead, for the sake of realism and computational tractability, I restrict the consideration decisions to be at the lender level. If a borrower i considers lender l, they consider every product j in the set of products l offers ($j \in J_l$) that is also in i's choice set ($j \in C_i$).

This approach means that the 'standard' alternative-specific consideration part of the model effectively occurs at lender rather than product level. Hence, the probability of lender l being considered by borrower i is given by:

$$\phi_{il}(\gamma) = \frac{e^{f(z_{il},\gamma)}}{1 + e^{f(z_{il},\gamma)}} \tag{3.1}$$

where γ are attention parameters and z_{il} is a vector of characteristics of lender l, including relational characteristics between l and i (e.g. the number of branches near i's home). I discuss the specific potential drivers of attention included in z_{il} in section 3.4.1.1.

Let $\bar{C_{ir}} \subseteq \bar{C_i}$ be a possible consideration set for i in a scenario r. As a shorthand, let $l \in \bar{C}_{ir}$ stand for the event when lender l is considered, leading to all of l's products being in i's consideration set $J_l \in \bar{C}_{ir}$. Then, the probability of any specific $\bar{C}_{ir}$ is:

$$\Phi_{ir}(\gamma) = \prod_{l \in \bar{C}_{ir}} \phi_{il}(\gamma) \prod_{s \notin \bar{C}_{ir}} (1 - \phi_{is}(\gamma)) \tag{3.2}$$

The total number of consideration sets is the $2^{n_{il}}$, where n_{il} is the total number of unique lenders in C_i.

3.3.1.2 *Choice*

Conditional on considering the subset of available alternatives $\bar{C}_{ir}$, a borrower's choice can be represented as the standard random utility model where borrower i's utility from a product j offered by lender l is given by:

$$V_{ijl} = x_{ijl}\beta + \chi_l + \xi_{jl} + \zeta_{ij} \tag{3.3}$$

11 For instance, typically, the website would contain a short sequence of eligibility questions (first time buyer status, desired loan amount, desired property value) which lead to a full list of products available to borrowers who satisfy these criteria.

where x_{ijl} are the characteristic of mortgage j as experienced by i (e.g. expected monthly interest payment), β is a vector of taste parameters, χ_l is a lender-level fixed effect (such as service quality and other unobservables that do not vary between products from the same lender), ξ_{jl} is an unobservable market-wide shifter of demand for j (discussed in more detail in section 3.4.1.2 on identification), and ζ_{ij} a random shock to i's taste for j.

The probability that i chooses j out of the consideration set $\bar{C}_{ir}$ is then:

$$s_{ijlr} = Pr([V_{ijl} > V_{iks} \quad \forall k \in \bar{C}_{ir}]) \tag{3.4}$$

3.3.1.3 *Borrower heterogeneity*

I allow the parameters that determine attention and choice to be affected by demographic characteristics of the borrower, using the latent class approach (Ben-Akiva and Bierlaire, 1999; Greene and Hensher, 2003). To keep the model tractable and interpretable, I allow for two borrower types, each with its own set of parameters, (β_1, γ_1) and (β_2, γ_2), respectively. Any given borrower's type is a stochastic function of their demographic characteristics. Borrower i belongs to type t with probability σ_{it}. Naturally, with only two types $\sigma_{i2} = (1 - \sigma_{i1})$.

After accounting for demographic variation, the unconditional probability of observing borrower i choosing j from lender l is:

$$P_{ijl} = \sum_{t=1,2} \sigma_{it} \sum_{\bar{C}_{ir} \subseteq \bar{C}_i} \Phi_{ir}(\gamma_t) \; s_{ijlr}(\beta_t) \tag{3.5}$$

3.3.2 Supply

Although the primary focus of this paper is on borrower behaviour, I develop a simple model of supply-side behaviour to analyse a policy counterfactual in section 3.6.

3.3.2.1 *Lender mortgage pricing*

There are N_L lenders who compete to sell mortgages to households by setting prices in a one-shot, noncooperative Nash equilibrium setting. Each lender l has a set of mortgage products J_l that cover a range of product characteristics and eligibility criteria. Lenders maximise expected profits by setting interest rates for each of these products. As before, there are no discrete markets in the standard sense: because different lenders specify eligibility criteria differently, mortgage choice sets vary between borrowers, as explained in section 3.4.1.1.

Lenders are assumed to have correct expectations about the pool of potential borrowers and sets of products for which those borrowers qualify. There is no default risk.

Given the demand probabilities defined in (3.5), lender l's expected profit of offering product j is:

$$\Pi_{jl} = \sum_{i \in I_j} P_{ijl} t_j (r_j - mc_j) \tag{3.6}$$

where I_j is the set of all borrowers i such that $j \in C_i$, t_j is the length of the initial deal period for product j, r_j is the interest rate for the product, and mc_j is the marginal cost of selling j. Additional elements of product pricing (fees or exit charges) are implicitly assumed to be exogenous, and enter lenders' profits through mc_j. Similar to earlier work (e.g. Robles-Garcia, 2019), I assume households switch to a new mortgage product at the end of the teaser period, and that the loan amount is exogenous (normalised at one).[12] For each borrower that qualifies for product j, the change in the interest rate affects the probability of choosing the product but not the value of the loan.

The lender chooses the interest rates for all products in J_l to solve the following total profit maximisation problem:

$$\max_{\{r_j\} j \in J_l} \Pi_l = \sum_{j \in J_l} \sum_{i \in I_j} P_{ijl} [t_j (r_j - mc_j)] \tag{3.7}$$

The first-order conditions of (3.7) can be rearranged to give a sequence of profit-maximising interest rates:

$$r_j^* = mc_j - \sum_{i \in I_j} \left(\frac{\partial P_{ijl}}{\partial r_j} \right)^{-1} \left[P_{ijl} + \sum_{k \neq j \in J_l} \frac{\partial P_{ikl}}{\partial r_j} \frac{t_k}{t_j} (r_k - mc_k) \right] \tag{3.8}$$

where the first term captures the marginal cost, the second is the mark-up and the third term reflects the effects of the r_j on the profits from l's other products.

12 The switching assumption is based on a high rate of prompt remortgaging in the UK market. A recent Financial Conduct Authority (2018a) report on the mortgage market shows that 77% of borrowers switch to a new mortgage product within 6 months of their teaser rate expiring.

3.4 Estimation

Having set out the theoretical framework in general terms, in this section I describe the practical detail of estimating it with the available data. I first go through the variable specification for each part of the demand model, and explain why the parameters are identified in this setting. I conclude the demand sub-section by discussing the performance of the estimated baseline model. I then cover these topics for the supply model as well.

3.4.1 Demand

3.4.1.1 *Specification*

choice sets Financial products, including mortgages, typically have strict and multidimensional eligibility criteria which make it invalid to assume that each borrower can access the whole market or even a standardised segment of the market. Instead, I construct borrower-specific counterfactual choice sets C_i using the observed choice data and the extensive information on product listing times and eligibility criteria for each product available from Moneyfacts.

Using the latter, I can identify products that (a) were on the market at the time the borrower made their choice, (b) were available for at least a month to allow for date measurement error) and (c) for which the borrower satisfied all the eligibility criteria. The explicit criteria I consider are: geographic, borrower type (e.g. first-time buyer only), existing customer exclusivity, as well as minimum and maximum limits on age, income, loan-to-value (LTV) ratio, loan-to-income (LTI) ratio, property value and loan amounts.

Then, for further robustness, I use the observed transaction data to calculate additional 'implicit' eligibility criteria: the minimum credit score, maximum LTV, and maximum LTI across the borrowers accepted for each product. I then exclude from the counterfactual choice sets any product for which the borrower does not meet these implicit criteria.[13]

consideration sets I further restrict the probability of i paying attention to product j which was set out in (3.1) by assuming that $f(z_{il})$ is linear in z:

$$\phi_{il}(\gamma_t) = \frac{e^{z_{il}\gamma_t}}{1 + e^{z_{il}\gamma_t}} \tag{3.9}$$

where γ is a vector of parameters to be estimated.

13 This approach is identical in terms of data and methodology to the first stage of identifying dominated products – defining the 'available choice sets' – in Iscenko (2018).

The potential drivers of attention in z_{il} aim to capture factors that might affect the salience of a particular lender in the borrower's without the borrower deliberately looking for information. One such factor is advertising intensity, which I include in the model as the average advertising expenditure by lender l in i's broad geographical region (e.g. North-East England) per capita in the six months before the application. Another factor is the visibility of the lender's brand in the surrounding area, captured by the number of l's bank branches within 5 mile radius of the borrower's postcode. Furthermore, larger lenders are likely to be more visible in other ways, such as featuring in news coverage more or being recently chosen by the borrower's friends and family. I control for this by including in z_{il} lender's size as measured by their mortgage lending volume in 2014 (i.e. before the estimation sample starts).

For investigating a more direct form of familiarity, I include an indicator for the borrower having an existing relationship (a personal current account (PCA)) with the lender[14] and the number of years the borrower had continuously had an account with this lender prior to application.[15]

UTILITY I further assume that the individual taste shock ζ_{ij} in the utility function (3.3) follows the type 1 extreme value distribution. This means that the probability of i choosing j out of a consideration set $\bar{C}_{ir}$ in equation (3.4) can be expressed as:

$$s_{ijlr}(\beta_t) = \frac{e^{x_{ijl}\beta_t + \chi_l}}{\sum_{k \in \bar{C}_{ir}} e^{x_{iks}\beta_t + \chi_s}} \tag{3.10}$$

The initial interest rate (r_j), appears in x_{ijl} both on its own and interacted with the teaser period length. The interaction is necessary to account for the fact that a borrower can be reasonably expected to give a greater weight to an interest rate that will apply for a longer period. For the same reason, I include the interaction of the reversion (post-teaser) interest rate with the teaser period length as well as the reversion rate on its own.

Other product and lender characteristics in x_{ijl} are: upfront fees, cashback, early repayment penalty, fixed effects for the length of the fixed-rate period, offer of free property valuation (a commonly offered incentive), whether the mortgage offers

14 To avoid the reverse causality affecting results (borrowers opening an account with a lender with which they want to get a mortgage), I require the PCA to be at least 6 months old at the time of mortgage application. The results in the paper are not materially affected by requiring that the account is at least a year old instead.

15 The results in this paper are also robust to a more complex ways of measuring existing relationships, as shown in section 3.7.

payment holiday and underpayment options. I also calculate several features that are unique to each borrower-product/lender pair: the distance from the borrower's address at application to the lender's nearest branch, and how far the borrower's LTV and credit score are from the minimum eligibility criteria for the product. These last two variables control for potential preference for products where the borrower has more 'headroom' over minimum standards and thus may perceive a lower risk of rejection.

I also allow all of the drivers of attention for the lender l, z_{il} other than lender size[16], to also influence product preferences by appearing in x_{ijl} as it appears plausible that e.g. a borrower might have a preference for having multiple branches nearby in addition to finding the lender more salient due to their presence.

BORROWER TYPES To complete the latent class model of borrower heterogeneity, I assume that the probability of a borrower i belonging to type 1, σ_{i1} has the following form:

$$\sigma_{i1}(\delta) = \frac{e^{d_i \delta}}{1 + e^{d_i \delta}} \tag{3.11}$$

where d_i is a vector of i's demographic characteristics and δ is a vector of parameters to be estimated. Borrower's characteristics that are allowed to affect the borrower type are: credit score, age, income, LTV ratio for the loan, as well as indicators for whether they are a first-time buyer, are applying for a joint mortgage or are self-employed. I do not observe borrower's education directly, but I include the percentage of population in low-skilled occupations in their postcode as proxy for educational attainment.

LIKELIHOOD Combining equations (3.9) and (3.10) in this section with the probability of observing i choosing a product j as specified in (3.5) gives the following observed demand probability:

$$P_{ijl}(\beta, \gamma, \delta) = \sum_{t=1,2} \sigma_{it} \sum_{\bar{C}_{ir} \subseteq \bar{C}_i} \prod_{l \in \bar{C}_{ir}} \frac{e^{z_{il}\gamma_t}}{1 + e^{z_{il}\gamma_t}} \prod_{s \notin \bar{C}_{ir}} (1 - \frac{e^{z_{is}\gamma_t}}{1 + e^{z_{is}\gamma_t}}) \frac{e^{x_{ijl}\beta_t + \chi_l}}{\sum_{k \in \bar{C}_{ir}} e^{x_{iks}\beta_t + \chi_s}} \tag{3.12}$$

where σ_{it} is as defined in (3.11).

16 Unlike other attention variables, size only varies across lenders and has no within variation, meaning that it is fully absorbed by lender fixed effects.

The corresponding empirical log-likelihood function for any observed set of choices for a set of individuals I is:

$$LL(\beta, \gamma, \delta) = \sum_{i \in I} \log \left(\sum_{j \in C_i} y_{ij} P_{ijl}(\beta, \gamma, \delta) \right) \tag{3.13}$$

where y_{ij} is an indicator variable that takes the value of 1 if i chose j and 0 otherwise.

Given the likelihood function in (3.13), I estimate the taste, attention and demographic parameters with exact maximum likelihood.[17]

3.4.1.2 *Identification*

This section provides an informal discussion of how the estimation procedure described above allows me to separately identify the attention, utility and borrower latent class parameters.

latent class parameters To the extent that the utility and attention parameters are identified (as will be shown later), the latent class element of the model just mixes between two fully identified alternative-specific consideration models for each type. All demographic variables that affect the type probability σ_{it} are excluded from the attention and utility specifications. One identification concern could be that some of the demographics (e.g. income, age), which affect borrower type, can also influence which products borrowers qualify for and thus contribute to the variation in choice sets across borrowers that helps to identify parameters in other parts of the model. This should not be a problem in practice for the following reasons: (a) I do not solely rely on choice set variation to identify other parameters, (b) the latent class specification also includes borrower education (% of low-skilled workers in borrower's postcode) which is not part of mortgage product eligibility criteria, and (c) some of the important drivers of choice set variation (e.g. timing of the mortgage choice) are excluded from the latent class specification.

attention parameters Abaluck and Adams (2017) show that attention and preference parameters are separately identified in the class of limited attention models that includes the approach I describe in section 3.4.1.1. The identification comes from the asymmetries in cross-derivatives of choice probabilities (i.e. restrictions from economic theory) even when exactly the same variables appear in the probability of attention, ϕ_{il}, and indirect utility, V_{ijl}. An important identifying exclusion

17 I verify that the solution is a global maximum using the basin-hopping global search algorithm (Wales and Doye, 1997; Jones, Oliphant, Peterson et al., 2001–) with gradient-based BFGS optimisation in the inner local loops.

restriction in their setting is that the probability of paying attention to the alternative j depends solely on characteristics of j and not any other alternatives. As can be seen in (3.9) above, the model I estimate satisfies this restriction and exclude characteristics of rival options from the probability of considering each lender.

To achieve a more realistic description of borrower choice, I also impose additional exclusion restrictions beyond the minimum required in Abaluck and Adams (2017). I assume that the detailed product characteristics — interest rate, fees, eligibility status, etc — affect borrowers' preferences but not their probability of attention. This granular product information is not fully covered by the headline advertising or branch shop fronts, and requires effort (i.e. attention) to find out, which makes it implausible that it would influence borrower's attention.

Utility parameters There are two recurring identification concerns in standard demand estimation: (a) market shares and (b) potentially endogenous product characteristics (e.g. price). I address these in turn.

With respect to (a), Abaluck and Adams (2017) prove that full consideration market shares are identified in limited attention discrete choice models under the assumptions described above. Without additional concerns about endogeneity of product characteristics, this is sufficient to identify the parameters in the utility function.

In my setting, identification is further strengthened by two sources of exogenous variation choice probabilities within the same 'market'. First, as described in Berry and Haile (2016), there is standard micro-data variation due to changes in characteristics that are specific to each borrower-lender combination (e.g. distance to the nearest branch). Second, my data on eligibility criteria and exact product availability dates allow me to create borrower-specific choice sets. As borrower characteristics (e.g. the desired loan amount) change, new products get incrementally added to and removed from their choice sets, leading to further variation in the individual choice probabilities.[18]

The concern regarding (b) endogenous characteristics, is that the unobserved product-level demand shifter ξ_{jl} in the utility function (equation (3.3)) may be correlated with the price (interest rate). I use a different approach to this problem from the standard instrument-based solutions originating from Berry, Levinsohn and Pakes (1995).[19]

18 My sample of c86,000 borrowers contains 63,859 unique choice sets, which all involve different combinations of 2,592 products.

19 Berry and Haile (2016) provide a recent summary of the approaches in the BLP tradition, by summarising the identification concerns in the original (and more general) settings, showing that

The reasons for adopting a different approach are twofold. First, in contract with most settings in which the BLP price instruments are used, mortgage prices are multi-dimensional — characterised by a combination of multiple interest rates, fees and penalties. Choosing to use instruments for some of these dimensions (e.g. the main interest rate) but not others can therefore be somewhat arbitrary. Second, mortgages, like many other financial products, are fundamentally different from cars or other consumer goods in that they are essentially just a detailed description of cash transfers between different time periods and states of the world. Conditional on a comprehensive specification of the financial attributes and, importantly, the lender, a mortgage product does not have meaningful intrinsic utility or quality. As a result, the interpretation and importance of the *product-specific* unobservable fixed effect ξ_{jl} for mortgages are different from many traditional IO settings.

In light of these considerations, I deal with the potential endogeneity of product characteristics with an approach that is standard in the literature on another financial market — US health insurance (see, e.g., Abaluck and Gruber, 2011; Ho, Hogan and Scott Morton, 2017). I rely on using my comprehensive dataset of the objective mortgage product characteristics (costs, time profile of interest rate changes, additional incentives, eligibility) to explicitly control in x_{ijl} for all observable information about products that a borrower might reasonably have access to when they make a decision. There can be differences between lenders in their service quality and speed, perceived or real risk of rejection, or even just the positive sentiment about the brand, but those are captured by lender fixed effects χ_l.[20] Given the extensive available data and the 'formal' nature of individual mortgage products, I then assume that $\mathbb{E}\left[\xi_{jl} \mid x_{ijl}, \chi_l\right] = 0$.

3.4.1.3 *Performance*

IN-SAMPLE FIT The full details of the estimated baseline demand-side model are reported in Table B.1 in the Appendix. The model has an in-sample McFadden R^2

---

the use of choice-level data does not in itself identify endogenous prices, and discussing the potential sources of price instruments.

20 In motivating his use of price instruments, Benetton (2019) gives an example of the identification challenge posed if a lender lowers screening standards while raising its interest rate. This is a legitimate concern that I address in two ways: one by enforcing the product's screening criteria at choice set construction stage (so as to not overstate potential demand for cheaper products) and also by including controls for borrower's 'headroom' relative to the product's minimum standards in the utility model. There could conceivably be more idiosyncratic lending criteria (e.g. different willingness to lend on flats in public housing blocks) that could vary with price. Those standards, however, are typically part of lender's overall policy and do not vary between products. As such, they are taken care of by lender fixed effects in the utility specification. (See the UK Finance Mortgage Lenders' Handbook for conveyancers, compiled by the UK mortgage lenders' trade association from individual firms, for additional evidence that the legal processes and minimum standards are set within a lender and do not vary across products.)

of 0.326. I test the baseline model against the null hypothesis that the attention and preference parameters are the same across both borrower types using the maximised likelihood from a model with one borrower type (model (1) in Table B.3).[21] The likelihood ratio test strongly rejects this null hypothesis with the LR test statistic of 12,358 and p-value of 0.000.

I further test and reject the null hypothesis that there is no limited attention by comparing the homogeneous limited attention model above to the simple conditional logit as defined in (3.10) (LR test statistic of 10,137 and p-value of 0.000).

out-of-sample performance I use the estimated parameters to make predictions for the hold-out sample of 21,572 borrowers that were not used in the estimation. The model fits the key moments of the out of sample data well. As shown in Figure 3.1, the predicted joint market shares for the largest four, middle four and the smallest four lenders are very close to those observed in the data. On the level of individual firms, the predicted market shares are within 2 percentage points of the observed ones for all but two lenders in the hold-out sample.[22] The model predicts whether or not each specific borrower chooses a specific lender with 88.4% accuracy.

The fit is also good with respect to other dimensions of the choices. Table 3.2 compares the average characteristics of predicted choices[23] to the average product characteristics of chosen and not chosen observations. For all product characteristics, the averages across predicted choices match the actual chosen products well and are distinct from options that were not chosen.

3.4.2 Supply

As explained in section 3.3.2.1, I focus on the lender's setting of initial interest rate. To the extent that any other product characteristics (including fees and the reversion rate) affect pricing, they do so by changing the marginal cost of the product.

I recover mc_j from equation (3.8) by using the estimated demand parameters and information about product, lender and borrower characteristics.

21 Although the two models may not appear to be nested, one can obtain the standard limited attention multinomial logit model from my latent class version in equation (3.12) by restricting $\beta_1 = \beta_2$ and $\gamma_1 = \gamma_2$ (39 degrees of freedom). Then the demographic factors no longer affect the likelihood.

22 I am unable to disclose lenders' individual market shares for comparisons due to confidentiality restrictions on the use of PSD.

23 These are calculated as probability-weighted averages of the variable of interest across all products in each borrower's choice set, with the predicted P(Choice)_{ij} as a weight. The results are not materially different if the product with the highest P(Choice)_{ij} for each borrower is treated as their predicted choice.

Figure 3.1: Out-of-sample performance: Lender market share prediction accuracy

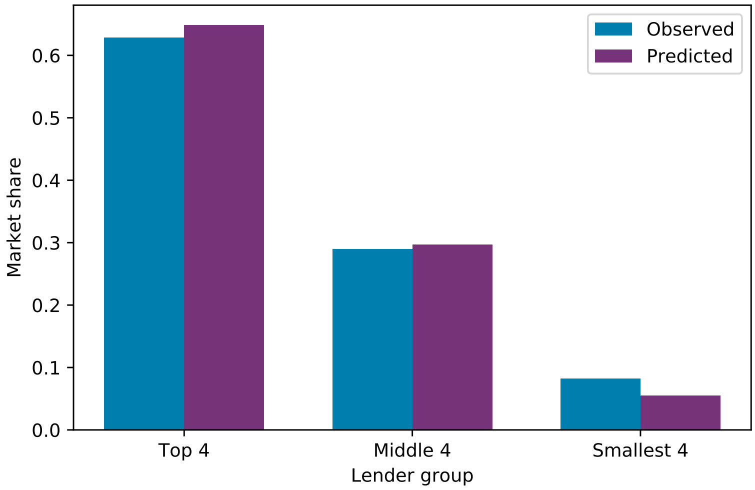


Table 3.2: Out-of-sample performance: mean product characteristics

	Predicted choices	True: chosen	True: not chosen
Initial interest rate, %	2.74	2.73	2.64
Reversion interest rate, %	4.03	4.02	3.96
Upfront fee, £	530.87	539.24	563.11
=1 if 2-year fixed rate	0.43	0.43	0.36
Teaser period, years	3.29	3.34	3.45
=1 if existing relationship with bank	0.51	0.54	0.17
N current accounts held	0.80	0.88	0.25
Distance to closest branch, mi	1.87	1.86	2.31
N of branches within 5mi radius	3.23	3.27	2.71
Observations (options)		21,572	748,637

identification In section 3.4.1.2 above, I argue that the demand side of the model is identified and so generates valid price-elasticities as inputs into the pricing equation. Subject to demand-side identification, in my simple Nash-in-prices setting with constant marginal costs, marginal costs obtained from the first-order conditions are identified without the need for further instruments (Berry and Haile, 2016). In essence, my supply side is identified through (admittedly, very substantial) theoretical restrictions.

Given the purpose that marginal costs serve in the simulations, I do not need to identify how they vary with specific contributing factors such as product characteristics, lenders' funding costs or capital regulations.

3.5 Results

In this section I report the results and predictions of the estimated model. First, I outline the results of the demographic heterogeneity part of the model, describing the two borrower types, their distribution and characteristics. In the following subsection, I discuss the extent of the limited attention predicted by the model and how it varies between the two borrower types. I conclude the overview of the results by elaborating on the role of bank familiarity in borrowers' attention and preferences, including the implied premiums that the two types are prepared to pay for choosing a familiar alternative.

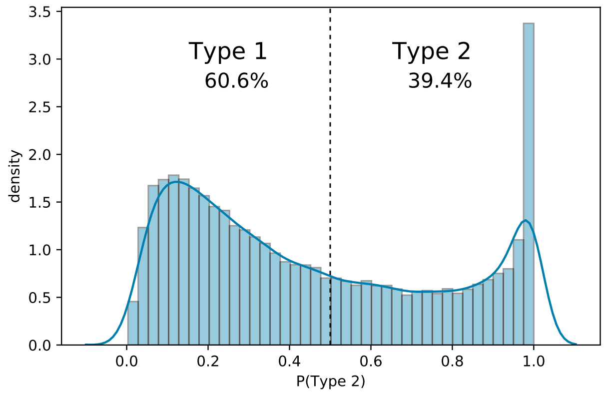
In discussing the results, I use the following shorthand terms to refer to the key outputs of the demand-side model:

- **P(Attention)**$_{il}$: probability that borrower i considers the lender l as defined in (3.9).
- **P(Preference)**$_{ij}$: probability that i prefers the alternative j to all others in their *full choice set*. This is equivalent to the overall probability of choice in the standard multinomial logit setting (equation (3.10)).
- **P(Choice)**$_{ij}$: the overall probability of observing i choosing alternative j in practice. It is the P_{ijl} as defined in (3.12) above.

3.5.1 Demographic variation

The model identifies two distinct groups of borrowers. As shown in Figure 3.2, the distribution of the probability that a borrower i belongs to type 2 is bimodal, with a clear mass at 1 and another concentration of borrowers around 0.1. If one uses P(Type 2) = 0.5 as a cut-off between the types, the majority (60.6%) of borrowers

Figure 3.2: Distribution of predicted probabilities of Type 2, (σ_{i2})



fall into type 1 and 39.4% are allocated to type 2. There is, however, a sizeable group of borrowers (around 15%) who have roughly equal chances of being either type as their P(Type 2) is between 0.4 and 0.6.

The parameter estimates from the borrower type probability model (reported as average marginal effects in P(Type 2) in Table 3.4a) paint a coherent picture of the two groups. The probability of belonging to type 2 increases strongly with household income and credit score. This probability gets lower, however, for borrowers who are older, more leveraged (as measured by the LTV ratio), self-employed or for those who live in areas with more unskilled or low-skilled workers. The negative effect of a joint application on the probability of belonging to type 2 is likely to be due to the fact that the net income in that case would be for the household as a whole, and the incomes of the two individual applicants are likely to be lower. Figure B.1 contains bivariate plots to illustrate how the predicted probability of being in type 2 changes across the distributions of borrower's income, loan amount and (postcode-level) education.

The factors that make the borrower more likely to be in Type 1 are remarkably similar to the demographic characteristics that are shown to be associated with the increased likelihood of dominated mortgage choices and larger avoidable costs in Iscenko (2018). In fact, the two groups - the poorer, less educated, more leveraged and the richer, more educated and better at managing credit - recur regularly in the literature as the unsophisticated and sophisticated consumers, respectively (e.g. Lusardi, Mitchell and Curto, 2014).

There are notable differences in the estimated preference parameters between the two borrower types. As can be seen from the selected average marginal effects on

Table 3.3: Estimated marginal effects on borrower type and on product demand across types, in pp

(a) Demographic effects on P(Type2)

	Δ P(Type 2)
Credit score	0.308
Age (years)	-0.514
Net income (£1000)	1.719
Loan-to-value (LTV, %)	-0.136
=1 if first-time buyer	10.568
=1 if joint mortgage	-16.157
=1 if self-employed	-9.672
Postcode: % low-skilled	-0.284

(b) Differences in effects on demand between types

	Δ P(Choice)	
	Type 1	Type 2
Initial interest rate (%)	-2.57	-3.00
Reversion interest rate (%)	-0.91	-1.26
=1 if 2-year fixed rate	2.06	3.12
=1 if 5-year fixed rate	-1.55	3.64
Distance to closest branch (mi)	-0.19	-0.18
Headroom to max LTV (pp)	2.10	1.17

the P(Choice)_{ij} in Table 3.4b [24], Type 2 borrowers tend to be more price-sensitive with respect to both interest rates, show a stronger preference for fixed interest rate mortgages and are less concerned about how close they are to the maximum LTV standards of the mortgage. In contrast, Type 1 borrowers appear to put a much larger weight on not being close to the maximum LTV for the product, potentially due to greater concerns about the application being rejected. Although Type 1 prefers a 2-year fixed rate to an adjustable rate mortgage, they in fact put a negative value on longer-term fixed rates relative to the rate varying throughout the contract. Both types appear to have broadly equal (and relatively small) preference parameters for the distance to the lender's nearest branch, suggesting it is not a major consideration for borrowers, at least as far as preferences are concerned.

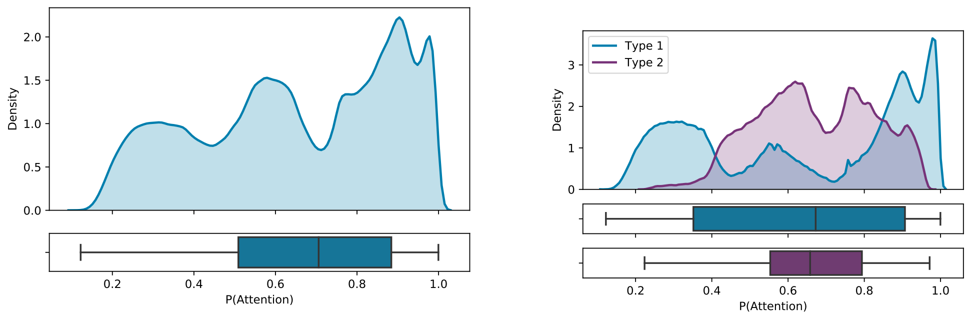
In addition to preferences in product characteristics, the two estimated types also differ in their degree of inattention and the extent to which lender familiarity plays a part in their attention and choices. The next section explores those differences in more detail.

24 For interest rate variables, the table reports the total marginal effect (both from standalone interest rate terms and their interaction with the teaser length). For instance, for type 1 the marginal effect of the initial rate on borrower i's choice of product j is:

$$\frac{\partial P_{ijl1}}{\partial r_j} = (\beta_1^{\text{ir}} + \beta_1^{\text{ir x tlength}} t_j) \sum_{\bar{C}_{ir} \subseteq \bar{C}_i} \Phi_{ir}(\gamma_1)\ s_{ijlr}(\beta_1)(1 - s_{ijlr}(\beta_1))$$

where t_j is the length of the teaser period for product j. The reported parameters are simple averages of these individual and product-specific partial derivatives. All parameters used to calculate the reported marginal effects are significant at 1% confidence level.

Figure 3.3: Distribution of the predicted $\mathrm{P(Attention)}_{il}$



(a) Overall (b) Split by borrower type

3.5.2 Limited attention

The estimated model suggests that there is non-negligible inattention. The average predicted $\mathrm{P(Attention)}_{il}$ across all borrower-lender pairs is 0.65. As can be seen from the whole distribution in Figure 3.4a, there is a lot of variation in the predicted probabilities, with some lenders having almost no chance of being considered by some borrowers.

Figure 3.4b compares the distributions for the two borrower types. Inattention appears to manifest differently in Type 1 and Type 2. The latter display a moderate degree of inattention fairly consistently. Over three quarters of predicted values of $\mathrm{P(Attention)}_{il}$ for this type lie between 0.5 and 0.9, meaning that they are more likely than not to consider most of the available lenders but often also have a small chance of not paying attention. In contrast, borrowers allocated to Type 1 have a highly bimodal distribution: they either consider a lender nearly with certainty or they are very unlikely to consider to them. As I explore in more detail in the section on familiarity, existing relationships between the lender and the borrower play a critical role in determining which side of the distribution the lender occupies. Type 1 borrowers are also more likely to consider only one lender, with probability of 0.143 compared to 0.118 for Type 2.[25]

The estimated parameters and fitted probabilities allow me to approximate the expected cost of inattention with a simulation. First, I use 1000 sets of Gumbel distribution draws for each taste shock to generate a realisation of borrower utilities from each product as defined earlier in equation (3.3). Second, I obtain 1000 sets of

25 These probabilities are calculated across 1000 simulated consideration sets for each borrower. In each simulation s, borrower i is deemed to consider a lender l if the predicted $\mathrm{P(Attention)}_{il}$ exceeds the simulation-specific independent random draw from the uniform distribution for this borrower-lender pair (R_{ils}).

Table 3.5: Summary inattention measures, split by type

	Mean among:	
	Type 1	Type 2
P(Attention)_{il}	0.641	0.669
$\text{P(Consider 1 lender)}_{i}$	0.143	0.118
$\mathbb{E}$[Utility forgone from missing best] :		
Initial interest rate equivalent (pp)	-0.306	-0.177
Income change equivalent (%)	1.224	0.731

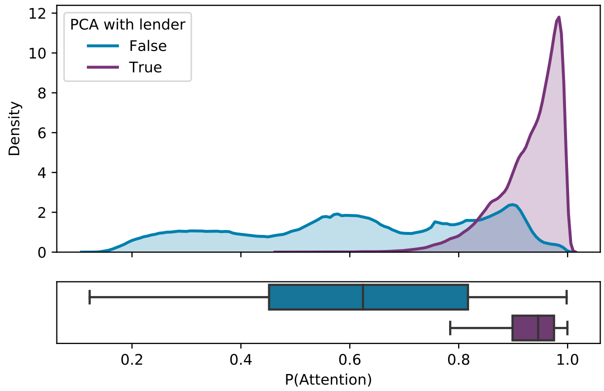
uniform distribution draws and compare them with the predicted P(Attention)_{il} to simulate which products get considered. For each of the 1000 scenarios, I calculate the difference between the maximum utility across all products available to borrower i and the maximum utility within their simulated consideration set. Naturally, if the 'best' product in that simulation is considered, the difference is 0. For each borrower, the expected cost of inattention is the average difference across the all scenarios. I then use estimated utility parameters and borrower demographics to convert the forgone utility into equivalent changes in the interest rate and annual borrower income which are reported in Table 3.5.

Individuals in Type 1 typically forgo larger utility improvements due not considering better products. Their average expected costs are equivalent to forgoing a 0.31 percentage point fall in interest rate, which in turn means around a 1.2% reduction in annual household income. For Type 2, the expected costs of inattention are considerably lower, equivalent to a 0.18 percentage point change in the initial interest rate or 0.73% of annual post-tax household income.

3.5.3 The role of familiarity

This section explores in more detail the powerful effect that a borrower's existing link with the lender though current product holdings ('lender familiarity') has on their attention towards that lender and the likelihood of choosing their products conditional on paying attention. As explained in the specification section (3.4.1.1), in the baseline model, a 'familiar' lender is one with which the borrower has an existing personal current account (PCA). This section's results about the importance of lender familiarity for attention and choice are robust to more complex specifications, for instance, those that take account of the number of existing PCA and non-PCA products with the lender.

Figure 3.5: Distribution of P(Attention)$_{il}$, by familiarity



3.5.3.1 *Familiarity and attention*

Having lender's familiarity has a large positive effect on their likelihood of being considered, shifting the whole distribution of the probabilities towards 1, as can be seen in Figure 3.5. Almost 9 out of 10 borrower-lender pairs that are linked by an existing PCA have P(Attention) of 0.8 or higher, whereas it gets above 0.8 for only 27% of pairs without an existing link.

Both types of borrowers are affected by familiarity when they decide whether to consider a particular lender. Its importance is, however, a lot more pronounced for Type 1. Those borrowers almost certainly consider lenders with which they have a relationship (mean P(Attention) of 0.97) and are much less likely to consider unfamiliar lenders (mean P(Attention) of 0.56). For Type 2, the mean likelihoods of considering familiar and unfamiliar lenders are less dramatically but still significantly different, at 0.88 and 0.63 respectively.

The differences in attention towards familiar and unfamiliar lenders summarised in Figure 3.7a for both types arise through several channels. First, there is the direct impact of the existing PCA relationship on attention. Even controlling for other drivers of attention in the model, the mere fact of the borrower i having a PCA with the lender l increases P(Attention)$_{il}$ by 24.7 percentage points for Type 1 and 21.2 percentage points for Type 2. This effect is further amplified by the length of this existing banking relationship. For Type 1, every additional year[26] of having the PCA with l increases the probability of considering them by 51.8 percentage points. In contrast, the length of the banking relationship does not have a statistically or economically significant effect on attention for Type 2 borrowers.

26 Beyond the initial six months required for the PCA with the lender to count as an existing relationship.

Table 3.6: Average marginal effects on P(Attention), by type

	Δ P(Attention)[a]		Corr with
	Type 1	Type 2	has PCA
=1 if has PCA with lender	24.70	21.24	1
Length of lender relationship (years)	51.84	0.17[b]	0.74
Branches within 5 mi (N)	1.45	0.49	0.11
Regional advertising spend (£per cap pcm)	1.97	-4.64	0.05
Lender size: lending volume in 2014 (£m)	0.63	0.19	0.13

[a] Marginal effects reported in percentage points for interpretability. All underlying parameters are significant at 1% level unless stated otherwise.
[b] Only significant at 5% level.

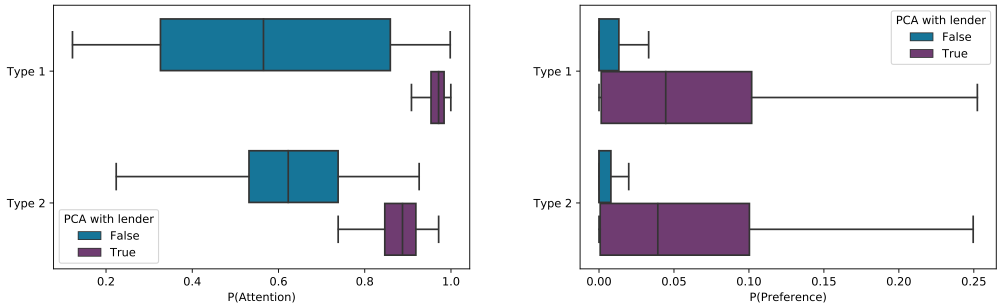
Second, there are additional characteristics of the each borrower-lender pair (also summarised in Table 3.6), which influence attention and are positively correlated with the likelihood that the borrower has an existing current account. These characteristics further contribute to the dramatic differences between familiar and unfamiliar lenders in Figures 3.5 and 3.7a. Both types of borrowers are somewhat more likely to consider lenders with more branches near their address, although attention is a lot more responsive to this for Type 1 borrowers (marginal effect of 1.45 percentage points per additional branch compared to 0.49 for Type 2). Lender's higher mortgage lending volumes in the past are also positively but weakly associated with greater attention for both borrower types. Curiously, lenders' advertising expenditure has the expected weakly positive marginal effect on $\mathrm{P(Attention)}_{il}$ for Type 1, but appears to be negatively associated with attention for Type 2 borrowers. Given that a £1 per capita per month change in advertising expenditure is a large change (approx. 65% of 1 standard deviation), the estimated effect on attention is not very economically significant even for Type 2. Advertising campaigns that anticipate and seek to counter (partially) slumping demand could be a possible explanation for the negative association with attention.

3.5.3.2 *Familiarity and preferences*

Lender familiarity affects choice through borrower preferences as well. As shown in Figure 3.7b, an existing relationship with a lender is associated with around 4.5 percentage point increase in median probability of the product offering the highest utility out of the whole choice set ($\mathrm{P(Preference)}_{il}$) for both consumer types. Curiously, none of the other familiarity and attention factors described in the preceding section – length of the relationship, branches, lender size or advertising – have economically significant effects on preferences.

Figure 3.6: Distributions of predicted probabilities, by lender familiarity and type

(a) P(Attention) (b) P(Preference)



Preferences for familiarity also translate into a higher probability of observing choices of products from familiar lenders, even when holding the attention channel constant. Table 3.7 reports the estimated marginal effects of lender familiarity on P(Choice)_{ij} through the preference channel. In this case, Type 2 borrowers are more sensitive to lender familiarity than Type 1. The average marginal preference effect of an existing PCA with the lender is to increase the probability of the lender's products being chosen by 4.34 percentage points for Type 2 and 3.33 percentage points for Type 1. Both are extremely large effects, more than doubling P(Choice)_{ij} from its sample average of 0.028.

Table 3.7 also shows the implied own-bank premiums that would be required to compensate the borrower for taking out an otherwise identical product from an unfamiliar lender without reducing their utility. The premium is equivalent to reducing the mortgage interest rate by 1.27 and 1.41 percentage points, respectively, for Type 1 and Type 2 borrowers. These interest rate changes are equivalent to reducing the annual mortgage payments as a percentage of income by 5 percentage points for Type 1 and 5.7 percentage points for Type 2 per year.

It is clear that borrowers being more likely to think about their own current account providers first is not the cause of the own-bank premium because it is separated by the model into a distinct attention channel. Other possible drivers of the premium include: (a) risk aversion with respect to some lender characteristics such as lending standards or service quality which are perceived to be more uncertain in unfamiliar lenders, (b) greater (perceived or real) effort costs of applying to a new provider for a mortgage, or (c) greater ongoing effort of managing accounts across multiple financial institution. In either case, the effects are clearly substantial and are not attenuated among otherwise apparently more sophisticated borrowers (Type 2).

Table 3.7: Marginal effects and willingness to pay for familiar banks, split by type

	Mean among:	
	Type 1	Type 2
Δ P(Preference)$_{il}$	3.44	4.55
Δ P(Choice)$_{il}$	3.33	4.34
Own-bank utility premium		
Initial interest rate change (pp)	-1.27	-1.41
Income change equivalent (%)	5.05	5.70

The fact that a non-price characteristic plays a significant role in borrower preferences also means that borrowers are not very price-elastic. The mean own-price demand elasticity in the sample is -3.64, suggesting that a 10% increase in interest rates would, on average, reduce demand by a third. Price elasticity varies a lot, however, across borrowers and products, with standard deviation of 1.73. This is not surprising given the amount of heterogeneity permitted by my demand model and the presence (or absence) of existing links with lenders for different borrowers. The implied marginal cost of an average loan is 1.33 percentage points (i.e. £133 on a £10,000 loan), but, again, with substantial variation across lenders and products. These estimated average marginal costs are broadly consistent with the size of the variable costs in the recent regulatory report on UK lenders (FCA, 2018b). Table B.2 in the appendix provides a breakdown of demand elasticities and marginal costs by lender size, loan-to-value band and interest rate type.[27]

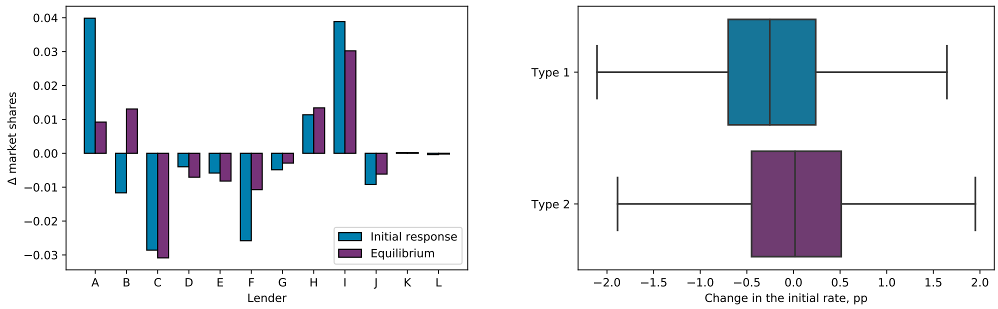
3.6 Counterfactual simulation: forced full attention

The extent of inattention documented in section 3.5.2 raises a question about the scope for an intervention to improve borrower welfare. To explore this I simulate the market effects of a hypothetical extreme policy where each borrower is made to pay attention to the whole choice set available to them (e.g. by making it mandatory to use a comprehensive comparison tool before taking out a mortgage). In practice, an intervention like this would likely also involve considerable search and time costs for borrowers arising from the larger number of options they would have to consider. It would also involve additional implementation costs. I abstract from these factors

27 Fixed costs such as IT, branch network, etc, tend to be very large in banking and can exceed costs that scale more directly with lending volume, such as funding costs (FCA, 2018b). As a result, the traditional mark-ups over marginal cost do not have a very meaningful interpretation in this context and are not reported.

Figure 3.8: Impacts of the counterfactual simulation with full attention

(a) Changes in market shares (b) Distributions of interest rate changes



in the current setting to consider the upper bound on the benefits borrowers might experience from the change. I simulate this change by turning off the attention channel (setting $\mathrm{P(Attention)}_{il} = 1$ for all i and l and letting consumer demand and lender pricing adjust to a new equilibrium.

In running this simulation, I assume that lenders' marginal costs, set of offered products, and characteristics of the existing products (other than the initial rate) are not affected by the intervention. There is no new entry or exit, and no changes to advertising or the branch network. Borrowers' preferences and existing PCA relationships with lenders remain the same.

Figure 3.9a, shows the initial and equilibrium effects on the market shares of individual lenders. Even under full attention, the market shares of most lenders appear little changed, with a small number of exceptions.The proportion of borrowers choosing familiar lender does, however, decline somewhat from 54% to 49%.

The effect on the equilibrium interest rates is more pronounced. Interest rates offered by lenders decline by 23 basis points on average across all mortgage products. The falls in the interest rates incurred by borrowers on their *chosen* products is a lot smaller, averaging just 4.3 basis points (because they were largely choosing better priced products to begin with).

The new policy affects the two types of borrowers very differently. As shown in Figure 3.9b, the consequences for Type 1 borrowers is broadly favourable: interest rates on their equilibrium mortgage choices fall for almost two thirds of this group with decline of 9.2 basis points on average. The mean cost saving is £130 per year. On the other hand, the interest rates increase for more than a half of the Type 2 borrowers. Consequently, this group has larger mortgage payments on average in the new equilibrium, on average paying 3.6 basis points (or just under £200) more

Table 3.8: Summary of average changes under full attention relative to the baseline

	Mean among:		
	Type 1	Type 2	All
Price effects:			
Δ interest rate (pp)	-0.092	0.033	-0.043
Δ annual payment (£)	-129.975	197.036	-1.214
Expected utility change:			
Equivalent Δ interest rate (pp)	-0.258	-0.023	-0.165
Equivalent Δ annual payment (£)	-289.523	-33.095	-188.554

in interest. Even though there are more Type 1 borrowers, the larger loan sizes in Type 2 mean that the changes in pound mortgage costs across the population as a whole broadly net out and remain constant on average.

The mechanism for rise in prices for Type 2 appears to be as follows. Some of the cheapest lenders prior to intervention are significantly more popular with Type 2 borrowers (both in terms existing PCAs and resulting mortgage choices), but not very likely to be considered by many borrowers, especially Type 1. As a result, those lenders enjoy a relatively larger positive demand shock from increased attention, allowing them to raise interest rates closer to market average while still gaining market share and retaining demand from their existing Type 2 PCA customers.

Considering all products could lead borrowers to discover options that were better for them on terms other than price and thus increase overall utility despite the very modest interest rate improvements. Comparing 1000 simulated counterfactual choices for each borrower under the new equilibrium and the original baseline suggests that the beneficial effects of full attention do, indeed, exceed the changes in prices alone.[28] The average expected utility increase for a Type 1 borrower is equivalent to a 26 basis point reduction in their initial interest rate, more than double the actual observed interest rate change for this group. This interest rate change is equal to a reduction of annual mortgage payments by £289.

28 The simulation follows a similar process to the one described in section 3.5.2 for calculating costs of limited attention. As before I generate 1000 independent random draws of the utility taste shocks, $\zeta_i j$ from the Gumbel distribution and uniform distribution draws to simulate which product get considered by each borrower under limited. For each draw and each borrower I calculate their maximised utility (a) under limited attention and the original product characteristics, conditional on the simulated consideration sets, and (b) under the new full attention equilibrium, with new equilibrium interest rates and other product characteristics. Each borrower's expected utility change from the new policy is the average difference between (b) and (a) across 1000 simulations.

On average, being able to discover products with better non-price characteristics under full attention improves utility for Type 2 just enough to compensate them for the rise in prices. The average change in consumer surplus for this group is equivalent to a decrease in mortgage costs by 2.3 basis points or £33 per year. Overall, the change to the full attention equilibrium results in average expected welfare gains equivalent to the interest rate being reduced by 16.5 basis points (or by 6.3% from the baseline average rate).

Importantly, the intervention is far from universally beneficial, even after accounting for consumer surplus from better matching on non-price characteristics. In fact, welfare declines in the new equilibrium for more than a third of the population (28% of Type 1 and 47% of Type 2 borrowers).

This simulation exercise highlights that even if an intervention to enforce full attention to alternatives without subjecting borrowers to search costs were feasible, it could still entail welfare losses for a significant number of borrowers. Even on average, the predicted improvements in welfare are perhaps less dramatic than would be expected of an intervention of this scale.

Most notably, the market structure and shares are barely affected by the mandated full attention, in important part because of the large role that existing links between borrower and lender play in shaping preferences – both for more and less 'sophisticated' types of borrowers. This means that if the policy motivation is concern about the concentration in the market, making consumers aware of alternatives, however clearly, is unlikely to be effective insofar as many of those borrowers have pre-existing links to incumbents.

The simulation results also show that in the context where existing links matter, even a costless hypothetical intervention to increase attention creates welfare transfers between consumer groups. The previously more inattentive borrowers benefit, on average, from the increased price competition among the more expensive lenders. But many of the more sophisticated borrowers, who searched more for the cheaper lenders (or were lucky to have an existing relationship with them) before the intervention, can lose out after full attention is enforced across the market, especially when price elasticity is low due to strong preferences for other factors (like lender familiarity). Given the demographic characteristics of the two types, the transfer is largely from the better-off consumers to the poorer ones, but there are also winners and losers within each type.

In practice, of course, there is also likely to be an additional (e.g. cognitive or time) cost of the enforced full attention even if search is simplified. Revealed preference would suggest that this cost is likely to be higher for those who are more inattentive ex ante, which could considerably reduce their welfare gains from the intervention as well.

3.7 Robustness

I explore several alternative specifications and data sub-samples to confirm the robustness of the results in this paper. For the sake of interpretability, ease of comparison between the different models, and computational efficiency, I estimate the standard (single borrower type) alternative-specific consideration models under these different approaches. Table B.3 shows the results of the alternative models alongside the single-type version of the baseline specification in this paper (1).

alternative specifications of familiarity I check that the estimated parameters, and especially the strength of the effects of lender familiarity, are robust to different specifications using two alternative models. In the first (model (2) in Table B.3), I keep the indicator for the existing link through personal current accounts as is, and include an additional indicator for the borrower having other credit products with the lender, such as credit cards, personal loans, mortgages, etc. In the second alternative (model (3)), instead of the binary indicators for the existence of the link with a lender, I instead control for the number of personal current accounts and the number of other credit products the borrower has with each lender. As can be seen from the regression results in Table B.3, these more sophisticated measures of the existing relationships between borrowers and lenders do not add much to the fit of the model or materially change the parameters on other product characteristics. Controlling for other products reduces the estimated effect of the link through current accounts slightly, but it still remains the main channel through which existing relationships affect attention and preferences.

first-time buyers only There could be a concern that borrowers who are not new to the housing market, and are taking out a mortgage to move from their existing property could have prior information about lenders through their previous mortgage loans. Those borrowers may have learnt about the lenders through their earlier (now refinanced) mortgage loans, but kept an open current account from that period. I can observe their 'vestigial' current accounts, but not the history of mortgage relationships that finished over 6 years ago. Thus I could misinterpret the effects of relevant learning from previously held mortgage products as more general preference for brand or lender familiarity acquired through current account links. To see the extent to which the results could be distorted by past experiences by 'home movers', I estimate the model baseline model specification in a sub-sample of first-time buyers: households who have not owned a property (and hence not had a mortgage) before. Model (4) in Table B.3 shows the results of this exercise.

The estimated parameters for first-time buyers are qualitatively (and often quantitatively) similar to those in the full sample for nearly all of the variables. The only notable exception is the distaste that first-time buyers have for fixing interest rates for more than two years, which is not found in the full sample. The likely reason for the differences is that there are penalties for moving house before the fixed rate period expires, first-time buyers are more likely to be buying a property to get on the housing ladder but with the expectation of having to move soon. These households tend to be younger than the sample as a whole (mean age of 31.8 vs 37.1), and thus are more likely to have their housing size and location needs change in the near future due to a job change or having (more) children.

The findings on lender familiarity reported in the paper are robust to using the first-time buyer sub-sample. They show an even stronger preference for lenders with whom they had an existing current account. Their estimated tendency to consider 'familiar' lenders more than 'unfamiliar' ones is also only marginally smaller.

3.8 Conclusion

This paper applies a novel combination of the latent class and limited attention multinomial logit models to explore the drivers of inattention and product preferences in the UK mortgage market. I identify two groups of borrowers, who broadly have characteristics associated with greater and lesser financial sophistication described in earlier research (in e.g. Lusardi, Mitchell and Curto, 2014). I find that borrowers who belong to the 'demographically' less sophisticated type tend to be more inattentive and are relatively more likely to focus their attention on 'familiar' lenders, with whom they already have a relationship through a personal current account.

I also find, however, that even after accounting for limited attention, mortgage borrowers exhibit a strong preference for products from 'familiar' lenders. Both types of consumers, effectively trade off borrowing cost savings of up to 5% of post-tax income for going to a lender with whom they have an existing relationship. This behaviour has implications for policy. As I show in the counterfactual scenario, even a hypothetical intervention that achieves full attention among borrowers without subjecting them to increased search cost, has an ambiguous effect in a setting with very loyal borrowers. Lenders' market shares change little, and improvement in prices is not dramatic. On average, prices paid and consumer surplus improve due to the previously less attentive consumers finding better deals. However, there is a significant minority (including nearly a half of the more 'sophisticated' type) whose welfare is reduced in the new equilibrium even before any search costs are taken into account. In the context of substantial and strong existing links to providers (often through other products), just making borrowers aware of alternatives has only a limited effect on their choices.

More generally, given the extent to which past choices of a personal current account provider shape borrowers' preferences and decisions about other products, it is important for policymakers and researchers to consider the 'portfolio' of consumers' product holdings rather than focus on individual markets. Likewise, when suppliers operate across a wide range of product lines, as is often the case in the finance and technology sectors, adopting a cross-market perspective could be necessary to understand their competitive behaviour and market outcomes.

The work in this paper could be extended in multiple ways. First, it could be insightful to consider the behaviour of brokers in intermediated transactions in the context of limited attention and preferences for familiar alternatives. Looking at both the direct and intermediated channels would also allow more comprehensive policy counterfactuals to be explored. Second, like all limited attention work in the alternative-specific consideration tradition (as defined by Abaluck and Adams (2017)), I have to implicitly assume that attention is not rival. For instance, having a lender's branch near the borrower's home has the same effect on attention regardless of how many other competitors have branches nearby. Search costs are also difficult to measure and conceptualise in this setting as each alternative is independent. To the extent identification permits, it may be instructive to revisit the topics in this paper using a model that combines sequential search with characteristics-based inattention.

4 The effects of mandatory mortgage advice: Evidence from a policy intervention[1]

4.1 Introduction

Choosing a mortgage product is one of the most important financial decisions for a household. The chosen product's features will directly affect the household's exposure to economic shocks, from interest rate changes to idiosyncratic events like unplanned relocations. Moreover, within sets of products with similar features and eligibility criteria there is a wide divergence in borrowing costs.[2] With mortgages being most households' largest liability by far, the choice of the right product thus has important consequences for both resilience and financial headroom.

As with many complex and important consumer choices, a market for mortgage advice has now emerged in most countries. Advisers and brokers promise to help consumers navigate the market, identify the best product for their circumstances and increase the chances of a successful application.[3] Often provided as part of a sale, mortgage advice typically requires a professional qualification and is bound by other regulatory constraints. Advice is sometimes mandated for product sales to certain types of consumers or, as is the case in our identification setting, all consumers.[4] Specifically, we exploit a nationwide policy intervention in the United

1 This paper is based on joint work with Jeroen Nieboer at the FCA and London School of Economics. Both authors contributed equally throughout the project.
We would like to thank Peter Andrews, Vimal Balasubramaniam, Richard Blundell, João Cocco, Karen Croxson, Stefan Hunt, and Ian Preston for their extensive comments and suggestions. We would also like to thank participants of the 2019 FCA-Imperial College conference on Household Finance in London. We are thankful to Adiya Belgibayeva, Teresa Bono, Damien Fennell, Claudia Robles-Garcia, and especially Simone Pedemonte at the FCA, for their contributions to the dataset and this research.
All views in this paper are those of the authors and do not represent the position of the FCA.

2 See Iscenko (2018) and Liu (2019) for the UK evidence. Considerable price dispersion in mortgage loans has also been documented in Canada (Allen, Clark and Houde, 2014b) and the United States (Bhutta, Fuster and Hizmo, 2019; Alexandrov and Koulayev, 2018).

3 See FCA (2018a) and Mysliwski and Rostom (2018) or Robles-Garcia (2019) for an overview of the role mortgage brokers play in the UK mortgage market.

4 Mandating advice is a common policy response to concerns that complex products are sold without regard for consumers' needs and circumstances. Following the prominent role of mortgage defaults in the Great Financial Crisis, many governments responded by issuing stronger advice requirements on mortgage sales - for example, the United States' Dodd-Frank Act (2015) and the European

Kingdom (UK) to estimate the effect of mandating advice on the population of consumers that, pre-intervention, obtained mortgages without getting advice. This intervention was part of the UK's Mortgage Market Review (MMR) in 2014.

What distinguishes mortgage advice from other types of guidance available to consumers is that it constitutes a tailored recommendation of one or more suitable mortgage product(s). Advice in our setting is subject to regulatory standards, and if the borrower follows the advisor's recommendation they have a right to compensation if the recommended product proves unsuitable. Rejecting the advisor's recommendation forfeits these statutory compensation rights. We therefore expect advice to have a strong steering effect on borrowers' choice of product, in line with the policy intention.

Our dataset comprises the universe of UK residential home purchase mortgage originations in the two years prior to and after the policy intervention, linked with Credit Reference Agency (CRA) data and the Moneyfacts mortgage product details database. These granular loan-level data allow us to control for key borrower and property characteristics that co-determine one's propensity to get advice, such as property price, credit score, age, income and employment status. Following the policy intervention, we see an increase in the proportion of advised sales from a steady base of three quarters to effectively the whole market. The consumer population of interest is the quarter of borrowers who, prior to the intervention, obtained a mortgage without advice. This borrower population is likely to be a mix of sophisticated individuals, who know exactly what products best meet their needs, and less savvy consumers who might not have realised they required assistance to make a good mortgage choice or had mistaken a sales pitch for advice in their best interest.

We use a difference-in-differences matching approach as described in, for example, Blundell and Costa Dias (2009), to estimate robust counterfactual outcomes for borrowers who were brought into advice by the the policy using our repeated cross-section data. We estimate the Average Treatment on the Treated (ATT) of advice requirement on key product features (e.g. fixed rate period) and on cost of borrowing. Overall, we find that one of the largest effects of the requirement was to substantially increase use of independent brokers among the group that previously did not take advice (even though advice as required by the policy can also be provided by lenders directly). Despite the growth in the use of brokers, however, we find only ambiguous and weak effects on the borrowing costs incurred by the treatment group.

Union's Mortgage Credit Directive (2014). In some countries, such as the UK and Netherlands, regulated mortgage advice was made mandatory.

We also find that the advice requirement materially changed consumers' product choices. Prior to the intervention, non-advised borrowers' choices were approximately evenly split between three product categories: fully variable rate mortgages, 2-year fixed rate mortgages and longer-term fixed (typically 5-year) fixed mortgages. The effect of advice requirement was to concentrate borrowers in 2-year fixed rate products, decreasing the shares of variable rate and 5-year fixed rate mortgages by 11 and 5 percentage points, respectively. Given the low and persistent interest rates during the period, this change has no obvious interpretation as advice correcting a previously widespread error in product choice.

This paper provides the first empirical evidence, to our knowledge, on the effects of a mandatory advice policy on consumers' financial outcomes. Since mandatory advice policies typically only affect a certain type of product, tracking (would-be) buyers can be prohibitively hard due to attrition and switching to alternative products. Even if a policy affects the whole market, there is the challenge of constructing a counterfactual for buyer types whose outcomes are only observed either before or after the intervention. We believe these issues are unlikely to apply to the policy intervention we study, for several reasons. First, the treated group of prospective borrowers did not see an increase in fees under the new rules, as advice does not incur a separate charge.More important, since we only study house purchases and the house purchase decision effectively implies the need for a mortgage, it seems very unlikely that the introduction of mandatory advice would have material effects on the extensive margin by changing the rate at which households participate in the mortgage market. Finally, surveys carried out before the intervention indicates that many borrowers were unable to distinguish regulated advice from sales talk or support with their mortgage application (FSA, 2009; FSA, 2010). It is thus likely that some proportion of the treated population thought they already received advice before the intervention (or were unaware of regulated advice being available), the fact that they did not strictly receive a personal recommendation being lost on them.

Our main contribution is to the literature on mortgage distribution, with a particular focus on regulatory intervention. Part of this literature focuses on the role of brokers in the mortgage market, which promise to reduce search costs and frictions in this complex market (Gavazza, 2016; Salz, 2017). Woodward and Hall (2010, 2012) find that less-informed consumers pay higher brokerage fees and that many consumers would have benefited from visiting multiple mortgage brokers. Jiang, Nelson and Vytlacil (2014) find that mortgage brokers originated loans with a higher probability of default in the years leading up to the 2007 global financial crisis. These findings suggest some caution is in order if mandatory advice is expected to increase market share for brokers, although Mysliwski and Rostom (2018) and Robles-Garcia (2019) estimate with UK data that brokers are still net-beneficial due to their role in the information search process. Moreover, even advisers employed by a lender may steer consumers towards certain mortgage products that are more profitable for the lender (Guiso et al., 2018; Foà et al., 2019).

This paper is also related to the empirical literature on the value of financial advice, much of which focuses on investment advice. A key stylized fact in this literature is that advised portfolios tend to differ from self-directed portfolios held at the same financial institution. Many of these papers find that advised portfolios are relative underperformers (Bergstresser, Chalmers and Tufano, 2008; Hackethal, Haliassos and Jappelli, 2012; Karabulut, 2013), driven by higher trading costs and greater investment in actively managed funds (Hoechle et al., 2017; Foerster et al., 2017). Linnainmaa, Melzer and Previtero (2018) show that financial advisers themselves hold similar portfolios, suggesting that attributing these differences to adviser commissions may be too simplistic. Although these papers raise important questions about the value of financial advice, they are still somewhat removed from the ideal experiment of offering financial advice to some investors whilst withholding it from others.[5] Studies that use (quasi-)experimental variation to identify effects of advice find that self-directed investors do not follow advice offered to them (Bhattacharya et al., 2012) and that those in employee pension plans only benefit from advice if their default portfolio allocation is poor value (Chalmers and Reuter, 2012). Collectively, these studies show some of the challenges of improving people's outcomes through financial advice.

4.2 Background

To provide context, it will be helpful to briefly mention some distinctive features of the UK mortgage market. The most notable feature is the absence of interest rates fixed until maturity. Residential mortgages typically have interest rates fixed for a relatively short period (2-5 years), followed by a (higher) variable reversion rate for the remainder of the contract. The reversion rate can be varied by the lender at their discretion, but in practice tends to move broadly in line with the Bank of England base rate. Two-year fixed rate contracts are by far the most common

5 A notable exception is the mystery shopping study by Mullainathan, Noeth and Schoar (2012), which confirms financial advisers' tendencies to recommend frequent trading and actively managed funds.

mortgage type; they accounted for almost two-thirds of all home purchase mortgage lending in 2016 (the last year of our sample).

A smaller part of the market (circa 10% in 2016) is made up of fully variable rate mortgages that typically track the Bank of England base rate for the contractual term. But many of these mortgages, too, often offer more attractive terms in the early years of the contract (typically in the form of a lower spread over the underlying rate). In this paper we refer to the initial years of the mortgage during which the borrower benefits from special terms, such as a fixed rate or a discount on the variable rate, as the *deal period*.

It will usually be in the consumer's interest to refinance the mortgage immediately at the end of the deal period. This is driven by the interest rate reset and the fact that early repayment penalties, another common feature, no longer apply after the deal period.[6] In contrast to sluggish refinancing observed in other countries (USA: Agarwal, Rosen and Yao (2015), Keys, Pope and Pope (2016); Denmark: Andersen et al. (2018)), most borrowers in the UK refinance promptly: 77% of consumers on fixed rate deals refinance within 6 months of their deal period ending (FCA, 2018a).

4.2.1 Product selection

At any given point in time, UK lenders post menus of products on offer. Borrowers therefore do not bargain with lenders over price and product features - they simply choose a product for which to apply. A product is characterized by its features, price structure and eligibility criteria. The different combinations of all these options result in a large number of products: Iscenko (2018) shows that the average prime mortgage borrower in 2015 was eligible on average for more than 400 products. Besides publicly posted eligibility criteria, such as loan-to-value ratio bands and borrower types (e.g. first time buyer), lenders also employ eligibility criteria that are less transparent to the borrower (e.g. past credit impairments or unusual property types). Uncovering the latter type of criteria will require information search or a conversation with a broker or a lender.

Choosing a mortgage product in this market is complex. Prospective borrowers need to choose which product features fit their preferences and circumstances, identify products for which they are eligible (including on criteria that are less comparable across lenders) and then compare costs of the suitable remaining options. In addition to the time-varying interest rates discussed above, products normally include at least one (and often more) fixed fee payable at mortgage origination, stepped early

6 The exit costs after the deal period expiry are typically low, just covering administrative charges of closing an account.

repayment penalties and account closure fees. Lenders may also offer promotional products that waive certain fees or promise the borrower a cash refund at origination or later in the contract.[7] Unsurprisingly, surveys of borrowers suggest many borrowers struggle with this complexity: 34% of recent borrowers thought there was "too much information to deal with", 31% stated that mortgage products were not simple to understand and 26% stated that it was not easy to compare mortgages from different lenders (FCA, 2017).

4.2.2 The role of advice

Various tools and services are available to help prospective borrowers with their choice. There is publicly available information on various platforms, including mortgage calculators provided by lenders and not-for-profit organizations. Some parties may also support prospective borrowers with their application paperwork. As long as these services do not include a personalised product recommendation, they do not constitute advice in a regulatory sense. The prospective borrower could, in principle, thus have interactions with various parties, including lenders and brokers without receiving advice.

Many borrowers will want more than just information provision and help with their application - they want advice. An adviser helps the borrower with product search and comparison, considers the likelihood of a successful application based on private information about different lenders, and finally recommends one or more product(s) based on the borrower's stated needs. Advisers are either employed by lenders or by brokerages, which means advice is tied to a (prospective) product sale. Brokerages receive commissions from lenders based on a percentage of the total loan amount, and both lenders and brokers may charge flat arrangement fees directly to the consumer (typically conditional on a successful application). Both lender-based and brokerage-based advisers thus have clear incentives, albeit slightly different in nature, to recommend to the borrower one or more product(s) that she is likely to qualify for. Advisers need to hold a relevant qualification and their advice also needs to meet a regulatory standard, backed up by statutory adjudication and compensation rights for the consumer.Should the consumer ignore the advice and apply for a product that was not recommended, she will waive her adjudication rights.

In the pre-intervention period, around three quarters of those obtaining home purchase mortgages got advice - roughly 80% of first time buyers and 70% of returning

7 To enable comparisons, lenders are legally required to disclose borrowing costs as an Annual Percentage Rate (APR). But, since the calculation of this 'mandatory' APR assumes that the borrower does not remortgage, and the vast majority of borrowers refinance well before maturity, this metric is of little practical use.

buyers. The fact that returning buyers are less likely to opt for advice suggests that experience plays a role in the demand for advice. This is consistent with prior research on the UK market, which also finds that consumers are more likely to seek advice if they are less confident and less experienced (ESRO, 2015). Lachance and Tang (2012) find in US data that mortgage borrowers are more likely to seek advice when they are younger, female, married, have more years of education and higher incomes. Some aspects of the demand for advice may be effectively random, such as whether the borrower encounters a likeable mortgage adviser or whether the selling real estate agent puts pressure on the borrower to receive advice.[8]

Lender-based mortgage advisers' recommendations are restricted to products offered by the lender, whereas a broker can recommend products from different lenders. Since some products are only available direct from the lender, the lender set is not always a subset of the broker set (although some brokers may informally advise the consumer to go direct instead). In general, however, brokers will offer a wider range of choice to borrowers. Since product search and comparison is central to brokers' business, brokers already tended to make the vast majority of sales on an advised basis (99% in the pre-intervention period). Lenders, by contrast, made a substantial number of sales without advice (61%). The demand for advice in the pre-intervention period is thus intertwined with the demand for broker-specific services, a fact we take into account when constructing our propensity score model.

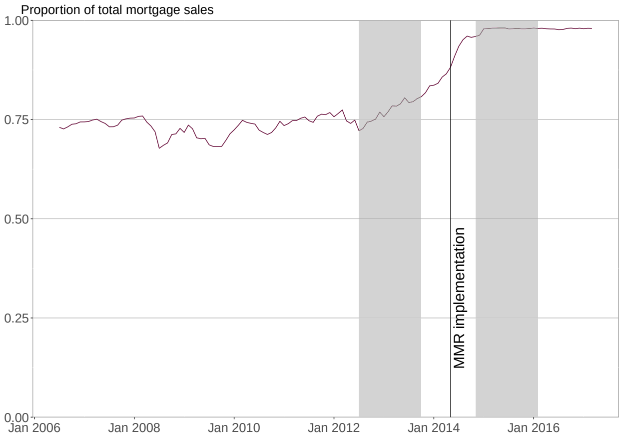
4.2.3 Mortgage Market Review (MMR)

The MMR was a substantial package of regulatory reforms implemented by the FCA in April 2014. The MMR introduced a number of changes to the rules on mortgage sales to retail customers, including a strengthening of the responsible lending (affordability verification) requirements for mortgage lenders and – the focal point of this research paper – the requirement that consumers receive regulated advice with every 'interactive' mortgage sale. Interactivity is defined as "spoken or other interactive dialogue ... [including] SMS, mobile instant messaging, email and communication via social media" and thus seems to encompass the overwhelming majority of sales.[9]

8 For instance, 26% of respondents to the FCA 2017 Financial Lives Survey survey reported feeling that the broker they used was recommended by their estate agent.

9 The definition of interactive sales only exempts mortgage sales conducted entirely by post or on-line. Consumers were effectively discouraged from remaining solely in those channels, however, leading to non-advised sales being only 1.6% of the market. In addition, the MMR rules allow high-net worth individuals and mortgage professionals to opt out of the advice requirement. See FCA Handbook, MCOB 4.8 for more details.

Figure 4.1: Use of regulated advice over time



Note: The figure shows the proportion of advised mortgage transactions among first-time buyers and home movers over time. The grey shaded areas show the pre- and post-MMR periods we use to estimate treatment effects of advice in the baseline model. See section 4.3.2 for more information about the dates used for estimation.

Figure 4.1 shows the proportion of mortgage borrowers receiving regulated advice over time. Note that the Figure, in line with the rest of our analysis, only describes mortgage contracts taken out to purchase a property (first-time buyers and movers).

Before the MMR, when receiving advice was optional, the proportion of borrowers receiving advice was stable at around three quarters of sales. This proportion starts to increase gently in 2013, followed by a sharp increase before reaching a plateau around 98% in late 2014. The greatest increase in advice takes place just after the official MMR implementation date of 26 April 2014 (represented by a vertical line in the graph). The MMR's advice requirement thus appears to have affected around 1 in 4 mortgage borrowers — the proportion of consumers who previously (pre MMR) obtained mortgages without advice.

4.3 Data

4.3.1 Sources

Our data cover 4 years (1 July 2012 to 30 June 2016) and are based on a merge of three main data sources at the transaction level:

1. **Product Sales Data (PSD):** the universe of regulated residential mortgage originations reported by UK lenders to the UK financial regulator (FCA);

2. **Credit files:** borrowers' credit files at mortgage application, including credit score, provided by a major Credit Reference Agency;

3. **Moneyfacts:** proprietary daily data on all available mortgage products in the UK market purchased from Moneyfacts, a commercial supplier.

We also supplement our main datasets with an indicator variable for newly built homes (derived from the UK land registry, a public database of property sales) and several postcode-level variables from the 2011 UK Census to proxy for borrower sophistication (educational attainment) and socio-economic conditions (average employment and socio-economic status).

Our dataset excludes mortgage originations due to remortgaging, as we do not have full visibility of refinancing activity.[10] Our analysis therefore focusses on first-time and returning borrowers obtaining a mortgage to finance the purchase of a residential property.

4.3.2 Estimation periods

Our analysis of the effects of the MMR, is computed by comparing outcomes in pre- and post-MMR groups of consumers. We define the 6 months either side of the official MMR implementation date (April 26, 2014) as the implementation period. This is the period in which most of the growth in advised sales took place, and the lenders were in transition to complying with the new requirements. To avoid our estimates being distorted by this transition, we exclude the implementation period from our estimation. Our pre- and post-MMR observation windows therefore consist of 16 months of data before and after the implementation period.24 This gives us a pre-MMR period between July 2012 and October 2013 and a post-MMR period between November 2014 and February 2016, illustrated by the shaded areas in Figure 4.1. [11]

4.3.3 Summary statistics

We start with the relevant PSD population containing approximately 2 million mortgage originations over four years. After omitting observations that cannot be matched to one or more of the necessary datasets, have missing data, or lie within

10 PSD for the relevant period does not include data on refinancing with the current lender ('internal remortgaging' and product transfers).

11 As discussed in section 4.5.2, our findings are robust to alternative assumptions about the length of the excluded implementation period.

the implementation period, we are left with the estimation samples of 601,683 for cost outcomes and 424,937 non-cost (product characteristics) outcomes.[12]

Table 4.1 shows key demographics and loan characteristics for our sample. To illustrate changes over time, we aggregate data for two periods: a pre-MMR period (July 2012 to October 2013) and a post-MMR period (November 2014 to June 2016). To provide context for our first research question, it is instructive to look at the non-advised consumer sample before the MMR. Note that these consumers are less likely to be first-time buyers, are slightly older, have higher incomes, take out bigger loans for more expensive properties and are much less likely than advised consumers to use the services of an intermediary. Not shown in the table, but easily constructed from the reported figures, we find that non-advised consumers have lower LTI (Loan-To-Income) and LTV (Loan-To-Value) ratios than advised consumers.

Table 4.1: Summary statistics for the main demographic variables

	Pre-MMR		Post-MMR
	Advised	Not advised	Advised
First-time buyer, %	44.7	35.4	42.9
Joint, %	62.0	60.3	64.4
Age, years	36.2	37.2	36.5
Income, £k	52.2	62.7	60.8
Loan value, £k	158.5	173.3	188.6
Property value, £k	231.4	279.0	278.3
Credit score (normalised)	0.63	0.64	0.64
Previous products, nr	2.50	2.55	2.68
N	130,887	27,655	259,208

4.3.4 Outcome variables

Our combined dataset allows us to calculate a range of outcome variables for the borrower. First, we measure key product characteristics: whether the mortgage had a fixed interest rate for the deal period, the length of the deal period and the total mortgage term. An adviser may take a different view to an non-advised borrower as to which features are more appropriate, for example based on an assessment of the need for stable monthly payments given the borrower's financial situation. In addition, we measure whether the mortgage was sold by a broker or directly by a lender.

12 The cost and non-cost outcome sample sizes differ because the estimation of non-cost outcomes does not require Moneyfacts data.

Second, we calculate three measures of the cost of borrowing: (a) the Annualised Percentage Rate (APR) over the deal period of the mortgage product, (b) the 5-year APR on a ”no remortgaging” basis (consumers are assumed to be inert and revert to the applicable Standard Variable Rate at the end of their deal period), and (c) the monthly mortgage payment during the deal period. [13] Both APR measures use Moneyfacts data on all relevant product fees (such as arrangement, booking and discharge fees) to represent the full cost of borrowing, including remortgaging expenses. They exclude contingent fees, such as late payment fees and early termination charges. The Bank of England interest rate was constant during our sample period at 0.5% throughout our sample period, which simplifies the comparison of borrowing costs before and after the intervention.

We also apply the methodology from Iscenko (2018) to identify mortgage product choices that were strongly dominated by other available alternatives. A mortgage choice is considered to be dominated when at the point of making the choice the borrower was also eligible for at least one product with had the same non-price features as the chosen mortgage but was strictly cheaper on cost: rates and fees no higher than the chosen product and at least one that was strictly lower. Furthermore, for a choice to be strongly dominated the consumer needs to incur excess costs of at least £250 per year and at least 5% of their annual mortgage cost compared to what they would pay on the cheaper but otherwise comparable available alternatives.

4.4 Method

4.4.1 Difference-in-difference matching

The MMR advice requirement meant that nearly a quarter of prospective borrowers, specifically those that had previously been willing and able to obtain a mortgage without receiving advice, would now receive advice. Given that the function of advice was not substantively changed by the policy, it seems a fair assumption that these consumers would not have received advice in the absence of the intervention. The question of policy interest, therefore, is how the advice requirement has affected their outcomes in the mortgage market.

Given the setting, we use difference-in-difference matching, as described, for example, by Blundell and Costa Dias (2000, 2009) as our policy evaluation method. Our interest in the changes in outcomes for the different groups over time favours

13 The deal-period APR measure is also very close (but not exactly equivalent) to a 5-year APR under the assumption that the borrower remortgages to a product with the same terms once the deal period expires.

Table 4.2: Summary statistics for the main outcome variables

	Pre-MMR		Post-MMR
	Advised	Not advised	Advised
Cost of borrowing			
Monthly payment, £	831.5	927.1	892.4
Deal period APR, %	3.89	3.82	2.87
5-year APR, %	4.06	3.94	3.44
Pr(Strongly dominated), %	10.4	6.7	18.0
Other characteristics			
Pr(fixed rate mortgage), %	88.4	74.8	94.2
Pr(2-year fixed rate mortgage), %	48.4	29.3	56.9
Pr(5-year fixed rate mortgage), %	27.5	33.3	28.6
Mortgage term, years	25.66	23.61	25.68
Pr(use a broker), pp	67.6	2.0	66.8

a difference-in-difference (DID) approach. However, the treated and control groups are not directly distinguishable in one of the periods, and when there is freedom to opt out of advice, it is reasonable to assume (as discussed in more detail in section 4.4.2 below) that selection into 'treatment' (i.e. receiving advice) is not random. But our extensive data about the consumers and their situation make it plausible that this selection is occurring on observables.

These characteristics are all closely aligned with matching DID in repeated cross-sections but in the inverted form. If the pre- and post-MMR periods are reversed (full advice, followed by freedom to opt-out) and the pre-MMR non-advised are designated as the treated group, the standard matching DID with would estimate the treatment effect of 'opting out from advice'. We are just interested in the negative of that as the impact of mandated advice (i.e. the opt-out option being withdrawn from those who used it).

4.4.1.1 *Identification*

We define outcomes as follows,
$Y^A(i,t|D)$ = outcome for individual i at time t after receiving advice;
$Y^N(i,t|D)$ = outcome for individual i at time t without advice;
where
$t = 0$ is the pre-intervention period;

$t = 1$ is the post-intervention period;
$D \in \{A, N\}$ indicates treatment status (Advice or Not) at $t = 0$.[14]

Our identification problem centres on constructing the counterfactual for the treatment group (which we refer to as the non-advised population). We wish to compute an estimate of how the outcomes of individuals in the non-advised population change depending on whether they do or do not receive mortgage advice,

$$ATT_t = \mathbb{E}\left[Y^N - Y^A | D = N\right] = \mathbb{E}\left[Y^N(i, t|D = N)\right] - \mathbb{E}\left[Y^A(i, t|D = N)\right] \quad (4.1)$$

The ATT at $t = 1$ can be rewritten as,

$$\begin{aligned} ATT_{t=1} = &(E[Y^N(i, t = 0|D = N)] - E[Y^A(i, t = 1|D = N)]) - \\ &(E[Y^N(i, t = 0|D = N)] - E[Y^N(i, t = 1|D = N)]) \end{aligned} \quad (4.2)$$

We can now see the identification problem more clearly. Two hurdles need to be overcome to construct a valid counterfactual. First, we do not actually observe the same person at two different points in time - we have repeated cross-sectional data. Treating the selection problem as one on observables, we match consumers in the non-advised population to those in control populations on the propensity score of taking up advice given their characteristics X. We represent this henceforth by conditioning on the predicted propensity score $p(X)$.

Second, we do not observe one of the expectation terms in the ATT equation and therefore cannot compute the estimate without making a further identifying assumption. We opt for the following conditional common trend (unconfoundedness in trends) assumption,

$$\begin{aligned} &E[Y^N(i, t = 1|p(X), D = N) - Y^N(i, t = 0|p(X), D = N)] = \\ &\qquad E[Y^A(i, t = 1|p(X), D = A) - Y^A(i, t = 0|p(X), D = A)] \end{aligned} \quad (4.3)$$

We thus assume that controlling for propensity score, non-advised outcomes in the non-advised population and advised outcome in the advised population follow a common trend over our implementation period. In Figures C.1 and C.3 in Appendix C, we plot the pre-intervention means of these two groups, conditional on their

14 It is not necessary to specify whether individuals receive advice at $t = 1$ (post intervention), since all individuals in the post-intervention population were treated with advice.

propensity score quartiles. Although this is not a statistical test, it allows us to investigate the plausibility of this assumption with data from our pre-treatment observation window. It appears that the common trends assumption is plausible, although for the likelihood of broker usage the trend over time for the non-advised population is hard to assess and the trends in the rate of dominated choices are not very stable.

Under the common trend assumption stated above, the ATT of mandated mortgage advice can be recovered by the following equation,

$$\begin{aligned} ATT_{t=1} = & (E[Y^N(i,t=0|p(X),D=N)] - E[Y^A(i,t=1|p(X),D=N)]) - \\ & (E[Y^A(i,t=0|p(X),D=A)] - E[Y^A(i,t=1|p(X),D=A)]) \end{aligned} \tag{4.4}$$

Note that our approach additionally requires sufficient common support in the treatment and control populations, in the pre-MMR cross-section and over time ($0 < Pr(D_i = 1|X_i = x) < 1$). In practice, this boils down to having a sufficient number of observations to match on propensity score across groups. As we show in section 4.4.2 below, this common support or overlap requirement is satisfied in our context.

4.4.1.2 *Estimation*

In practice, we estimate the ATT through the following sequence of steps.

1. Estimate the propensity score model on pre-MMR data to obtain propensity scores to take advice as a function of borrower, loan, property and location characteristics (X_i) as described in mode detail in the next subsection. Predict estimated propensity scores for both post- and pre-MMR borrowers. ($\Rightarrow \hat{p}(X_i)$)

2. Run a Nadaraya-Watson kernel regression with a Gaussian kernel and Silverman (1986) rule-of-thumb bandwidth selection rule to match the pre-MMR borrowers to post-MMR ones on propensity scores $\hat{p}(X_i)$. Obtain the expected post-MMR outcome for each pre-MMR borrower ($\Rightarrow \hat{Y}_{i,post}(\hat{p}(X_i))$).

3. Calculate the individual changes between the observed pre-MMR outcomes and the predicted post-MMR outcomes for the individual with the same characteristics. $\Delta\hat{Y}_i(\hat{p}(X_i)) = \hat{Y}_{i,post}(\hat{p}(X_i)) - Y_{i,pre}$

4. Run the second Nadaraya-Watson Gausian kernel regression with Silverman (1986) rule-of-thumb bandwidth to compare $\Delta\hat{Y}_i$ for each non-advised pre-MMR borrowers to advised pre-MMR borrowers. ($\Rightarrow \Delta^2\hat{Y}_i$)

5. Take the mean of $\Delta^2 \hat{Y}_i$ is for the pre-MMR non-advised sub-sample. Under unconfoundedness in trends and common support, is the ATT for the outcome of interest Y.

We use cluster bootstrapping (clustered by year-quarter and postcode area) with 100 replications to calculate standard errors around the ATT estimates. Bootstrap is a valid approach in our context because we use kernel-based matching (Abadie and Imbens, 2008).[15]

4.4.2 Propensity score model

Our implementation of matching on the propensity score has the joint benefits of dimension reduction and controlling for the treatment decision by the (estimated) propensity to get advice. To model this propensity, we must first ask what factors affect consumer choice on whether to use an adviser.

As is standard with treatment effect evaluation methods (Blundell and Costa Dias, 2000; Blundell and Costa Dias, 2009), the first step to answering this question is to set out a decision rule that determines whether consumer i chooses to take advice. Using the notation set out above this decision can be summarised as,

$$D_i = \begin{cases} A, & \text{if } E_i[Y^A(i,t|D_i = A)] - E_i[Y^N(i,t|D_i = N)] \geq e_i^A + \epsilon_i \\ N, & \text{otherwise} \end{cases}$$

This means that a consumer i will choose to get advice if its expected benefit exceeds the cost of receiving advice, modified by a random variation in preferences. For practical applications including our propensity score literature, we think of the selection decision as a linear index model which can be represented by:

$$D_i = A \iff X_i \phi + v_i \geq 0$$

where X_i are observable mortgage borrower characteristics and other aspects of the sale that affect benefits or effort cost of advice or both (and their interactions), ϕ are coefficients, and v_i is the residual component in the decision to take advice that is driven by factors we do not observe. As suggested above, for difference-in-differences matching to be valid, we need the change in Y_i^A and Y_i^N over time to be independent of D_i (and thus v_i), conditional on all X_i.

15 Abadie and Imbens (2008) demonstrate that bootstrap inference is invalid in nearest-neighbour matching. They explicitly state, however, that kernel-based methods to not have the characteristics that drives their bootstrap failure result, and so bootstrap is likely to be valid for those methods.

In our case of modelling the pre-MMR decision on whether to take advice, there are two additional important considerations. First, consumers who use intermediaries virtually always use advice, whereas those who go directly to a lender vary in their advice choices. This means that selection into intermediation is nested in thinking about selection into advice as it appears very difficult to opt out of advice when using an intermediary. Second, recall that not receiving or following advice meant that the consumer would lose the right to seek redress from the Financial Ombudsman if their chosen mortgage proved unsuitable (but keep rights to redress for other reasons). This means that the benefits of advice might include not just a personalised recommendation to help with product type choice but also the value of the additional regulatory protection.

We discuss the potential confounders that affect the benefits of this personalised recommendation (and regulatory protection) *and*, potentially, the outcomes of interest (Y_i) in the next sub-section about the control variables we use in the propensity score model. It is equally important for the validity of our approach, however, that there are also multiple factors that may affect whether the borrower takes advice without an obvious direct link to our outcomes of interest (let alone to the change in those outcomes over time). Many of these factors 'shift' the effort and time costs of obtaining advice. For instance, a consumer might not feel they need advice but prefers to clarify something about the terms of their chosen product in the branch before they apply for the mortgage. The ability of bank staff available at the time to answer these specific queries could make a difference between an advised and non-advised transaction. If advice is only weakly preferred, another example of a situational source of randomness is the availability of appointments with an adviser in the chosen branch and, convenience of the wait time or how strongly their estate agent pushes an affiliated mortgage broker/advisor. Finally, the consumer's trust in the value of advice can be affected by exogenous shocks, e.g. reported experiences in their social network.

Conditional on the large number of observed demographic and situational characteristics we include in the model, factors such as bank branch staff expertise and availability of appointments provide an important source of random variation in the uptake of advice that makes it possible to satisfy the common support assumption that underlies our matching approach.

4.4.2.1 *Control variables*

In this section we explain the borrower, property and area characteristics we use to estimate the propensity score for getting advice because of their possible direct effects on outcomes, such as price and product choice. The control variables we use are also listed in more detail Table C.1 in the appendix.

The choice of variables for the model is informed by the conceptual selection model we set out above, as well as specific variables suggested in the relevant literature on demand for financial advice.[16]

First of all, the consumer's need for advice will depend on experience, knowledge of the market and financial capability more generally. We therefore include in our model demographics such as age, gender, education (measured directly for those with postgraduate titles and indirectly through postcode-level educational attainment and unskilled worker shares) and whether the mortgage is jointly held (i.e. marital and relationship status). We also include behavioural variables that earlier literature (e.g. Scholnick, Massoud and Saunders, 2013; Shen, 2016) links to consumer's experience with financial products and sophistication in using them: the expensive practice of taking out cash advances on one's credit card and whether the consumer has been making minimum payments on their credit card. Finally, the consumer may be more daunted by choice and likely to seek advice if eligible for many different products. We therefore include the number of products available to the borrower as a proxy for choice complexity.

Furthermore, the value of information on the products available to the consumer will depend on the credit commitment relative to income (LTI) and property value (LTV), plus concerns consumers may have about their eligibility. The variables representing these concerns in our model are income, employment status, credit score, an additional impaired credit flag, property value, LTI and LTV. Income and borrower type (first-time buyer or a home owner moving to a new property) also proxy for opportunity cost of time. Note also that property price is likely to be a good proxy for household wealth and credit score is also likely to reflect consumer ability to manage their finances (in addition to measuring their riskiness to the lender).[17]

Our model also accommodates an additional dimension of eligibility and potential need for advice: lenders can be selective on the kind of property they will lend

16 Existing research on demand for advice and intermediation typically does not distinguish between these two services, which is helpful given that intermediation decision is linked to the choice to receive advice in our application. The literature also typically focuses on financial advice on investments rather than mortgages, but the relevant demographics are likely to be similar. For a more detailed discussion, see Nieboer, Dolan and Vlaev (2017) and references within.

17 Given that one of our matching dimensions is longitudinal, we normalise variables that may change over time (e.g. due to inflation) to the borrower's percentile in the distribution of the variable within the calendar quarter-year of their property transaction. We apply this conversion to measures of income, LTI, LTV and property value.

on. This includes newly built properties, properties in areas with low rates of home ownership and properties that have idiosyncratic features that make them less desirable. We control for these property features either directly – in the case of newly built properties – or by proxy: various postcode-level characteristics that correlate with the rate of home ownership. We also create a dummy variable ('problem property proxy') that indicates whether the property has a significantly lower value than other housing of the same size and in the same postcode area).

As summarised in Nieboer, Dolan and Vlaev (2017), existing research suggests that the likelihood of using a financial adviser can vary with financial knowledge, income, age, education, gender, marital status, employment status and stakes (e.g. amount invested). Studies based on survey data also report an association, albeit small, between stated attitude to risk and demand for advice (Robb, Babiarz and Woodyard, 2012). We are able to control for all these factors directly, with the exception of risk aversion. However, risk aversion has been shown to correlate with demographic variables in our model – notably gender, income, marital status, self-employment and age (Halek and Eisenhauer, 2001). Moreover, our measures of the borrower's credit history and choice of leverage would also be affected by their risk appetite. It therefore does not appear likely that conditional on observed behaviour and relevant demographics, the residual variation in borrower risk aversion is large or correlated with mortgage outcomes over time in a way that would materially threaten DID matching identification.

4.4.2.2 *Propensity score estimation*

Our propensity score model is fully interacted and allows for non-linearity in numerical variables by incorporating their squared terms. To represent our model specification, it will be convenient to split the explanatory variables into two sets based on variable type: X_1 contains all continuous variables and X_2 contains all categorical variables in binary form (i.e. variables with $k > 2$ categories have been transformed into a set of $k-1$ binary dummy terms). We can now write our model as follows,

$$Pr(Advice)_i = \alpha + \sum_{x \in X_1} (\beta_x x_i + \gamma_x x_i^2) + \sum_{x \in X_2} \delta_x x_i + \sum_{z_1 \in Z} \sum_{(z_2 \neq z_1) \in Z} \xi_{z_{i1}, z_{i2}} (z_{i1} \times z_{i2}) + \chi_i \tag{4.5}$$

where $Z = X_1 \cup X_2$ represents the full set of interaction terms and χ_i is a postcode region fixed effect.

We estimate the propensity score model using the LASSO (Tibshirani, 1996) logit regression, as implemented in Friedman, Hastie and Tibshirani (2009). This method has the advantage of performing fully data-driven variable selection on our initial set of 570 covariates. LASSO regression has one regularisation parameter, λ, that sets the penalty on non-zero coefficients in the model and therefore determines how many variables get 'selected'. We set λ in an entirely data-driven way by using a standard 10-fold cross-validation algorithm (explained in e.g. Athey and Imbens, 2019). Our choice of the LASSO estimator over other alternatives is motivated by the need to choose between a large number of potential covariates without arbitrary assumptions. Use of machine learning techniques for propensity score estimation is, however, also recognised in the matching literature as a potential source of methodological improvement (Imbens, 2015).[18]

Table 4.3 shows the predictive performance of our model, estimated on all available pre-MMR observations (mortgage sales between July 2012 and November 2013) in our estimation sample.[19] Grouping the borrowers by decile of propensity score, we construct two measures of the model's ability to discriminate between those who are more or less likely to get advice: (a) *% non-advised*, which is the proportion of borrowers within the decile who did not get advice, and (b) *% of all non-advised* — the ratio of the number of non-advised borrowers whose propensity scores are in the given decile to the total number of non-advised borrowers in the sample.

Because of the high proportion of advised transaction in the overall data, the *% non-advised* metric is below 50% in all deciles. Nonetheless, ranking by propensity scores has considerable predictive power, with the share of non-advised consumers rising from 5% in the top decile of predicted propensity scores to 39% in the top decile.

The second metric focusing on *% of all non-advised* also shows that propensity scores have good discriminating power. More than 40% of non-advised borrowers are in the bottom two propensity score deciles, and less than 8% are allocated into the top two bands. This suggests that the propensity score model performs well in ranking borrowers' propensity to take advice. Note that a perfectly discriminating model is neither realistic nor desirable. As described above, there is a large circumstantial (random) component to borrowers' advice choice that is not related to outcomes in which we are interested. Moreover, propensity scores from a nearly perfectly

18 As discussed in e.g. Athey and Imbens (2017), there are recent developments in machine learning econometrics that offer 'doubly robust' methods for treatment effect estimation using propensity scores that have substantial advantages over using LASSO in the stage of propensity score estimation. Applying more advanced machine learning to the question we are investigating is a natural area for further work.

19 It would not be meaningful to test this after the MMR implementation, since $Pr(Advice) = 1$ for this period.

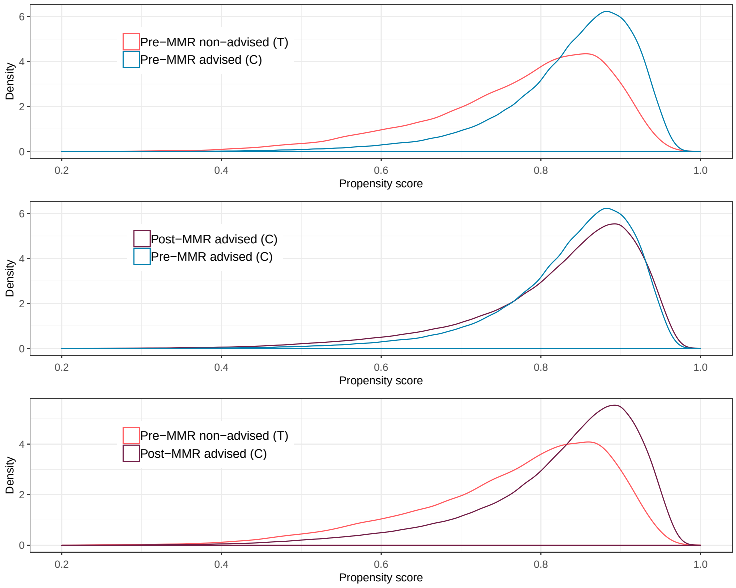
Table 4.3: Predictive power of the propensity score model

$\hat{p}(X_i)$ decile	% non-advised	% of all non-advised
1	0.389	0.238
2	0.271	0.166
3	0.206	0.126
4	0.173	0.106
5	0.149	0.091
6	0.124	0.076
7	0.108	0.068
8	0.091	0.058
9	0.074	0.045
10	0.050	0.031

discriminating model would have little overlap between advised and non-advised consumers and would thus violate the common support assumption.

In fact, to construct a plausible counterfactual (and estimate unbiased treatment effects), matching requires that the distribution of observables in the treatment group has support in common with the control group(s). The three panels in Figure 4.3 show the extent to which this holds for predicted propensity scores between the treatment group and the relevant control groups. The top panel compares the pre-MMR treatment (non-advised consumers) and pre-MMR control group (advised consumers); the middle panel compares the treatment group and the post-MMR control group; and the bottom panel compares the pre-MMR control group and the post-MMR control group. As the Figure shows, the domain of the propensity score metric for the treatment group is covered well by the respective control groups; furthermore, each control group is larger in size than the corresponding treatment group. The common support assumption therefore seems warranted.

Figure 4.3: Propensity score densities, by group



Note: The figure shows the distribution of propensity scores in the sub-groups of borrowers defined by pre-/post- intervention and whether they received advice. (T) denotes the treatment group. (C) denotes the control group.

4.5 Results

4.5.1 ATT estimates

Table 4.4 shows our estimates for the impact of advice for the previously non-advised group on key outcome variables: the type of product chosen and the near-term cost of borrowing. Each estimate (Average Treatment effect on the Treated, or ATT) is presented alongside the mean in the treatment group before the MMR, which serves as a baseline.[20]

Product characteristics One of the main effects of advice is to materially change the interest rate type of mortgage products being taken out. Consumers in the treatment group are more likely to choose fixed rate mortgages (+11 percentage points), effectively halving the share of variable rate mortgages relative to baseline. This effect is particularly pronounced for 2-year fixed rate products, where the estimated ATT is +16 percentage points, more than 50% of the baseline share of this product type. Longer term fixed-rate mortgages, however, became less attractive, with the likelihood of choosing a 5-year fixed product falling by 5 percentage points.

20 For more detail, including the standard error estimates, see Table C.2 in the appendix.

Table 4.4: Effects of the MMR advice requirement on previously non-advised

	ATT	Baseline
Cost of borrowing		
Initial monthly payment, £	-17.1	927.1
Deal period APR, bps	-3.8	382.5
5-year APR, bps	6.1	394.1
Pr(strongly dominated), pp	3.5	6.7
Other characteristics		
Pr(fixed rate mortgage), pp	10.8	74.8
Pr(2-year fixed rate mortgage), pp	15.8	29.3
Pr(5-year fixed rate mortgage), pp	-5.0	33.2
Mortgage term, years	0.63	23.6
Pr(use a broker), pp	63.0	2.0

Note: All presented treatment effects are significant at 1% confidence level, using bootstrap standard errors clustered by on year-quarter and postcode area.

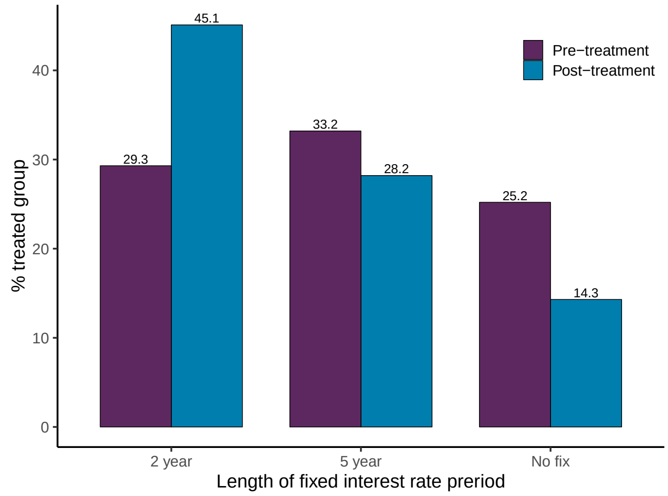
As illustrated in Figure 4.4, the overall effect of advice was to significantly increase the concentration of borrowers in 2-year fixed rate mortgages.

This concentration occurs in a period of historically low interest rates in the UK (followed, in fact, by another rate cut in August 2016 just after our post-MMR sample ends).[21] It is therefore not obvious that the 11 percentage point increase in 2-year fixed-rate mortgages at the expense of the (typically cheaper) fully variable rate mortgages was a result of advice correcting a pre-existing error of expectations about the future path of short-term interest rates as per Koijen, Van Hemert and Van Nieuwerburgh (2009). The hypothesis that advice helped the previously non-advised consumers make better trade-offs between variable and fixed rate mortgages to manage current mortgage costs (Campbell and Cocco, 2003; Campbell and Cocco, 2015) is also not very helpful in understanding the decline of variable rate mortgage use, although it could be relevant for the shift from 5-year to 2-year fixed mortgages.[22]

21 The Bank of England interest rate was constant during our sample period at 0.5% throughout our sample period.

22 A more formal investigation of the change in appetite for 2-year fixed mortgages associated with the advice requirement in the context of the existing literature on the topic of fixed vs variable (adjustable) rate mortgages (see, e.g. for a summary) is likely to be instructive in future research.

Figure 4.4: Effect of advice requirement on fixed rate periods



cost of borrowing Table 4.4 shows the impact of advice on near-term cost of borrowing, expressed as APR measures over either the deal period or on a 5-year basis. The effects appear small. For the deal period, we estimate a reduction in costs of around 4 basis points of APR. Under the assumption that borrowers remortgage after five years and stay on the reversion rate until then, our estimate of the impact of advice is an increase of 6 basis points. To illustrate: a 4 basis-point change in APR on a £152,000 loan (the median loan amount in our 2016 data) corresponds to a £5 difference in monthly payments. The direct estimate of the impact on monthly payments during the deal period is a sightly larger reduction of £14, although unlike the APR metrics this figure does not take into account up-front fees. To sum up, we find little to no evidence that receiving advice materially affected the borrowing costs for the treatment group on average.

We also find, however, that receiving advice increased the probability of consumers in the treatment group choosing a strongly dominated product by 3.5 percentage points—a considerable increase given the baseline of just 6.7%.

What is the reason for the potentially counter-intuitive increase in dominated choices and how can it be reconciled with no material effects of the advice requirement on borrowing costs? The change in the mix of products borrowers take out towards a concentration in 2-year fixed rate mortgages is an important factor. Typically 2-year fixed rate mortgages are the most standard product on the market, offered by virtually all lenders and in all LTV bands. Longer term fixed rate products, such as a 5-year fix, or even variable rate trackers are somewhat less widespread. Meanwhile, the dominance metric in Iscenko (2018) compares each borrower's choice with alternatives with exactly the same non-price characteristics (such as the length of the fixed rate period) as their chosen product. As long as there are imperfections in borrower's search, there is a higher likelihood of there being better alternatives on the market in product types, with a greater number of alternatives on the market.

In light of these facts, our findings are consistent with the situation where some consumers induced to take advice as a result of the MMR, who were previously choosing relatively good products (on cost), got advice to move to a generally cheaper product type of 2-year fixed rate loans but were not necessarily recommended a cost-effective product of that type (recall, for instance, that lenders can provide advice just on their own product range). As a result, the potential cost savings from considerably lower introductory rates (compared to 5-year fixes) are not fully realised, leading to our finding of small changes in borrowing costs. The rate of dominated choices, however, reflects the fact that there were better 2-year fixed rate products the borrower could have chosen.

Another likely interaction of the effects we estimate could be between the slight increase in mortgage term and decrease in monthly payments may be related. This may be driven by suitability concerns: a longer mortgage term mechanically reduces monthly payments, which may be deemed more suitable for the consumer by reducing their immediate debt servicing obligations. We do not produce estimates of the cost of borrowing over the term of the mortgage (this would require implausibly strong assumptions on repayment patterns and the availability of mortgage products in future), but it is likely that a lower rate of repayment will increase the total cost of credit.

intermediation Another substantial change is the increase in the use of brokers. The proportion of intermediated transactions among the treatment group, which had overwhelmingly preferred to go directly to the lender, grows to over 60%. Some of this change is likely due to capacity constraints that banks have in providing advice to all clients in branches. As housing transactions are usually time sensitive, this meant that more borrowers would have opted to seek out an available intermediary to get their mortgage if faced with a delay in getting an advice appointment with their preferred lender.

Naturally, this large change can raise a question about separating the effects of those borrowers starting to use brokers from the effects of getting advice per se, which can be a fruitful avenue for future research. At present, however, we are interested in documenting the effect of the advice requirement as a policy intervention. In this context, the effects discussed are relevant consequences of the policy whether or not they occurred through the mechanism of the policy de facto encouraging broker use or otherwise.

4.5.2 Robustness

bandwidth We test the extent to which our results might be affected by setting bandwidth using the Silverman (1986) rule of thumb. We re-estimate the model for a subset of outcomes with four alternative approaches. Three of these involve scaling the (constant) baseline bandwidth by a factor of 0.5, 1.5 and 2. In the fourth, we use the adaptive bandwidth approach from Blundell and Duncan (1998), which allows bandwidth at different points of the propensity score distribution to vary with the density of the data.

Table C.3 in the appendix compares the baseline estimates with those obtained with alternative bandwidth assumptions. Overall, the ATT estimates remain broadly the same under all of the four alternatives we test.

implementation period As discussed in 4.3.2, we exclude 12 months of data around the implementation date of the MMR (November 2013- October 2014) from estimation of treatment effects. A natural robustness check is to vary the excluded implementation period to ensure that our findings are not unduly affected by the somewhat arbitrary assumption about the transition period. Table C.4 in the Appendix C shows the results of alternative specifications of the implementation window (6 months and 0 months). The differences between ATTs obtained under different implementation periods are not economically significant.

4.6 Conclusion

We apply a difference-in-differences matching model to evaluate the effects of a real-world policy intervention that introduced mandatory advice for around a quarter of the UK mortgage consumers that previously opted to make their mortgage choices without it. We measure the ATT for the consumers who were 'brought into' advice in terms of their borrowing costs and the characteristics of products they chose.

One of the largest effects we observe is on the propensity to obtain the mortgage through a broker. Whereas before the intervention the treatment group strongly preferred to go directly to the lender, over a half of their post-MMR counterparts used brokers. Despite the rise in the use of brokers (who are meant to search the market for deals on behalf of the consumer), the effects of advice on borrowing costs were very muted. Even the direction of the effect depends on whether borrowers are assumed to remortgage immediately after their teaser period expires (a small reduction in costs from pre-MMR baseline) or stay on their current product for 5 years (a small increase in costs). Reliably decomposing the observed effects into the impact of using brokers and the impact of advice per se would require additional data to enable separate causal identification of the two mechanisms, but would be an interesting topic for further investigation.

The second large impact of the advice requirement was to substantially increase borrower concentration in 2-year fixed rate mortgages. Given the low and stable interest rate environment during the post-MMR period when this effect occurred, the existing literature on fixed-rate choices (Koijen, Van Hemert and Van Nieuwerburgh, 2009; Campbell and Cocco, 2003; Campbell and Cocco, 2015, , etc.) offers no obvious interpretation of this change in product mix (especially, the shift away the from variable rate mortgages) as advice correcting an previously existing error among non-advised borrowers. Further research that investigates this change with better data on changes (if any) in interest rate expectations during our sample period would likely prove insightful.

Another obvious extension to the existing work, which we are already pursuing, is exploring heterogeneity in the effects of the advice requirement. In recent years, there have been substantial advances in machine learning methods that allow for estimation of asymptotically consistent heterogeneous treatment effects and their variances, for instance using random forests (Wager and Athey, 2018; Athey, Tibshirani and Wager, 2019). These methods make it possible to formally explore whether the effects of being subjected to mandatory advice varied significantly with borrower demographics or circumstances.

Finally, as more time passes since the original policy intervention, it will become possible to explore some of the potential longer-term effects of advice on delinquency rates and other measures of loan performance.

5 Conclusion

This thesis is comprised of three papers that advance academic understanding of consumer behaviour in the mortgage market. As discussed in more detail in the individual chapters, I apply innovative methods to a unique combination of datasets to (a) detect and measure ineffective product choices in the mortgage market without requiring assumptions about preferences, (b) quantify the extent of limited attention in the market and the influence from links with lenders through existing product holdings and (c) explore whether and how introduction of mandatory mortgage advice in the UK in 2014 changed decisions.

This research has already started having real-world impact, as evidenced, for example by the extensive references to the findings in Chapters 2 and 4 in the 2018 Mortgages Market Study by the UK Financial Conduct Authority. [1] In addition to the specific findings about the mortgage market in the papers, I make a further contribution to the academic literature (and, potentially, regulatory work) with respect to methods for analysing consumer decisions. This is because the novel combinations of techniques I use in this thesis are relevant for other contexts with a proliferation of differentiated products, multi-dimensional pricing, and potential pre-existing links between consumers and some of the suppliers. Other relevant regulated markets where these methods are likely to be useful are, for instance, other financial products, utilities, and telecommunication services.

While the chapters in this thesis answer some pertinent questions in the literature, they also point the way for further research. For instance, there is a lot of potential in exploring the findings of Chapter 3 on the importance of existing relationships, to how mortgage brokers choose lenders on behalf of the borrowers. It is likely to be instructive to compare and contrast what familiarity and brand loyalty mean in the context of a broker and a consumer choosing the product directly. The findings about the impact of advice (Chapter 4) also raise important further questions about the heterogeneity of the potential impacts of mandatory advice and long-term impacts that receiving advice might have on loan performance.

1 See, for example Chapter 5 in the FCA's Mortgages Market Study Interim Report.

Bibliography

Abadie, A. and G. W. Imbens (2008). 'On the failure of the bootstrap for matching estimators'. *Econometrica* **76**, 1537–1557.

Abaluck, J. and A. Adams (2017). *What do consumers consider before they choose? Identification from asymmetric demand responses*. Tech. rep. National Bureau of Economic Research.

Abaluck, J. and J. Gruber (2011). 'Choice inconsistencies among the elderly: evidence from plan choice in the Medicare Part D program'. *American Economic Review* **101**, 1180–1210.

Adams, P. D., S. Hunt, C. Palmer and R. Zaliauskas (2019). *Testing the Effectiveness of Consumer Financial Disclosure: Experimental Evidence from Savings Accounts*. Tech. rep. NBER Working Paper No. 25718.

Agarwal, S., J. C. Driscoll, X. Gabaix and D. Laibson (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., J. C. Driscoll and D. I. Laibson (2013). 'Optimal Mortgage Refinancing: A Closed-Form Solution'. *Journal of Money, Credit and Banking* **45**, 591–622.

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

Alexandrov, A. and S. Koulayev (2018). 'No Shopping in the U.S. Mortgage Market: Direct and Strategic Effects of Providing Information'. *Consumer Financial Protection Bureau Office of Research Working Paper* **No. 2017-01**.

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

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

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

Allen, J., R. Clark and J.-F. Houde (2019). 'Search Frictions and Market Power in Negotiated-Price Markets'. *Journal of Political Economy* **127**, 1550–1598.

Andersen, S., J. Y. Campbell, K. M. Nielsen and T. Ramadorai (2018). 'Sources of inaction in household finance: Evidence from the Danish mortgage market'. *Available at SSRN 2463575*.

Andersen, S., J. Y. Campbell, K. M. Nielsen and T. Ramadorai (2019). 'Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market'.

Armstrong, M. (2015). 'Search and ripoff externalities'. *Review of Industrial Organization* **47**, 273–302.

Athey, S. and G. Imbens (2019). 'Machine Learning Methods Economists Should Know About'. *arXiv preprint arXiv:1903.10075*.

Athey, S. and G. W. Imbens (2017). 'The state of applied econometrics: Causality and policy evaluation'. *Journal of Economic Perspectives* **31**, 3–32.

Athey, S., J. Tibshirani and S. Wager (2019). 'Generalized random forests'. *The Annals of Statistics* **47**, 1148–1178.

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 T. Majer (2018). 'Six of One...? Choice of intermediary in the UK mortgage market'. *Financial Conduct Authority Occasional Paper* **35**.

Ben-Akiva, M. and M. Bierlaire (1999). 'Discrete choice methods and their applications to short term travel decisions'. *Handbook of transportation science*. Springer, 5–33.

Ben-Akiva, M., D. McFadden, M. Abe, U. Böckenholt, D. Bolduc, D. Gopinath, T. Morikawa, V. Ramaswamy, V. Rao, D. Revelt et al. (1997). 'Modeling methods for discrete choice analysis'. *Marketing Letters* **8**, 273–286.

Benetton, M. (2019). 'Leverage Regulation and Market Structure: An Empirical Model of the UK Mortgage Market'.

Benetton, M., P. Bracke, J. F. Cocco and N. Garbarino (2019). 'Housing consumption and investment: evidence from shared equity mortgages'.

Bergstresser, D., J. M. Chalmers and P. Tufano (2008). 'Assessing the costs and benefits of brokers in the mutual fund industry'. *The Review of Financial Studies* **22**, 4129–4156.

Berry, S. and P. Haile (2016). 'Identification in differentiated products markets'. *Annual review of Economics* **8**, 27–52.

Berry, S., J. Levinsohn and A. Pakes (1995). 'Automobile prices in market equilibrium'. *Econometrica: Journal of the Econometric Society*, 841–890.

Best, M. C., J. Cloyne, E. Ilzetzki and H. J. Kleven (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., G. Loewenstein and J. Sydnor (2017). 'Choose to Lose: Health Plan Choices from a Menu with Dominated Option'. *The Quarterly Journal of Economics* **132**, 1319–1372.

Bhattacharya, U., A. Hackethal, S. Kaesler, B. Loos and S. Meyer (2012). 'Is unbiased financial advice to retail investors sufficient? Answers from a large field study'. *The Review of Financial Studies* **25**, 975–1032.

Bhutta, N., A. Fuster and A. Hizmo (2019). 'Paying too much? Price dispersion in the US mortgage market'.

Blundell, R. and M. Costa Dias (2000). 'Evaluation methods for non-experimental data'. *Fiscal Studies* **21**, 427–468.

Blundell, R. and M. Costa Dias (2009). 'Alternative approaches to evaluation in empirical microeconomics'. *Journal of Human Resources* **44**, 565–640.

Blundell, R., M. Costa Dias, C. Meghir and J. Shaw (2016). 'Female Labor Supply, Human Capital, and Welfare Reform'. *Econometrica* **84**, 1705–1753.

Blundell, R. and A. Duncan (1998). 'Kernel regression in empirical microeconomics'. *Journal of Human Resources*, 62–87.

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

Bronnenberg, B. J., J.-P. H. Dubé and M. Gentzkow (2012). 'The evolution of brand preferences: Evidence from consumer migration'. *American Economic Review* **102**, 2472–2508.

Bronnenberg, B. J., J.-P. Dubé and S. Moorthy (2019). 'The economics of brands and branding'. *Handbook of the Economics of Marketing, Volume 1*. Ed. by J.-P. Dube and P. E. Rossi. Vol. 1. Handbook of the Economics of Marketing.

Brown, G. H. (1953). 'Brand loyalty – fact of fiction'. *Trademark Rep.* **43**, 251.

Calvet, L. E., J. Y. Campbell and P. Sodini (2009). 'Measuring the Financial Sophistication of Households'. *American Economic Review* **99**, 393–98.

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

Campbell, J. Y. (2006). 'Household finance'. *The journal of finance* **61**, 1553–1604.

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

Campbell, J. Y. and J. F. Cocco (2015). 'A model of mortgage default'. *The Journal of Finance* **70**, 1495–1554.

Chalmers, J. and J. Reuter (2012). 'What is the impact of financial advisors on retirement portfolio choices and outcomes?' *Available at SSRN 2078536*.

Cloyne, J., K. Huber, E. Ilzetzki and H. Kleven (2019). 'The effect of house prices on household borrowing: a new approach'. *American Economic Review* **109**, 2104–36.

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

Coval, J. D. and T. J. Moskowitz (1999). 'Home bias at home: Local equity preference in domestic portfolios'. *The Journal of Finance* **54**, 2045–2073.

Crawford, G., R. Griffith and A. Iaria (2019). *Preference estimation with unobserved choice set heterogeneity using sufficient sets*. Tech. rep. Discussion Paper 11675, Center for Economic Policy Research.

Dubé, J.-P., G. J. Hitsch and P. E. Rossi (2010). 'State dependence and alternative explanations for consumer inertia'. *The RAND Journal of Economics* **41**, 417–445.

Egan, M. (2019). 'Brokers vs. retail investors: Conflicting interests and dominated products'. *Journal of Finance* **74**, 1217–1260.

Eliaz, K. and R. Spiegler (2006). 'Contracting with diversely naive agents'. *The Review of Economic Studies* **73**, 689–714.

Elton, E. J., M. J. Gruber and J. A. Busse (2004). 'Are Investors Rational? Choices among Index Funds'. *The Journal of Finance* **59**, 261–288.

ESRO (2015). *Understanding consumer expectations of the mortgage sales process: A report prepared for the Financial Conduct Authority*. Tech. rep.

FCA (2017). *Financial Lives Survey 2017*. Tech. rep. Financial Conduct Authority.

FCA (2018a). *Mortgages Market Study: Initial findings*. Tech. rep. Financial Conduct Authority.

FCA (2018b). *Strategic Review of Retail Banking Business Models: Final report*. Tech. rep. Financial Conduct Authority.

FCA (2018c). *Strategic Review of Retail Banking Business Models: Progress report*. Tech. rep. Financial Conduct Authority.

Finke, M. S., J. S. Howe and S. J. Huston (2017). 'Old Age and the Decline in Financial Literacy'. *Management Science* **63**, 213–230.

Foà, G., L. Gambacorta, L. Guiso and P. E. Mistrulli (2019). 'The supply side of household finance'. *The Review of Financial Studies* **32**, 3762–3798.

Foerster, S., J. T. Linnainmaa, B. T. Melzer and A. Previtero (2017). 'Retail financial advice: does one size fit all?' *The Journal of Finance* **72**, 1441–1482.

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

Friedman, J., T. Hastie and R. Tibshirani (2009). 'glmnet: Lasso and elastic-net regularized generalized linear models'. *R package version* **1**.

FSA (2009). *Discussion Paper 09/03: Mortgage Market Review*. Tech. rep. Financial Services Authority.

FSA (2010). *Consultation Paper 10/28: Mortgage Market Review - Distribution and Disclosure*. Tech. rep. Financial Services Authority.

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

Gavazza, A. (2016). 'An empirical equilibrium model of a decentralized asset market'. *Econometrica* **84**, 1755–1798.

Goeree, M. S. (2008). 'Limited information and advertising in the US personal computer industry'. *Econometrica* **76**, 1017–1074.

Greene, W. H. and D. A. Hensher (2003). 'A latent class model for discrete choice analysis: contrasts with mixed logit'. *Transportation Research Part B: Methodological* **37**, 681–698.

Grubb, M. D. (2015). 'Failing to choose the best price: Theory, evidence, and policy'. *Review of Industrial Organization* **47**, 303–340.

Guiso, L., A. Pozzi, A. Tsoy, L. Gambacorta and P. E. Mistrulli (2018). 'The cost of steering in financial markets: Evidence from the mortgage market'. *Available at SSRN 2951042*.

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

Hackethal, A., M. Haliassos and T. Jappelli (2012). 'Financial advisors: A case of babysitters?' *Journal of Banking & Finance* **36**, 509–524.

Halek, M. and J. G. Eisenhauer (2001). 'Demography of risk aversion'. *Journal of Risk and Insurance*, 1–24.

Handel, B. R. (2013). 'Adverse selection and inertia in health insurance markets: When nudging hurts'. *American Economic Review* **103**, 2643–82.

Heiss, F., D. McFadden, J. Winter, A. Wuppermann and B. Zhou (2016). *Inattention and Switching Costs as Sources of Inertia in Medicare Part D*. Tech. rep. NBER Working Paper No. 22765.

Ho, K., J. Hogan and F. Scott Morton (2017). 'The impact of consumer inattention on insurer pricing in the Medicare Part D program'. *The RAND Journal of Economics* **48**, 877–905.

Hoechle, D., S. Ruenzi, N. Schaub and M. Schmid (2017). 'The impact of financial advice on trade performance and behavioral biases'. *Review of Finance* **21**, 871–910.

Honka, E., A. Hortaçsu and M. A. Vitorino (2017). 'Advertising, consumer awareness, and choice: Evidence from the us banking industry'. *The RAND Journal of Economics* **48**, 611–646.

Honka, E., A. Hortaçsu and M. Wildenbeest (2019). 'Empirical search and consideration sets'. *Handbook of the Economics of Marketing, Volume 1*. Ed. by J.-P. Dube and P. E. Rossi. Vol. 1. Handbook of the Economics of Marketing, 193–257.

Hortaçsu, A., S. A. Madanizadeh and S. L. Puller (2017). 'Power to choose? An analysis of consumer inertia in the residential electricity market'. *American Economic Journal: Economic Policy* **9**, 192–226.

Huberman, G. (2001). 'Familiarity Breeds Investment'. *The Review of Financial Studies* **14**, 659–680.

Imbens, G. W. (2015). 'Matching methods in practice: Three examples'. *Journal of Human Resources* **50**, 373–419.

Iscenko, Z. (2018). 'Choices of dominated mortgage products by UK consumers'. *Financial Conduct Authority Occasional Paper* **33**.

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

Jiang, W., A. A. Nelson and E. Vytlacil (2014). 'Liar's loan? Effects of origination channel and information falsification on mortgage delinquency'. *Review of Economics and Statistics* **96**, 1–18.

Jones, E., T. Oliphant, P. Peterson et al. (2001–). *SciPy: Open source scientific tools for Python*.

Karabulut, Y. (2013). 'Financial advice: an improvement for worse?' *Available at SSRN 1710634*.

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

Koijen, R. S., O. Van Hemert and S. Van Nieuwerburgh (2009). 'Mortgage timing'. *Journal of Financial Economics* **93**, 292–324.

Lachance, M.-E. and N. Tang (2012). 'Financial advice and trust.' *Financial Services Review* **21**.

Linnainmaa, J. T., B. Melzer and A. Previtero (2018). 'The misguided beliefs of financial advisors'. *Kelley School of Business Research Paper*.

Liu, L. (2019). 'Non-salient fees in the mortgage market'. *Available at SSRN 3437374*.

Lusardi, A. and O. S. Mitchell (2014). 'The Economic Importance of Financial Literacy: Theory and Evidence'. *Journal of Economic Literature* **52**, 5–44.

Lusardi, A., O. S. Mitchell and V. Curto (2014). 'Financial literacy and financial sophistication in the older population'. *Journal of pension economics & finance* **13**, 347–366.

Marmorstein, H., D. Grewal and R. P. H. Fishe (1992). 'The Value of Time Spent in Price-Comparison Shopping: Survey and Experimental Evidence'. *Journal of Consumer Research* **19**, 52–61.

Mullainathan, S., M. Noeth and A. Schoar (2012). *The market for financial advice: An audit study*. Tech. rep. National Bureau of Economic Research.

Murray, K. B. and G. Häubl (2007). 'Explaining cognitive lock-in: The role of skill-based habits of use in consumer choice'. *Journal of Consumer Research* **34**, 77–88.

Mysliwski, M. and M. Rostom (2018). 'Value of Information and the Impact of Mortgage Intermediaries on Lender Competition and Households' Financial Positions'.

Nelson, S. T. (2018). 'Private Information and Price Regulation in the US Credit Card Market'.

Nieboer, J., P. Dolan and I. Vlaev (2017). 'Financial Advisory Services'. *Financial Behavior: Players, Services, Products, and Markets*. Ed. by K. Baker, G. Filbeck and V. Ricciardi, 285–301.

Oliver, R. L. (1999). 'Whence consumer loyalty?' *Journal of marketing* **63**, 33–44.

Palmatier, R. W., R. P. Dant, D. Grewal and K. R. Evans (2006). 'Factors influencing the effectiveness of relationship marketing: a meta-analysis'. *Journal of marketing* **70**, 136–153.

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

Robb, C. A., P. Babiarz and A. Woodyard (2012). 'The demand for financial professionals' advice: The role of financial knowledge, satisfaction, and confidence.' *Financial Services Review* **21**.

Robles-Garcia, C. (2019). 'Competition and Incentives in Mortgage Markets: The Role of Brokers'.

Salz, T. (2017). 'Intermediation and competition in search markets: An empirical case study'.

Scholnick, B., N. Massoud and A. Saunders (2013). 'The impact of wealth on financial mistakes: Evidence from credit card non-payment'. *Journal of Financial Stability* **9**, 26–37.

Shen, N. (2016). 'Consumer rationality/irrationality and financial literacy in the credit card market: Implications from an integrative review'. *Financial Literacy and the Limits of Financial Decision-Making*. Springer, 155–176.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Chapman & Hall/CRC, 48.

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

Swait, J. and M. Ben-Akiva (1987). 'Incorporating random constraints in discrete models of choice set generation'. *Transportation Research Part B: Methodological* **21**, 91–102.

Terui, N., M. Ban and G. M. Allenby (2011). 'The effect of media advertising on brand consideration and choice'. *Marketing science* **30**, 74–91.

Tibshirani, R. (1996). 'Regression shrinkage and selection via the lasso'. *Journal of the Royal Statistical Society: Series B (Methodological)* **58**, 267–288.

Tucker, W. T. (1964). 'The development of brand loyalty'. *Journal of Marketing research* **1**, 32–35.

Tyers, R., M. Sweeney and B. Moon (2019). 'Harnessing behavioural insights to encourage consumer engagement in the British energy market: Results from a field trial'. *Journal of Behavioral and Experimental Economics* **80**, 162–176.

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

Van Nierop, E., B. Bronnenberg, R. Paap, M. Wedel and P. H. Franses (2010). ‘Retrieving unobserved consideration sets from household panel data’. *Journal of Marketing Research* **47**, 63–74.

Wager, S. and S. Athey (2018). ‘Estimation and inference of heterogeneous treatment effects using random forests’. *Journal of the American Statistical Association* **113**, 1228–1242.

Wales, D. J. and J. P. Doye (1997). ‘Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms’. *The Journal of Physical Chemistry A* **101**, 5111–5116.

Woodward, S. E. and R. E. Hall (2010). ‘Consumer confusion in the mortgage market: Evidence of less than a perfectly transparent and competitive market’. *American Economic Review* **100**, 511–15.

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

A | Appendix for Chapter 2

A.1 Additional robustness checks

a.1.1 Cost specification

I undertake two 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 (reported in Table 2.2).

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.[1] 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.3 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, 2018a), 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.1 for the full distribution). However, this scenario again involves implausible borrower behaviour as it implies that nearly all borrowers stay on their

1 In cases where remortgaging after the given number of years (e.g. 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.

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 2.2.1). 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.

A.1.2 Sample representativeness

As discussed in section 2.3.2, 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.4 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.

A.1.3 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 is 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 2.4.1) for each borrower and identify all alternatives, whether chosen or not, 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.[2] 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.2 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, for a 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 than chance would suggest, regardless of their demographic profile and use of specialist support (e.g. an independent intermediary).

A.2 Additional figures and tables

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

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

	Sample			Population[a]		
Variable	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 2.3.2, e.g. LTV$\geq$20%. (b) Income is post-tax household earnings added across all borrowers named on the loan.

Figure A.1: Distribution of excess costs incurred due to dominated mortgage choice

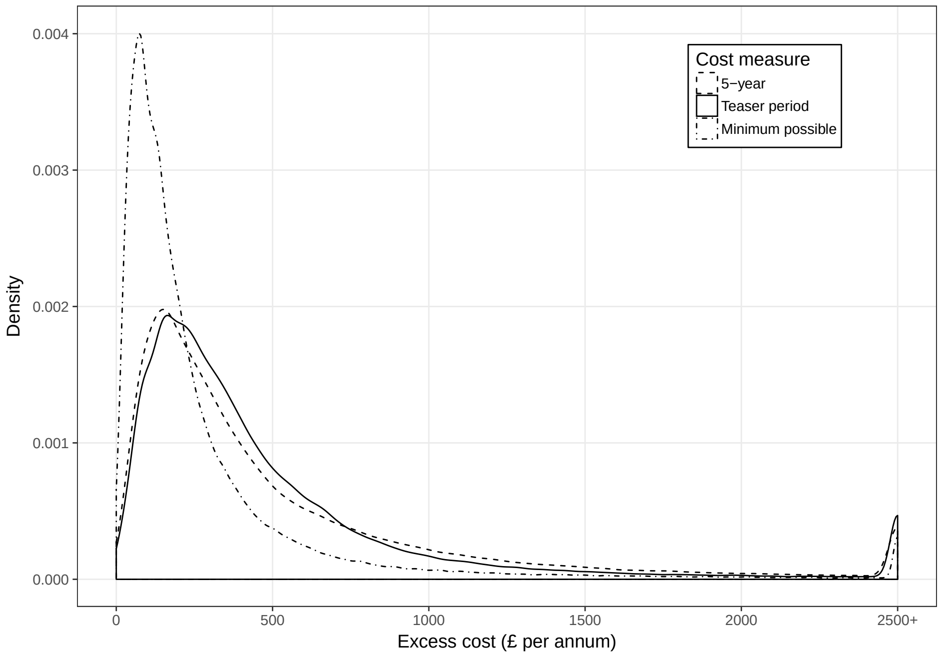


Table A.2: Regression results for main borrower characteristics

	=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×log(Income (£1000))		-1.942^{***}	-1.934^{***}	14.160^{***}	12.991^{***}
=1 if joint×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 2.3.

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

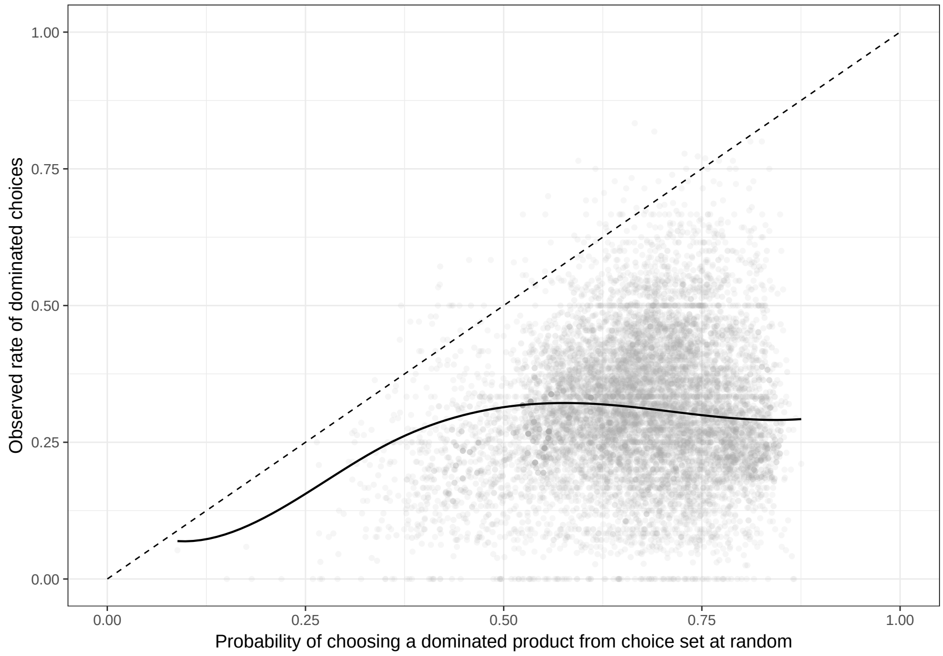
Table A.3: Distributions of annual excess cost (£pa) under different time horizons

			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

Table A.4: 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

Figure A.2: 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 2.3) 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.

Table A.5: 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×*log*(Income (£1000))	−2.058***	0.207	0.292	−3.777***	−2.670***
=1 if joint×*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.

B | Appendix for Chapter 3

B.1 Extra materials

Figure B.1: Distribution of predicted type by demographic characteristics

(a) Income vs loan amount (b) Income vs education

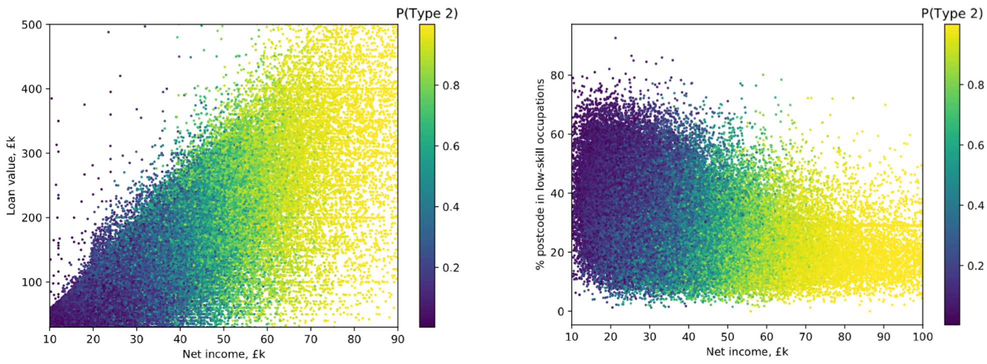


Table B.1: Estimated parameters from the baseline model: Full

	Type 1		Type 2	
	parameter	p value	parameter	p value
Preference parameters (β)				
Initial interest rate (%)	-0.7353	0.0000	-0.9739	0.0000
Reversion interest rate (%)	-1.5250	0.0000	-1.3602	0.0000
Initial rate X teaser period length	-0.1505	0.0000	-0.1615	0.0000
Reversion rate X teaser period length	0.3142	0.0000	0.2146	0.0000
Upfront fee (£)	-0.0012	0.0000	0.0005	0.0000
Cashback amount (£)	0.0008	0.0000	0.0012	0.0000
=1 if free property valuation	-0.2173	0.0000	0.0929	0.0104
=1 if 2-year fixed rate	0.9859	0.0000	1.5438	0.0000
=1 if 3/4-year fixed rate	0.5666	0.0000	1.0443	0.0000
=1 if 5-year fixed rate	-0.7440	0.0000	1.8045	0.0000
=1 if 10-year fixed rate	-3.8914	0.0000	1.3662	0.0006
=1 if has payment holiday option	0.4340	0.7057	-2.1733	0.0586
=1 if has underpayment option	0.9902	0.3892	1.1925	0.2998
Early repayment penalty (% of loan)	0.1918	0.0000	-0.3439	0.0000
Distance to closest branch (mi)	-0.0924	0.0000	-0.0904	0.0000
Headroom to max LTV (pp)	1.0062	0.0000	0.5776	0.0000
Headroom to min credit score	0.0098	0.0000	0.0199	0.0000
Branches within 5 mi (N)	0.0326	0.0000	0.0285	0.0909
Regional advertising spend (£per cap pcm)	-0.0211	0.0418	0.3810	0.0000
=1 if has PCA with lender	1.5951	0.0000	2.1499	0.0000
Length of lender relationship (years)	0.0252	0.0000	0.0983	0.0000
Lender FEs	Yes		Yes	
Attention parameters (γ)				
Intercept	-2.0467	0.0000	0.7997	0.0000
Lender size: lending volume in 2014	0.0564	0.0000	0.0093	0.0000
Branches within 5 mi (N)	0.1289	0.0000	0.0245	0.0002
Regional advertising spend (£per cap pcm)	0.1758	0.0000	-0.2324	0.0000
=1 if has PCA with lender	2.2008	0.0033	1.0647	0.0000
Length of lender relationship (years)	4.6188	0.0000	0.0083	0.0444
Demographic parameters for P(Type 2) (δ)				
Intercept	-3.1991	0.0000		
Credit score	0.0211	0.0000		
Age (years)	-0.0353	0.0000		
Net income (£1000)	0.1180	0.0000		
Loan-to-value (LTV, %)	-0.0093	0.0000		
Postcode: % in low-skill occupations	-0.0195	0.0000		
=1 if first-time buyer	0.7254	0.0000		
=1 if joint mortgage	-1.1090	0.0000		
=1 if self-employed	-0.6639	0.0000		
N (individuals)	64,716			
N (observations)	2,301,341			
Fitted log-likelihood	-149,268			
McFadden's adjusted R^2	0.3265			

Table B.2: Estimated marginal costs and demand elasticities

	Demand elasticity[a]		Marginal cost[b]	
	Mean	σ	Mean	σ
All	-3.641	1.730	1.327	1.165
Lender size:[c]				
Larger 6	-3.625	1.739	1.291	1.195
Smaller 6	-3.674	1.712	1.400	1.101
Max LTV band:[d]				
(0, 50]	-1.952	1.863	1.144	1.046
(50, 70]	-2.533	1.849	1.219	1.156
(70, 85]	3.259	1.800	1.358	1.240
(85, 100]	-4.630	1.843	1.890	1.124
Fixed rate:				
Yes	-3.808	1.741	1.475	1.158
No	-2.240	0.693	0.715	1.018

Mean and standard deviation across products that satisfy the given condition. (a) Elasticity for each product j is an unweighted average of own-price elasticities for j across all borrowers who have it in their choice set. Borrower-product level elasticities use the individual marginal effects of the interest rate on P(Choice) described in footnote 24. (b) Marginal cost is obtained using the approach described in section 3.4.2. (c) Lender size ranking is based on total mortgage lending volume. (d) Bands are based on the maximum loan-to-value ration accepted for each product.

Table B.3: Alternative specifications in a single type model

	Full sample			FTB only[a]
	(1) Baseline	(2)	(3)	(4)
Preference parameters (β)				
Initial interest rate (%)	−0.8383***	−0.841***	−0.8332***	−0.8423***
Reversion interest rate, (%)	−1.5328***	−1.5335***	−1.5285***	−2.1994***
Upfront fee (£)	−0.0004***	−0.0004***	−0.0004***	−0.0006***
Cashback amount (£)	0.0009***	0.0009***	0.0009***	0.0009***
=1 if free property valuation	−0.3118***	−0.3119***	−0.3103***	−0.5089***
=1 if 2-year fixed rate	1.1115***	1.1113***	1.1100***	0.9363***
=1 if 3-year fixed rate	0.6187***	0.6183***	0.6139***	−0.3113***
=1 if 5-year fixed rate	0.1188	0.1184	0.1169	−2.296***
=1 if 10-year fixed rate	−2.0717***	−2.0737***	−2.0737***	−8.3731***
Initial rate X teaser period length	−0.1445***	−0.1444***	−0.1444***	−0.1233***
Reversion rate X teaser period length	0.2737***	0.2738***	0.2733***	0.4248***
=1 if has payment holiday option	−0.3983	−0.3086	−0.3015	−0.2014
=1 if has underpayment option	0.5152	0.4260	0.4013	1.1420
Early repayment penalty (% of loan)	−0.0655***	−0.0654***	−0.0651***	0.1315***
Distance to closest branch (mi)	−0.1048***	−0.104***	−0.1042***	−0.1422***
Headroom to max LTV (pp)	0.7204***	0.7155***	0.7352***	0.4609***
Headroom to min credit score	0.0146***	0.0146***	0.0147***	0.0179***
Branches within 5 mi (N)	0.0458***	0.0463***	0.0415***	0.0803***
Regional advertising spend (£per cap pcm)	0.1324***	0.1310***	0.1281***	0.0790***
Length of lender relationship (years)	0.0443***	0.0334***	0.0504***	0.0195***
=1 if has PCA with lender[b]	1.6661***	1.5865***		2.225***
=1 if has other products with lender[c]		0.6785***		
Number of PCAs with lender[b]			0.6752***	
Number of other products with lender[c]			0.3705***	
Attention parameters (γ)				
intercept	0.6109***	0.6169***	0.6176***	1.0031***
Lender size: lending volume in 2014	0.0302***	0.0301***	0.0293***	0.0287***
Branches within 5 mi (N)	0.0451***	0.0451***	0.0512***	0.0175*
Regional advertising spend (£per cap pcm)	−0.1213***	−0.1201***	−0.1192***	−0.0743***
Length of lender relationship (years)	0.0406***	0.0336***	0.045***	0.0196***
=1 if has PCA with lender[b]	1.4670***	1.4390***		1.3036***
=1 if has other products with lender[c]		0.2897***		
Number of PCAs with lender[b]			1.0688***	
Number of other products with lender[c]			0.3042***	
Lender FEs	Yes	Yes	Yes	Yes
N (individuals)		64,716		28,260
McFadden's adjusted R^2	0.2988	0.3004	0.2975	0.2899

(a) Model estimated in the sub-sample of first-time home buyers only, with no prior home or mortgage ownership. (b) Measure of the relationship with the lender through personal current accounts (PCA) opened over 6 months before to the mortgage application date. (c) 'Other products' include credit cards, personal loans, and mortgages and exclude lines of credit directly linked to current accounts (e.g. overdraft facility).

*p<0.05; **p<0.01; ***p<0.001.

C | Appendix for Chapter 4

Table C.1: Covariates for the propensity score, Pr(Advice)

Variable group	Variable description
Borrower	First time buyer dummy
	Joint mortgage dummy
	Main borrower age
	Main borrower gender (derived from title)
	Main borrower employment status
	Main borrower postgraduate qualifications (derived from title)
	Combined gross income (percentile in overall distribution in year-quarter of transaction)
	Credit score at time of application*
	Impaired credit flag
Borrower financial capability	Combined number of financial product types held in past 6 years*
	Measures of how concentrated the borrower's product holdings are with a small number of lenders (HHI all credit products, HHI current accounts)*, ****
	Main borrower number of credit card cash advances in last 12 months*
	Main borrower number of credit cards with minimum payment flag set*
Loan	Loan-To-Income (LTI) ratio (percentile)
	Loan-To-Value (LTV) ratio (percentile)
Property	Property value (percentile)
	New build
	Problem property proxy
Property area	Postcode area % unskilled/partly skilled workers**
	Postcode area % with university degree or higher**
	Postcode area % unemployment**
	Postcode area % of houses with zero county court judgements**
	Postcode area average number of defaulted credit accounts**
	Property region fixed effects
Choice complexity	Number of mortgage products for which the borrower is eligible***

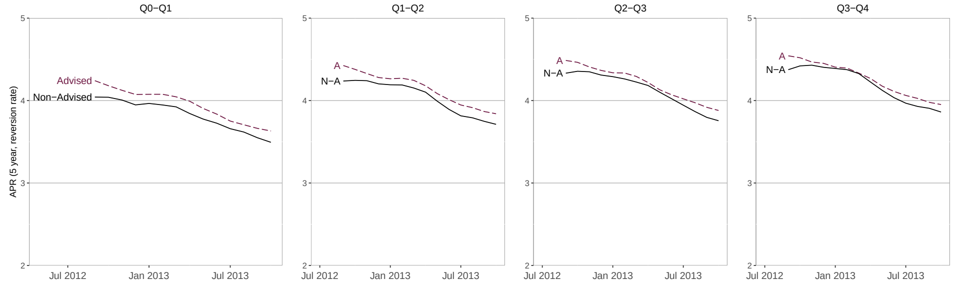
Notes: Data is taken from PSD unless indicated otherwise: * indicates variable taken from credit file; ** = variable taken from UK Census; *** = variable taken from Moneyfacts product data. **** HHI stands for Herfindahl-Hirschman index, a common measure of concentration.

C.1 Parallel trends

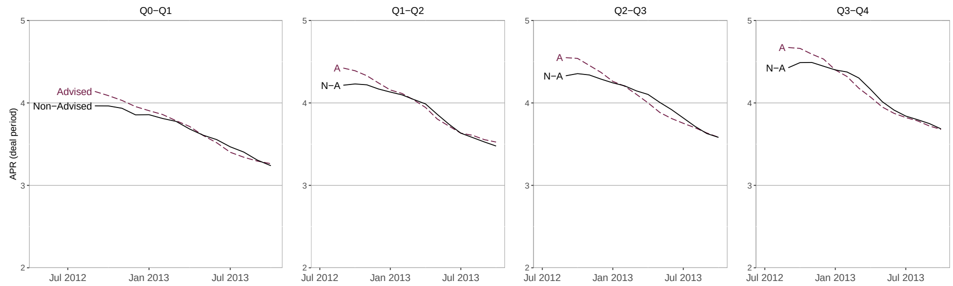
Figure C.1 and C.3 show 3-month moving averages, separate for advised (pre-MMR control group) and non-advised consumers (pre-MMR treatment group), grouped by the (population) quartiles of propensity score. The overall picture is that trends over time in the two groups are consistent, although the trends for intermediation are hard to evaluate for the non-advised sample.

Figure C.1: Parallel trends in cost outcomes, by quartile of propensity score

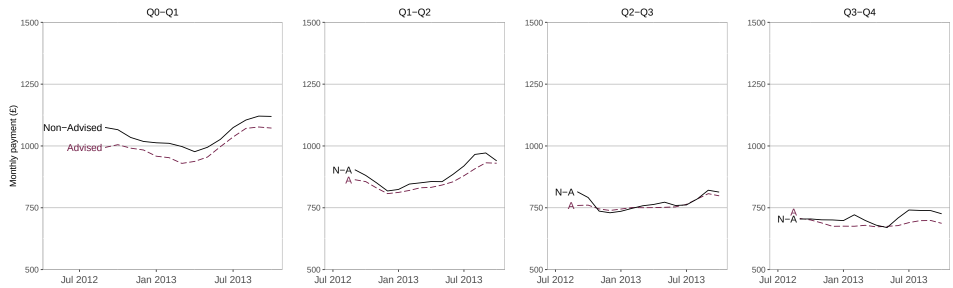
(a) 5-year APR, %, if stay on reversion rate after the teaser period



(b) Deal period APR,%



(c) Initial monthly payment, £



(d) Proportion of strongly dominated choces, £

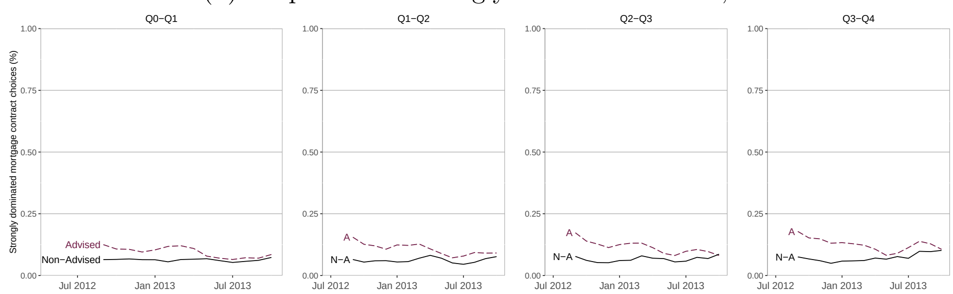
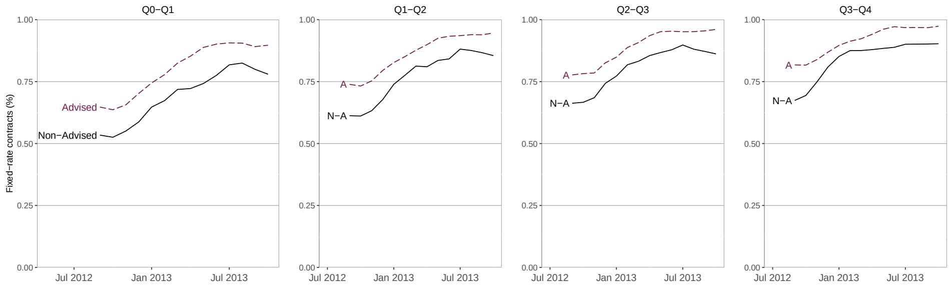
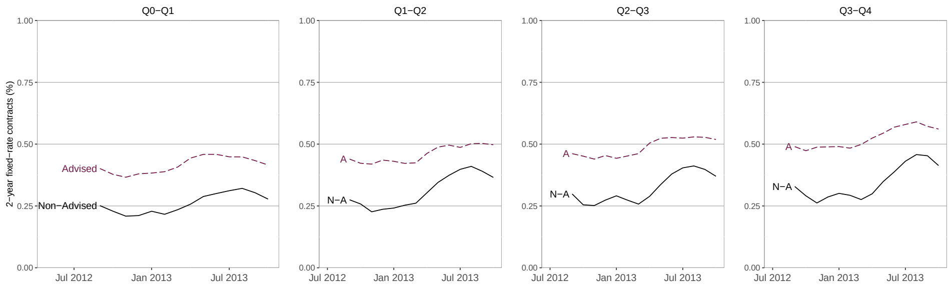


Figure C.3: Parallel trends in product outcomes, by quartile of propensity score

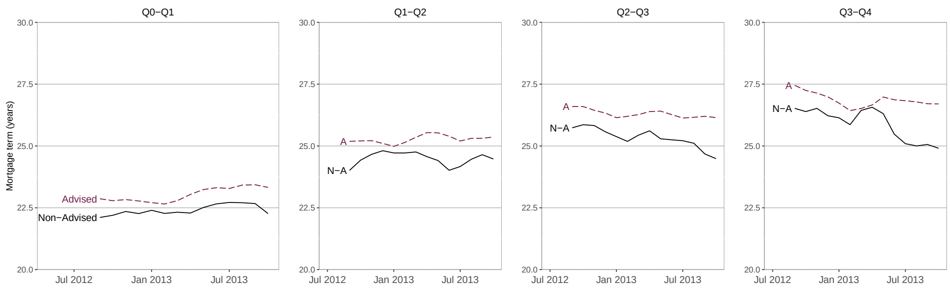
(a) Proportion of fixed rate-mortgages



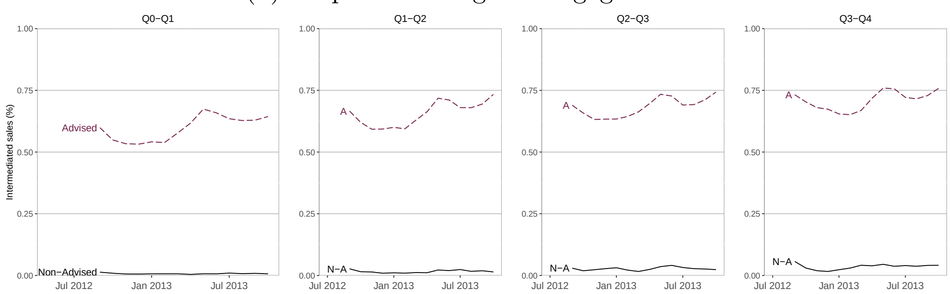
(b) Proportion of 2-year fixed rate mortgages



(c) Mortgage term, years



(d) Proportion using a mortgage broker



C.2 Findings and robustness checks

Table C.2: Effects of the MMR advice requirement on previously non-advised, ATT

	Baseline	ATT	SE
Cost of borrowing			
Initial monthly payment, £	926.178	-17.107	2.573
Deal period APR, bps	382.48	-3.826	0.534
5-year APR, bps	394.12	6.093	0.394
Pr(strongly dominated), pp	6.744	3.478	0.169
Other characteristics			
Pr(fixed rate mortgage), pp	74.779	10.905	0.349
Pr(2-year fixed rate mortgage), pp	29.318	15.820	0.368
Pr(5-year fixed rate mortgage), pp	33.281	-5.003	0.500
Mortgage term, years	23.612	0.627	0.037
Pr(use a broker), pp	2.044	63.025	0.006

Note: This table shows the estimated ATT of mandatory mortgage advice for borrowers who did not receive advice before the MMR. Baseline is the mean outcome of interest within the pre-MMR non-advised group. Standard errors are calculated using 100 bootstrap replications clustered on on year-quarter and postcode area.

Table C.3: Sensitivity of the ATT to different bandwidths

	(1) Baseline (h)	(2) 0.5*h	(3) 1.5*h	(4) 2*h	(5) Adaptive
Pr(fixed rate mortgage), pp	10.91	11.06	11.01	11.10	11.11
Pr(2-year fixed rate mortgage), pp	15.80	15.94	15.81	15.88	15.97
Deal period APR, pp	-0.04	-0.04	-0.04	-0.04	-0.04
5-year APR, pp	0.06	0.06	0.06	0.06	0.06

Notes: Baseline is the Silverman (1986) rule of thumb bandwidth used for the headline results of the paper. Columns (2) to (4) show estimates from matching estimation where the baseline bandwidths are scaled by a factor of 0.5, 1.5 and 2, respectively. Adaptive bandwidth uses an approach from Blundell and Duncan (1998), where the bandwidth is allowed to vary with the density of the propensity score at each point.

Table C.4: Sensitivity of the ATT to alternative lengths of the MMR implementation period (excluded from estimation)

Outcome	Baseline model 12-month	Alt. implementation 6-month	0-month
Cost of borrowing			
Initial monthly payment, £	-17.11	-15.94	-14.85
Deal period APR, pp	-0.04	-0.06	-0.06
5-year APR, pp	0.06	0.04	0.04
Pr(strongly dominated), pp	3.45	2.73	2.37
Other characteristics			
Pr(fixed rate mortgage), pp	10.91	9.67	9.32
Pr(2-year fixed rate mortgage), pp	15.80	14.58	14.29
Pr(5-year fixed rate mortgage), pp	-5.00	-3.72	-4.21
Mortgage term, years	0.63	0.58	0.59
Pr(use a broker), pp	63.03	58.30	58.18