

MS16/2.2: Annex 4

Market Study

---

# Mortgages Market Study

Interim Report: Annex 4 - Methodology of the  
dominance and matching analyses

May 2018

## Annex 4: Methodology of the dominance and matching analyses

### Introduction

1. In this Annex we describe the data used and explain our approach to modelling product choice in Chapter 5 of the Interim Report. For further details, please see Occasional Papers 33 and 34 published alongside this interim report. This Annex is intended to assist readers who wish to understand better the analysis we have conducted but who may not need the full detail of the Occasional Papers.

### Data used for the dominance and the matching analyses

2. The dominance and matching analyses take advantage of a broad evidence base that is unavailable in most retail financial markets. These include: our extensive regulatory data on UK mortgages (PSD001), market-wide detailed product data (Moneyfacts) and data on several million mortgage consumer credit files.
3. The core dataset is our data of historical mortgage transactions in our Product Sales Data, PSD001. These are confidential data reported to us by lenders whenever they sell a mortgage. The data contain detailed information about each transaction (including date, lender, intermediary, amount borrowed, rate type, postcode, purchase price, date of birth, and historical adverse credit events).
4. We intentionally restricted our analysis of PSD001 in several ways. It only includes residential first-charge mortgages as these are the focus of the MMS. It also excludes mortgages that are part of government initiatives (eg Help to Buy) and interest-only mortgages. This is because these loans involve nuanced conditions that are not very well captured in the data. It also excludes mortgages with LTVs less than 20%, loans smaller than £20,000, and those with terms less than 5 years. These final exclusions mean that we focus our analysis on mortgages representing significant financial stakes.
5. PSD001 data are also matched to two other data sets. We match these to Moneyfacts data on the products that were available on a daily basis from 2011 to 2016 H1. This dataset has detailed information on the products on the market at the time - the complete price structure, extra features and lending standards. These data are used to identify the products on the market at the time when each mortgage was transacted.<sup>1</sup> The data also do not cover internal switches, as PSD001 does not including internal switches.
6. Our analysis only considers products that are captured in Moneyfacts data. Where a product is missing, this could lead to the analysis underestimating potential harm to consumers who bought comparable products. The main example of missing products is deals that are exclusive to intermediaries, some of which are not captured in

<sup>1</sup> To take into account the lag between when choices are made and when the mortgage was reported in PSD (date of mortgage origination) the date of the mortgage choice is assumed to be 3 months before an origination date in PSD.

---

Moneyfacts. The Moneyfacts dataset does contain a large proportion of products on the market, so effects here should be limited.

7. Consumers' PSD001 data are also matched to their credit reference data on selected variables at the time of the mortgage transaction. These credit reference data were provided by the one of the UK's main agencies. The data are used to add information on consumers' credit risk at the time of their mortgage transaction. These data cover over 90% of UK mortgage transactions.
8. In addition to these data sources, we use some more specific datasets to supplement the main data sources above. For example, we use HM Land Registry Price Paid Data to identify whether a mortgage property is newly built. We also use selected variables from the UK 2011 Census as a proxy for demographic information. These include average employment rates, educational attainment and socioeconomic status by postcode.

## Dominance analysis

---

9. The aim of the dominance analysis is to identify the extent to which consumers are missing out on cheaper mortgages that have similar features and would also meet their needs. It also assesses how much consumers might save if they were to buy a cheaper alternative mortgage.

## Methodology

10. For the dominance analysis, we analyse the sub-sample from 2015 to 2016 H1 to cover the most recent post-MMR data over a reasonable time period.
11. For each mortgage transaction in the PSD001 sample the model identifies the other mortgages sold at the time (using the Moneyfacts data) with the same non-price features as the purchased mortgage. To avoid being skewed by products that were only briefly on the market, products had to be on the market for a month or more. A large number of product features are captured, such as deal type (eg fixed, tracker etc.), deal period (eg 2 years, 5 years), maximum LTV for the product, ability to make over- or underpayments. This helps to ensure we compare like-for-like products when identifying alternative mortgages.
12. Among these comparable mortgages it then identifies which of these the consumer was likely to be eligible for. To do this it uses eligibility data in the PSD001 and Moneyfacts and the credit score of the mortgage consumer (using the credit reference data). It only treats mortgages as available where these had already been sold to someone else whose credit score was the same or lower than the consumer's, and where the consumer satisfied the eligibility criteria captured in the PSD001 and Moneyfacts data.
13. Finally, it identifies cheaper available alternatives from these. It does this by ruling out any mortgage with a price feature more expensive than a corresponding price feature in the purchased mortgage (ie it has a higher headline rate or a fee that is higher). It also rules out mortgages where price features are all identical because these mortgages would have been the same cost, not cheaper.
14. The mortgages that remain are taken to be the alternative, comparable and available mortgages that were cheaper in some way and not more expensive in any other way. These products arguably represent 'better value' because they provide the

same (observable) key features at a lower price. These mortgages are said to 'dominate' the purchased mortgage. Where the model finds cheaper, comparable alternatives to the purchased mortgage, then that sale is described as dominated.

15. For these dominated sales the model estimates the forgone savings per year. It regards these as the average difference in the cost between the purchased mortgage and the cost of the cheaper mortgages over the introductory deal period. In some cases where reversion rate is relevant it also estimates savings over a five year period. This assumes borrowers stay on reversion rate after the introductory deal period. We do not use this cost metric as our estimate of savings because of the high proportions that switch, it is used instead to capture some potential savings from reversion rate differences.

### Limitations of the dominance analysis

16. Our approach does not make assumptions about consumers' preferences on mortgage features, and so does not assess or compare suitability of different alternative mortgages. Where consumers selected a mortgage with some features that the consumer did not need or want, the estimate will not capture savings from available products without these unwanted features.
17. There will be cases where a consumer bought a mortgage where an alternative suitable mortgage was available with a different fee and rate structure. Some element of the cost was higher but the overall cost for them would have been lower. For example, higher interest rate or higher fee. Such a mortgage would not dominate the purchased mortgage, because it has some price element that is higher. By not including these as cases where a consumer could have purchased a cheaper mortgage the estimates do not capture all cases where a consumer may have saved money.
18. Customers may have knowingly chosen a mortgage that was not the cheapest. This could be because they value speed of service, brand recognition or lender reputation over cost, or where their eligibility depends on criteria not in our data. In these circumstances the potential savings would be an over-estimate. We estimate that these possible explanations will have little effect on our harm estimates overall. For example, we compared borrowing costs for those switching with borrowing costs for first-time buyers and home-movers with similar characteristics. Switching customers are less focused on speed of service, but this is more important for first-time buyers and home movers. Similarly, we have compared those whose eligibility is more likely to be well captured in the data (low LTV, high credit score borrowers) with the others where it is less well captured. In both cases differences in the outcomes were relatively small suggesting limited bias in our estimates of potential savings.

### Matching analysis

19. The matching analysis builds on the dominance work. It assesses what difference taking advice or buying through an intermediary makes to the likelihood of buying a dominated product and to how much could have been saved.
20. The matching analysis has 2 parts:
  - i. assessing the impact of advice on product outcomes on those who before the MMR would not have received advice; and

- ii. assessing the consumer impact of using an intermediary rather than approaching a lender direct.

- 21. The matching analysis only covers first-time buyers and home movers because PSD001 does not capture internal switches with existing lenders.

### Methodology for assessing the impact of advice

- 22. The usual way to identify the impact of advice using matching would be to compare outcomes for those who obtained advice against the outcomes for those who did not. This would compare similar consumers with each other to avoid 'selection bias'.<sup>2</sup> We cannot apply this method here because, since MMR, very few new mortgages are bought without advice (ie execution-only). So the sample of non-advice taking consumers is too small to compare outcomes and identify the effect of advice.
- 23. Instead, we bring in additional data from before the MMR and apply a more complex 'difference-in-difference' method to isolate the effect of advice on the relevant population of consumers - those who before the MMR would not have received advice. The rationale is to bring in data on advised and non-advised sales from before the MMR. We then use data over time to correct for other market changes, yielding an estimate of how advice has affected previously non-advised consumers.
- 24. To estimate the effect of post-MMR advice, we estimate the trend in outcomes for consumers who took mortgage advice. We do this by comparing outcomes for those who got advice before the MMR rules came into effect with the outcomes of similar consumers (by propensity score) who got advice after that point. We assume there has been the same trend in outcomes for consumers who don't take advice compared to those who do.

### Methodology for assessing the impact of intermediation

- 25. The methodology for assessing the impact of intermediation where a consumer gets advice is simpler. Here, we have sufficient post-MMR data to compare outcomes for those who got advice in an intermediated sale, against those who got advice in a direct sale.
- 26. The methodology is similar to the first part that looks at advice, but without the difference-in-difference element. Here we also group similar consumers together and within these compare outcomes for those who used an intermediary against outcomes for those that went direct.

### Limitations of the matching analysis

- 27. The modelling of both the impact of advice and intermediation rests on an assumption that all relevant factors have been taken into account in grouping similar consumers together.
- 28. To overcome the lack of data from few consumers getting advice following the MMR, we use the difference-in-difference method which relies on a 'common trend' assumption. This assumes that similar consumers who got advice and those did not get advice have their outcomes affected to the same extent by other factors.

---

<sup>2</sup> For example, if consumers with complex circumstances tended to go to intermediaries and consumers with straightforward circumstances tend to go direct, then without comparing similar consumers, we might wrongly attribute the impact of unusual circumstances on outcomes to intermediation.

**Financial Conduct Authority**



© Financial Conduct Authority 2017  
25 The North Colonnade Canary Wharf  
London E14 5HS  
Telephone: +44 (0)20 7066 1000  
Website: [www.fca.org.uk](http://www.fca.org.uk)  
All rights reserved