Technical Annex to Evaluation Paper 18/3

December 2018

Introduction

This technical annex provides supporting detail to the analysis conducted within the evaluation of reducing barriers to entry (EP18/3) report.

The annex covers the following areas:

- Section 1: details the analysis to determine whether the number of joiners to the UK banking market has increased relative to the corresponding number of joiners to the rest of the EU banking market since barriers to entry were lowered
- Section 2: describes the analysis on whether post-review entrants have any perceptible difference in performance relative to incumbents and older entrants
- Section 3: explains the analysis to determine whether the interest rates offered by post-review entrants on cash savings deposits are higher or lower than those offered by incumbents
- Section 4: details the analysis to determine whether the interest rates offered by post-review entrants on mortgages are higher or lower than those offered by incumbents

Each section gives a short background, an explanation of the tested hypothesis, an overview of the data and dataset construction and the methodology adopted to examine the data, and the results and conclusions of each piece of analysis.

Section 1: Relative Number of EU and UK Entrants - Difference-in-Difference Analysis

Background

We wish to identify the effect of the regulatory reforms to barriers to entry within the UK market versus other wider changes which may have affected firms' decisions to enter. This allows us to separate the impact of UK-specific reforms relative to changes that may have been occurring in the global banking sector (such as technological changes or macroeconomic conditions).

A simple method to achieve this is a comparison of the rate of entry in the UK market relative to 26 other EU countries. If we do not observe a similar increase in the number of licences issued in these other EU countries during the 2013-2017 period, and we previously observed similar trends in the number of entrants prior to 2013, then the increase which is observed in the UK can credibly be attributed to UK specific changes in 2013 (such as the 2013 review).

Hypothesis

We expect the publication of the report in 2013 should lead to an increase in the expected number of joiners in the UK banking market when compared to the expected number of joiners to the rest of the EU banking market.

Using a dataset based on European Central Bank (ECB) figures on the number of entrants and a difference-in-difference regression approach, we formalise this comparison to test whether there is a statistically significant increase in the average number of banks entering the UK relative to the rest of the EU.

Data and Dataset Construction

For the regression analysis, we have used data on the number of Monetary Financial Institutions (MFIs) joining each EU country that is submitted to the ECB by the respective central banks on a monthly basis.[1] We noted that the ECB series for the UK does not match data held by the FCA on the number of banking licenses issued, likely due to differences in reporting procedures. We have therefore replaced the UK MFI joiners' series with FCA data on the number of banking licenses issued each quarter collected as part of the authorisations process.[2]

1 MFIs encompass the following undertakings: central banks, credit institutions (including banks), other deposit-taking corporations and money market funds.

2 We note that the major difference between the MFI series and the FCA data is the timing of when banking permissions were granted. In aggregate, the number of MFI joiners submitted to the ECB is very similar to the aggregate number of banking joiners drawn from the FCA data for the 2009 to 2017 period. In particular, the

The 27 countries which are included in our analysis are: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden and the United Kingdom.[3]

For the purposes of the difference-in-difference (DiD) regression, we aggregated the analysis to quarterly intervals, and supplemented with quarterly GDP data for 27 EU countries (the EU 28 minus Ireland) from 2009 Q1 to 2017 Q2. This data is extracted from Eurostat.

Methodology

We are comparing the number of banking licences issued (ie the number of successful entrants) in other EU countries to the UK in each quarter. As the distribution of entry in each quarter tends to be positively skewed, a regression analysis which assumes a normal distribution of the data points will not necessarily provide the best approximation. As a traditional OLS model assumes a normal distribution, we also consider three alternative specifications for the DiD regression, which rely on the following distributions: Negative Binomial, Poisson and zero-inflated Poisson.[4]

We use the following DiD model using the Negative Binomial and Poisson distributions in conjunction with OLS (which estimates the baseline case):

$$y_{it} = \sum_{t=1}^{36} \beta d_t + \gamma \cdot T_i + \delta \cdot DID_{it} + \theta \cdot GDP_{it} + \varepsilon_{it}$$

This model is estimated using a 'fixed-effects' approach at the country level, ie we attempt to remove any time-invariant factors which may influence the number of joiners in a given country.[5] Notation is as follows:

- y_{it} represents the number of joiners for a given country 'i' in time period 't'
- d_t is a dummy variable which is 1 in time period 't' and zero otherwise
- T_i is a ''treatment'' dummy, which is 1 for the UK and 0 otherwise
- DID_{it} is an interaction dummy,[6] which is 1 for the UK in time period 't', and zero otherwise
- GDP_{it} is GDP for country 'i' in time period 't'
- ε_{it} is the iid error term

For this regression design to function correctly, we require 2 core assumptions to be met:

- Common Trends: prior to the intervention the number of banks entering the UK and the number of MFI joiners to the EU were following a similar trend

use of FCA data on when banking permissions were granted corrects for a large influx of MFI licenses issued in Q1 2013 in the ECB data, which is inconsistent with the FCA's data on the timing of these licenses.

3 We also noticed that there appear to be issues with the data for Ireland. We have removed the data observations for Ireland, treating the series as an outlier.

4 A zero-inflated Poisson distribution is a variant of the Poisson distribution which accounts for distributions where there are a large number of expected zeros. This may be appropriate when considering that for a number of countries, including the UK series no banks enter for a number of quarters.

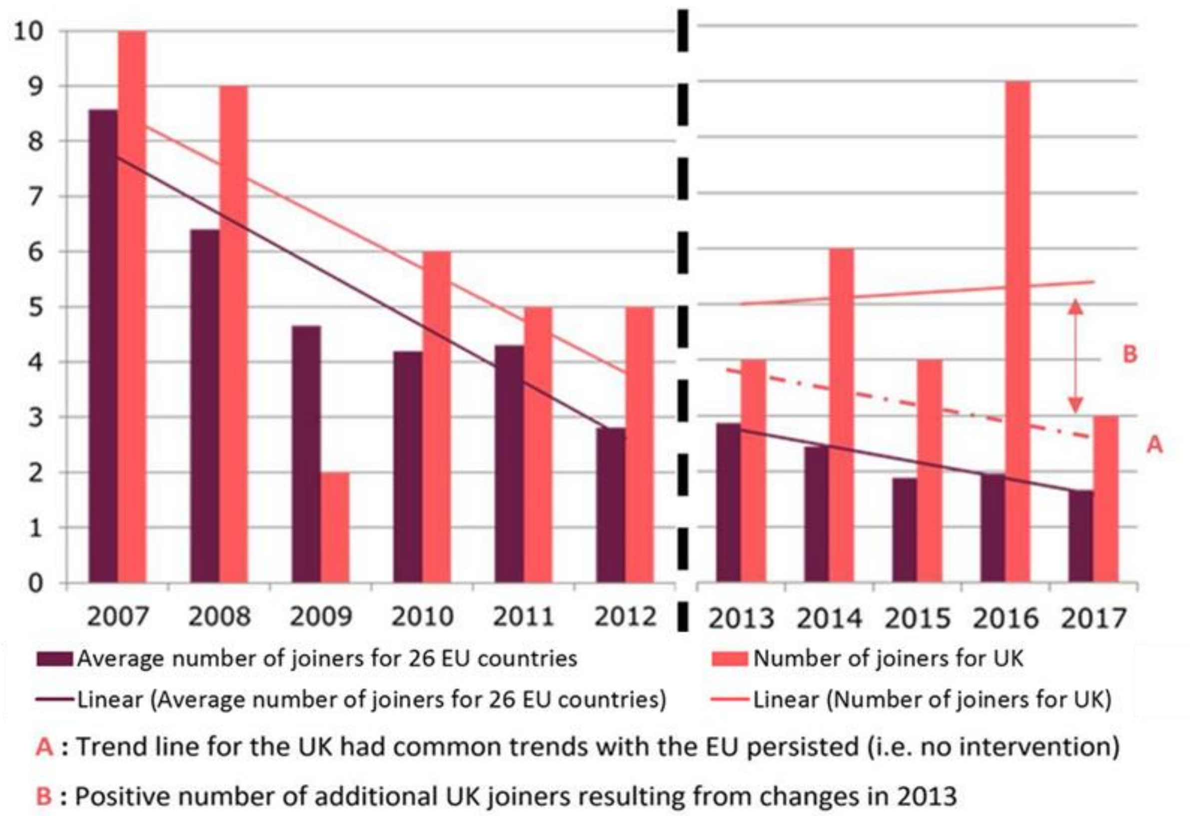
5 Note that with the exception of the zero-inflated Poisson model these models are estimated using a fixed-effects approach ie each coefficient represents a deviation from its time invariant mean e.g. $y_{it} - \bar{\bar{y}}_i$.

6 The interaction dummy captures the effect of the post-review time dummy and the UK country dummy, effectively isolating the effect for the UK post-review on the number of MFI joiners to the UK market.

- Group Exogeneity: The impact of the reform should not cause joiners to move from one group to another, ie the reforms should not lead to firms in the EU market becoming a UK joiner by leaving the EU first.

Figure 1 below demonstrates a very similar declining trend for both UK and EU joiners between 2007 and 2013, suggesting the common trends assumption is satisfied.

Figure 1: Banking licences issued in the UK and EU



Source: FCA Analysis of ECB data and FSA data. Note the data on banks authorised includes all firms that have been authorised in the UK, including both UK domiciled and foreign banks.

By construction, the group exogeneity assumption holds: firms which enter the EU cannot become UK joiners as they will be treated as a passporting firm. Passporting firms are not captured within the figures reported in the ECB dataset, nor in the figures we have used which were drawn from the authorisations process.

It should be noted that this analysis does not necessarily ascribe causality of the increase in the number of joiners in the UK relative to the EU *solely* to our 2013 intervention. Indeed, were there to be any other factors affecting the EU or UK banking market around the same time as the report was published, it would not be possible to disentangle the effect of this alternative factor relative to the effects of the 2013 review.[7]

7 We recognise that these EU countries have experienced specific shocks which the UK was not subject to, such as the Euro-crises. We consider it unlikely that these shocks have affected all 26 countries at the same time; i.e. the Euro-crisis would have impacted entry into the banking sectors in these countries in different ways and at different times. We have controlled for the impact these shocks may have had on these countries, through the inclusion of country-specific GDP and dummy variables within the regressions.

Results

The results of these models are shown below. With regards to the zero-inflated Poisson model, it is not possible to use the fixed-effects approach directly.[8] We therefore adopt a least-squares dummy variable (LSDV) approach, inserting a dummy variable for each country. This is mechanically equivalent to using a fixed-effects approach.[9] That is, each dummy variable will capture any country–specific effects and will not vary over time.

We adjust our standard error estimates for heteroscedasticity using heteroscedasticity robust standard errors.[10]

Table 1: Difference-in-difference Regression Output[11]

	OLS	Negative Binomial	Poisson	Zero Inflated Poisson
DID	2.077*** (4.82)	0.667* (2.00)	1.020*** (4.63)	1.074*** (3.85)
GDP	-0.0000117** (-3.19)	0.000000580 (0.70)	-0.00000215 (-1.19)	-0.00000415* (-2.35)
Constant	2.203*** (3.97)	0.479 (1.13)		-0.342 (-1.20)
Time Dummies	Yes	Yes	Yes	Yes
inflate (Joiners)				-45.59*** (-539.51)
Constant				22.90*** (454.59)
Observations	826	826	826	826

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: FCA regression analysis of ECB and FCA data

For each model reported above, we see (the DID coefficient) a positive and statistically significant effect of the 2013 reforms on the number of UK banking licences granted. This indicates that the UK has experienced a comparative increase in the expected number of banking licences issued each quarter relative to 26 other countries in the EU.[12]

8 The particular programme we have used to generate our regression results (Stata) does not directly allow for fixed effects estimation of a Zero-inflated Poisson model

9 This is a special case of the Frisch-Waugh theorem.

10 Vogelsang, T.J., 2012. Heteroskedasticity, autocorrelation, and spatial correlation robust inference in linear panel models with fixed-effects. Journal of Econometrics, 166(2), pp.303-319.

11 Note that while our initial specification includes a "treatment" dummy as in a standard DID model, since we are estimating using a fixed effects approach, this coefficient will be dropped during estimation. The reason for this is that the status of the "treatment" indicator does not vary over time and cannot be estimated separately from the constant term.

12 Given the data concerns we noted above, we have also tested the robustness of our findings through a series of falsification tests. Except for the exclusion of Germany from the negative binomial model, we find all of our results remain positive and significant (at 10%) to the exclusion of any individual country's data. Note the exclusion of Germany leads to a lack of convergence in the maximum likelihood estimation procedure as

Nevertheless, the magnitude of the expected increase differs across models. The Poisson and zero-inflated Poisson model indicate that the expected number of licences issued by the UK when compared to other countries in the EU has increased on average by 1 licence per quarter. In contrast, the OLS (and negative binomial) regression find that on average the UK grants two (and 0.67) licences per quarter compared to other countries in the EU.

Based on the most conservative result (0.67), we estimate that in the 4 years since the 2013 review (ie 16 quarters) the UK authorised 10 more banks relative to 26 other countries in the EU. That is, accounting for EU-wide factors we show that since the 2013 review the UK has experienced a comparative increase in the expected number of licences issued per quarter relative to other countries in the EU.

Conclusion

Using a variety of specifications, we have demonstrated a statistically significant, positive difference in the expected number of UK licenses issued when compared to the expected number of EU licenses following the 2013 reforms. More precisely, the expected number of entrants into the UK banking sector when compared to 26 other countries in the EU has increased by at least 0.67 banks per quarter since the reforms were introduced in 2013.

opposed to an estimated insignificant result ie the issue arises from the iterative procedure used to search for a concave solution rather than the data demonstrating an insignificant coefficient. As such we are reasonably confident our results are not driven by any one country’s data.

Section 2: Performance of Entrants Post-Review - Descriptive Analysis

Background

We wish to examine whether entrants post-reform show any perceptible difference in performance relative to incumbents and older entrants. Establishing whether there is a difference in performance for post-review entrants is an important indicator of potential benefits which may be passed onto consumers, as identified by the causal chain in the main evaluation report.

We address this question by looking at regulatory returns provided to the FCA on balance sheet items, such as a firm's total volume of retail deposits and lending. Considering performance in terms of the growth of these two particular activities, we compare the activities of entrants after the intervention relative to entrants before the intervention. These entrants provide a reasonable comparator group to post-2013 entrants, having entered the market following the financial crisis, and we have data on their balance sheet activities in the early years of operation. We benchmark these two entrant groups relative to incumbents to help account for any market-wide factors which may have affected the performance of the banking sector.[13]

Hypothesis

We expect both entrant groups to grow faster than incumbents: faster growth will be more achievable from a smaller absolute starting value.

We also expect that the growth for post-review entrants is faster relative to pre-review entrants when compared at the same point in their respective lifecycle.[14] For example, we are comparing the growth of pre-review entrants in the first 2 years of their lifecycle to the first 2 years of post-review entrants.

Data and Dataset Construction

The basis for the primary dataset used for performance analysis is regulatory returns data from FSA001. This information is supplemented with internal FCA data on permissions, limitations and the group structures of firms.

Dataset construction

We focus on firms with permissions to accept deposits. We consider the ability to accept deposits as a primary feature of a bank.

13 We note that that throughout this time-period the interest rate environment has remained constant.

14 As part of the changes to the authorisations process the cost structure for post-review entrants would be expected to differ from pre-review entrants. We anticipate a lower variable cost base, which in turn may result in the ability to grow deposits and lending at a faster rate.

Alongside the types of permissions held by a firm, the FCA permissions dataset also provides detail on the date authorisation took place.

As a number of financial institutions may accept deposits, we use information on the limitations applied to firms to remove those types of institutions who accept deposits but who would not be considered banks e.g. credit unions, insurers, friendly societies, and building societies.

We also use data on the limitations applied on firms to identify the length of time that firms are in the mobilisation process, since firms have limitations applied to the quantity of deposits they can take whilst in mobilisation. We therefore adjust the age of firms to begin from the point they exit mobilisation ie a firm is considered authorised as soon as they exit mobilisation.

To add an additional level of analysis to the consideration of what constitutes entry, the dataset was supplemented with information relating to the corporate group(s) a firm is part of. We established a unique group reference for each firm by aggregating up corporate groups, where they are part of the same corporate structure. This process is performed iteratively until there is no further aggregation that can be performed. As discussed in the report, we do not consider banks to be 'entrants' in instances where the firm is already part of a pre-existing banking group (ie a group with a bank already in it).

Finally, the list of banks under consideration is corroborated by the list of banks which have been processed and identified as new entrants by the FCA Authorisations team.

The dataset

The base dataset contains the following information in a cross-sectional format:

- entry and exit of the firm, provided by the passporting and permissions data
- which group each firm belongs to
- a list of 'Banks' by a traditional definition of 'Accepting Deposits' as a primary activity
- filters to identify those banks considered to be 'entrants'

This dataset is then supplemented with data on the balance sheet of each bank. This information is all provided within FSA001 and FSA002, both of which are mandatory submissions for banks.

FSA001 provides broken-down information on Assets and Liabilities of each firm, while FSA002 focuses on the Income Sheet of the firm. These two data sources are in panel format, providing the balance sheet information on a quarterly or half-yearly basis. Combining this with the core dataset provides a panel dataset allowing us to track a) the balance sheet information and b) the authorisation/permissions status of a given firm in any given quarter between Q1 2009 and Q2 2017.

For a given firm, in a given quarter, we are therefore able to identify time-varying information (such as balance sheet figures, the age of the firm, and whether the firm is still conducting deposit-taking activities), as well as time-invariant information such as the initial entry-date into the dataset and whether the firm is considered an entrant or incumbent.

Data Limitations

For younger banks, balance sheet data is limited due to the slight lag in firm submissions, which leaves limited information available for analysis.

We also note that following our characterisation of what constitutes an entrant, we are left with a small population of firms either categorised as a 'pre-review entrant' or as a 'post-review entrant' for comparison against the incumbent group. This will mean that our results may be susceptible to distortion by the lending and deposit taking activities of individual firms in each category.

Methodology

In the main report, we consider the performance of post-review and pre-2013 entrants relative to incumbents. We do this by considering the growth in retail deposits and lending relative to incumbents.

For the pre-review entrants, we conduct this comparison between 2011 Q2 and 2013 Q1, indexing the bank's activity from 2011 Q2. For the post-review entrants, a similar comparison is carried out, starting from 2015 Q3 until 2017 Q2, with the indexing starting from 2015 Q3. This allows us to consider the growth of entrants over a 2-year period for both entrant groups while using the growth for incumbents as a visual indicator for any potential market-wide factors which may have influenced performance.

We extend the analysis shown in the report by breaking down the incumbent group into 'small', 'medium' and 'large' firms. The categorisation of firms is based on the size of their total assets at the indexing point (2011 Q2 for comparisons with the pre-2013 entrants, and 2015 Q3 for the post-review entrants). We categorise firms as follows:

- small: £0-20bn
- medium: £20-100bn
- large: £100bn or more

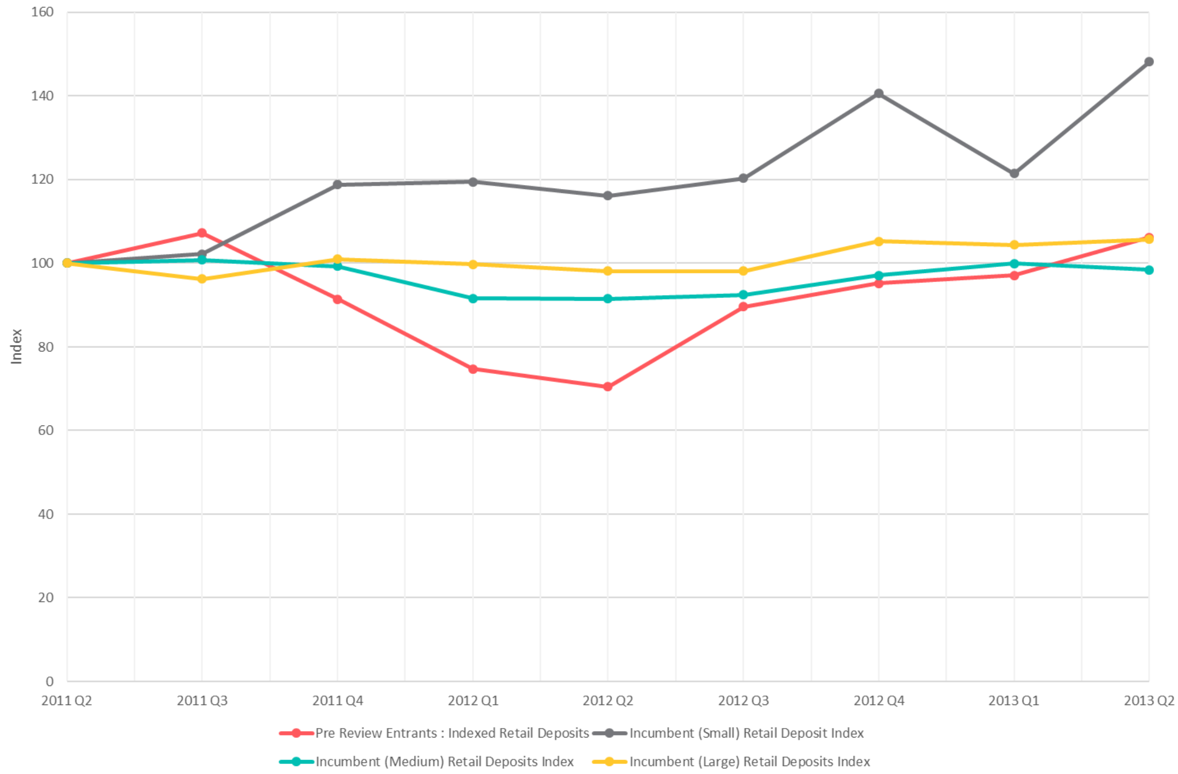
The intention of this analysis is to provide a more detailed look at the actions of the incumbent group relative to the 2 entrant groups, as well as to provide a set of more focused comparator groups for assessing the entrant groups' performance.

We further extend the analysis by comparing the absolute value of retail deposits and lending of post-review entrants to pre-review entrants by the age of the bank.[15] This analysis allows us to compare the evolution of lending and deposits between the two entrant groups and observe if there is an obvious difference in the growth of these activities by entrant group. This provides a more direct comparison of performance between the two groups of entrants.

15 Age here is defined as time since exiting mobilisation. The reason for this is that mobilisation imposes restrictions on the activities a bank can engage in, as discussed in the main report. As such we attempt to confirm firms starting from the point at which there are no restrictions applied to deposit-taking activities.

Results

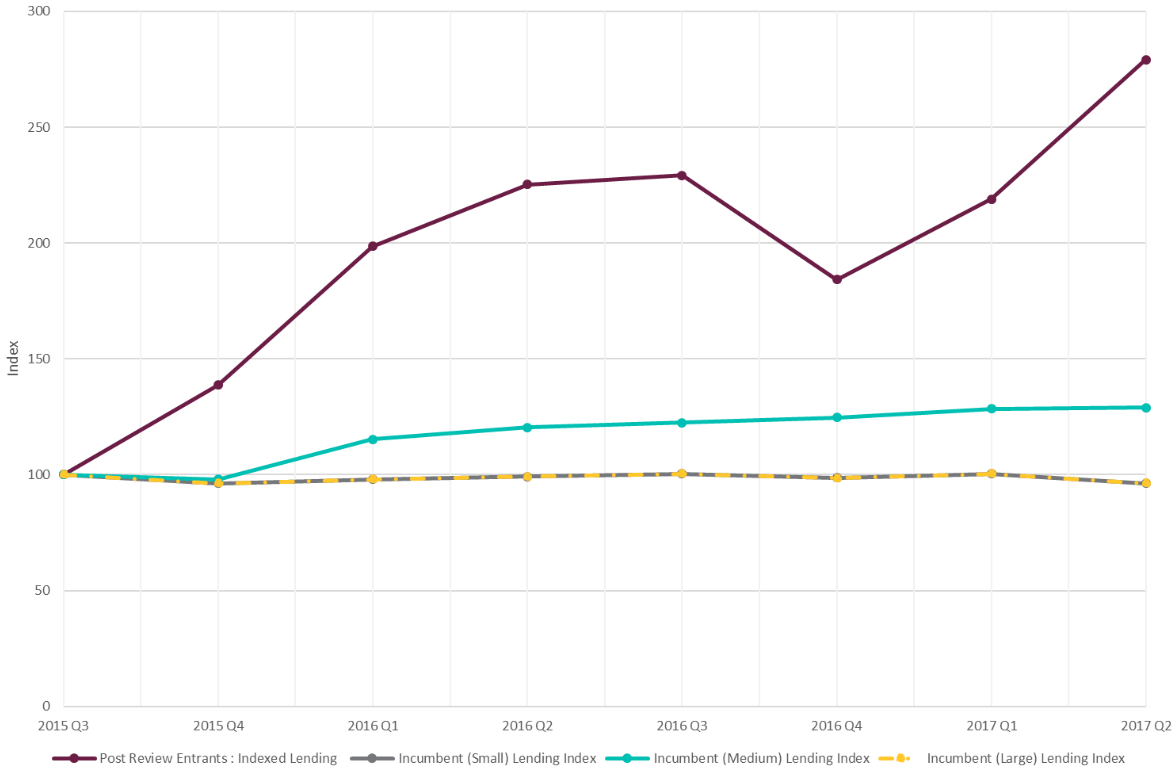
Figure 2:Average Growth in Retail Deposits Received by pre-2013 Entrants



Source: FCA Analysis of FSA001 submissions

Figure 2 breaks down the incumbent group's retail deposit growth in the pre-review comparison period. The small-sized firms are the primary drivers of growth in the incumbents, consistently outstripping growth in the other groups.

Figure 3: Average Growth in Retail Deposits Received by Post-2013 Entrants

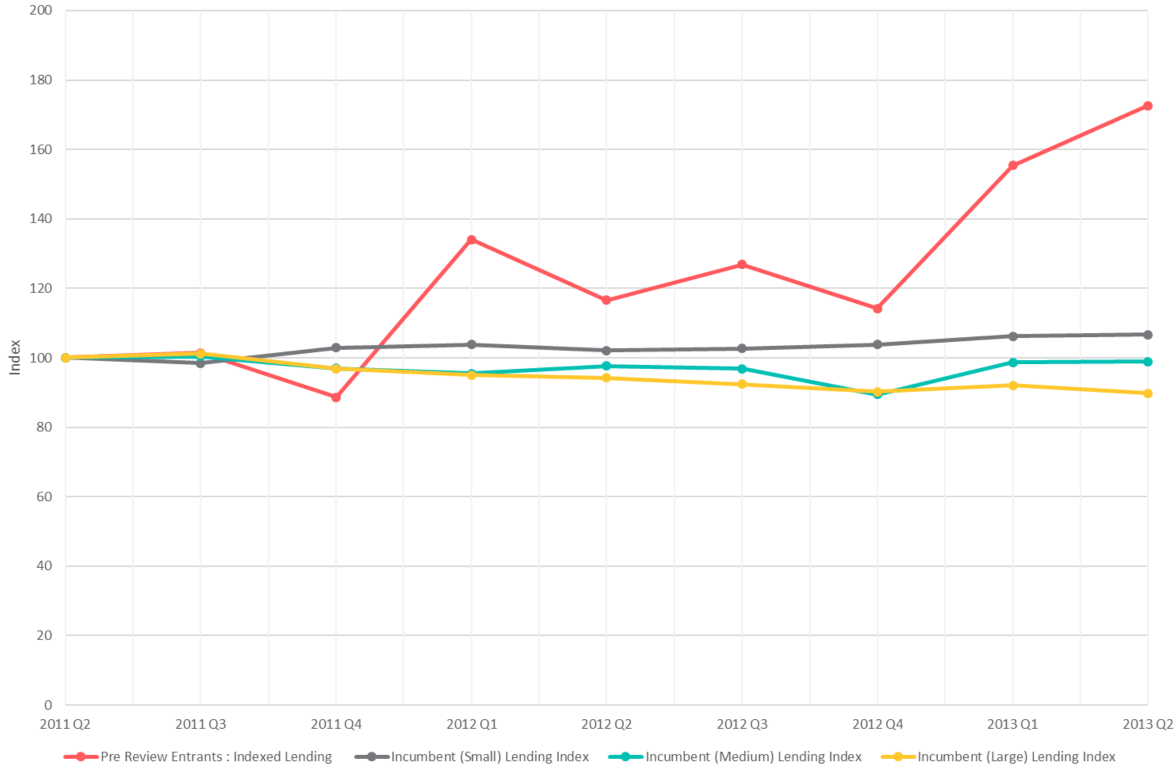


Source: FCA Analysis of FSA001 submissions

Figure 3 shows less of a spread amongst the incumbent group. While medium firms still outstrip smaller and larger firms in terms of retail deposit growth, the general trend is very similar across all incumbents. Given the smaller incumbents do not seem to be experiencing the same levels of growth when compared to the entrants (almost exactly mirroring the growth of the large incumbents), we can infer that the strong performance of the post-2013 entrants is not purely attributable to their size.

The general success of medium-sized firms when compared to the rest of the incumbent group may also suggest certain sized firms are particularly engaged in the retail-deposit space, ie there is a correlation between total assets and the business models of these firms.

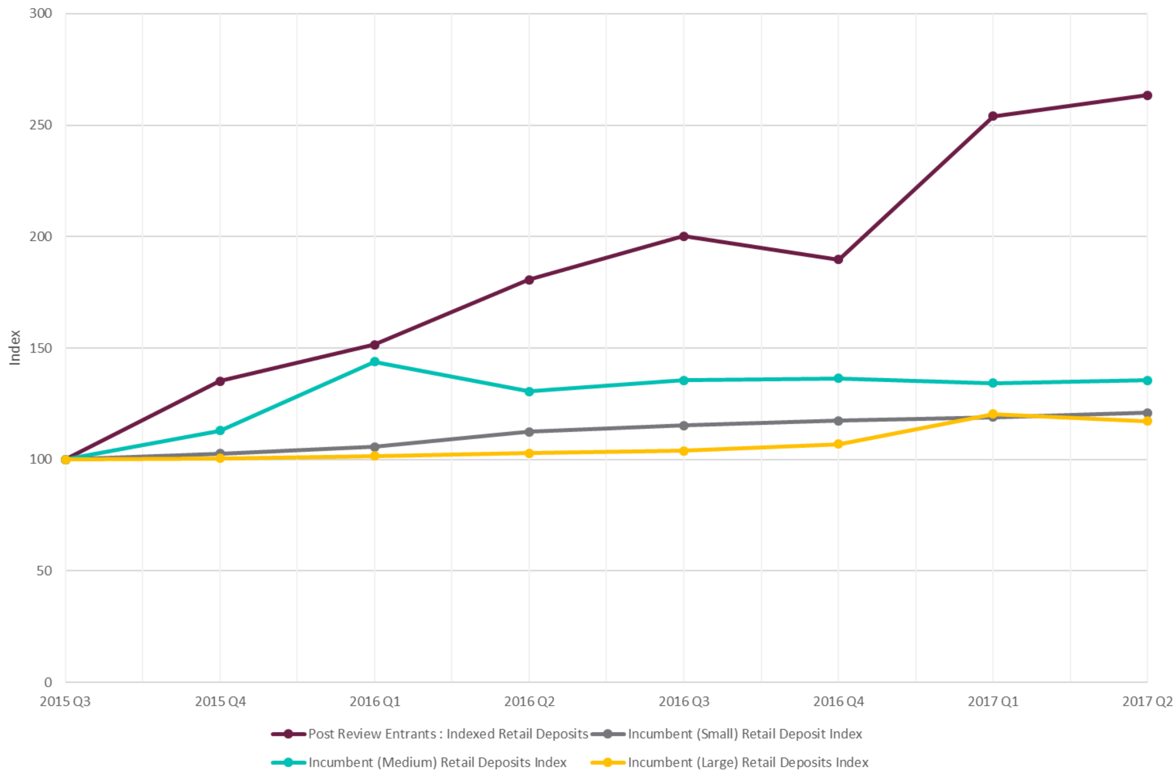
Figure 4: Average Growth in Lending by Pre-2013 Entrants



Source: FCA Analysis of FSA001 submissions

In contrast to the retail deposit analysis, Figure 4 shows that lending has grown relatively faster for the smallest incumbent firms, and slowest for the larger firms. This is intuitive and consistent with our original hypothesis: we would expect higher marginal growth for smaller firms. The lack of any major differences in the incumbent groups implies that lending over this period has not been particularly affected by external market factors.

Figure 5: Average Growth in Lending by Post-2013 Entrants



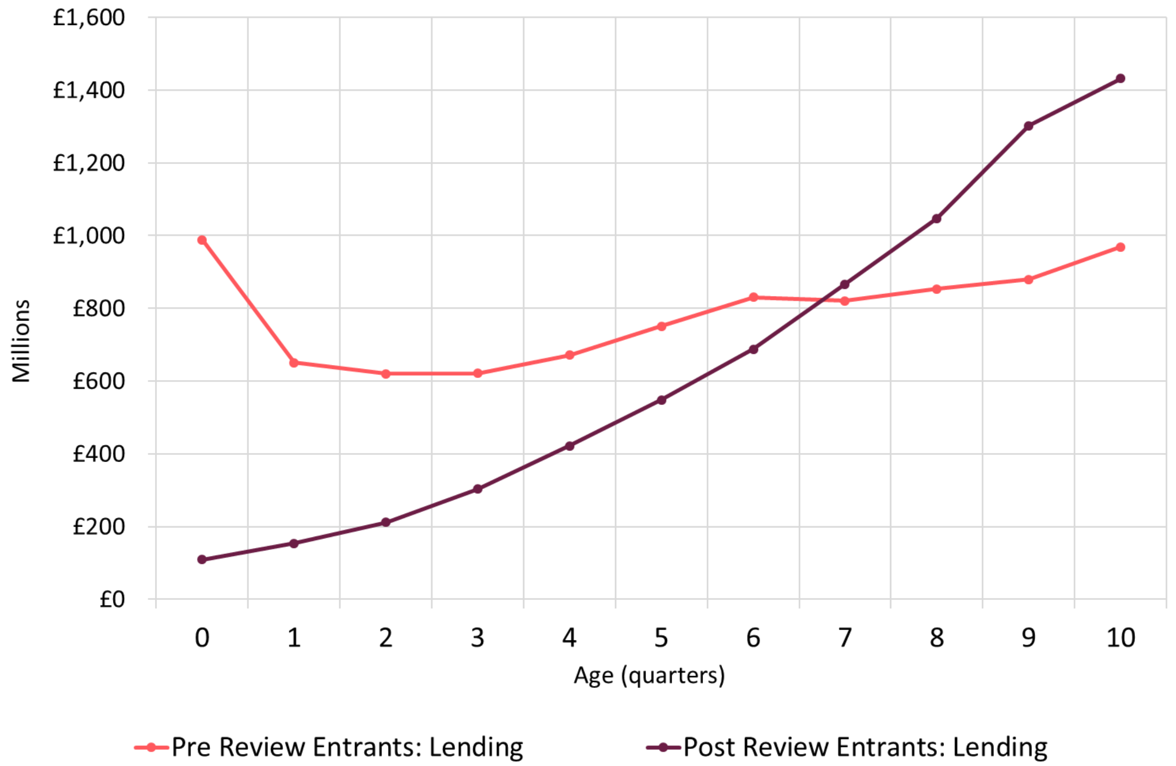
Source: FCA Analysis of FSA001 submissions

The story for growth in lending for the post-2013 entrant comparator group is, however, a lot more like that seen in relation to retail deposits, as shown in ed by external market factors.

Figure 5. Medium-sized firms once again outstrip smaller and large firms; however, all three groups of firms are broadly following a very consistent growth over time.

Again, the significant difference for both entrant groups relative to the small incumbents suggests that the strong performance of entrants is not purely down to them benefitting from high marginal growth on account of their size.

Figure 6: Average Absolute Value of Lending by Entrants

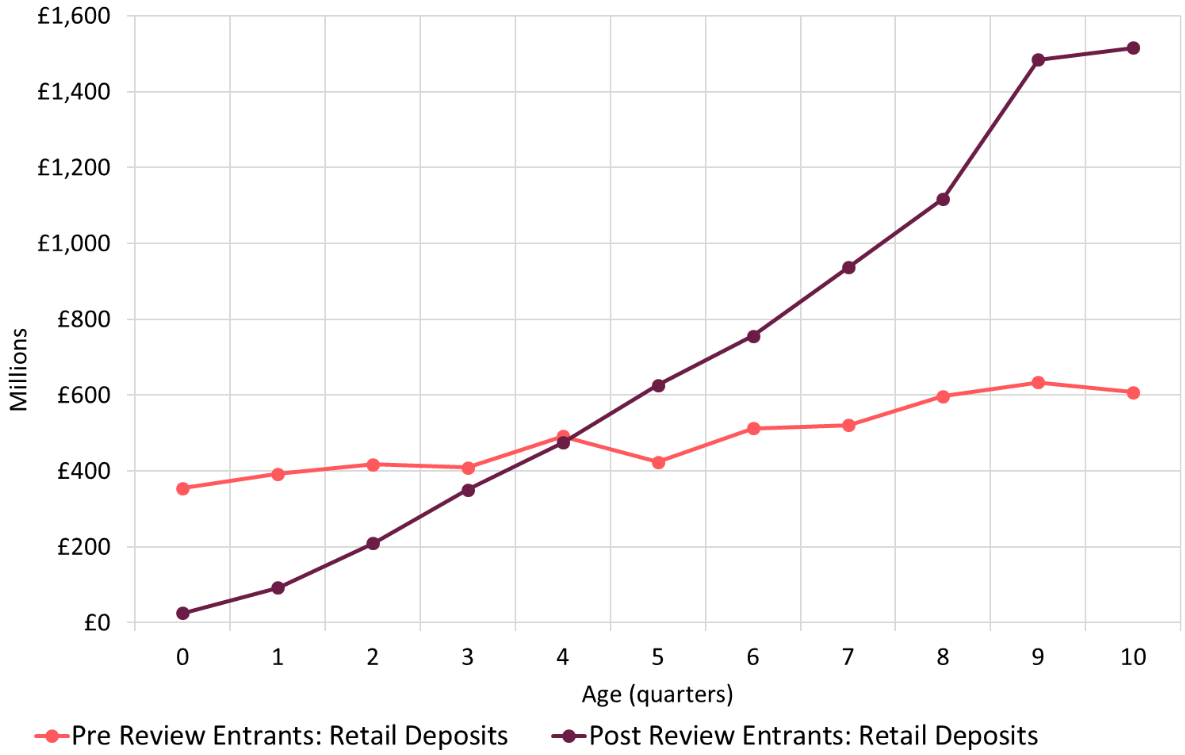


Source: FCA Analysis of FSA001 submissions

Figure 6 compares the lending of the two entrant groups by the age of the banks. This exercise allows us to compare banking groups at similar stages in their lifecycle.

Typically, post-review entrants start with a lower level of lending, but grow much faster than the pre-review entrants. Post-2013 entrants are consistently lending more on average than the pre-review cohort after their 6th quarter of operation.

Figure 7: Average Absolute Value of Retail Deposits for Entrants



Source: FCA Analysis of FSA001 submissions

Figure 7 depicts a broadly similar story for retail deposits, albeit the pre-review entrants have less variability, with consistent, but slow growth over the course of their lifecycle. The post-review entrants in contrast, display very consistent high growth in retail deposits over the first 2 years of their average lifecycle, with the post-2013 entrants outstripping the pre-review entrants by the 4th quarter of operation.

Conclusion

The additional detail provided by segregating the incumbent groups by total asset sizes suggests that there does appear to be a difference in the speed of growth depending on the size of firms. For pre-review entrants, while retail deposit performance is consistent between medium and larger incumbents, small-sized incumbents seemed to perform relatively well, suggesting that banks engaged in retail deposit-taking activities from the pre-review entrants group were not particularly successful in this area relative to certain elements of the rest of the market (as represented by the smaller incumbents). In contrast for post-review entrants, all incumbent groups were consistently outperformed by the entrants with fairly little variation between the different incumbent groups. This may be indicative of a stable retail deposit market in this time period.

For lending, the story is very similar for both pre- and post-review entrants: incumbents of all sizes are consistently outperformed, and despite some spread in performance in incumbent groups, there is nothing visually to suggest that particular sub-segments of the market have experienced any atypical market effects, or that firms of a specific size particularly benefitted from market conditions in this time period.

In general, the strong relative performance of entrants compared to smaller incumbents suggests that performance of entrants is not being driven purely by the benefit of starting from a smaller size base.

In terms of comparison by age, there is a very clear story for both lending and retail deposits: post-review entrants on average start from a smaller base, but typically grow faster, surpassing their pre-review entrant comparators within 1.5 years.

Thus, post-review entrants appear to grow faster than incumbents and faster than older entrants when compared at the same point in their lifecycle. This is consistent with our original hypotheses.

Section 3: Deposits Interest Rates - Pooled OLS Analysis

Background

We want to assess the impact of the 2013 review on outcomes for consumers. The first aspect of this we consider is the variation in deposit interest rates offered to consumers by post-review entrants compared to incumbents. This comparison allows us to partially assess the extent to which the presence of post-review entrants has improved outcomes for consumers.

We use data from Moneyfacts to assess what type of savings products post-review entrants typically offer, and use a simple regression framework to compare the interest rate offered by post-review entrants relative to incumbents.

Hypothesis

Post-review entrants appear to target retail deposits more than incumbents do. We therefore might expect them to offer higher interest rates relative to incumbents on fixed term cash savings accounts to attract consumers to their retail deposit offerings.

Data and Dataset Construction

The Moneyfacts cash savings dataset draws on the database provided to the FCA via the Moneyfacts data-trackers. We wished to construct a dataset which would allow us to reasonably compare savings/deposit products offered by the entrants against the post-review entrant and incumbent groups.

This dataset contains a subset of the banking sector for 2 reasons:

- some banks do not provide data to Moneyfacts
- some banks do not engage in activities within the cash savings sector

There are 10 entrants present within the dataset (3 pre-review and 7 post-review entrants).

The dataset provides information on the broad category of account (e.g. fixed, variable, cash ISA etc) in addition to dummy variables for specific account types such as affinity[16], telephone or online accounts. For fixed products, the term length is provided.

Summary Statistics

Based on the statistics in Table 2, we see post-review entrants typically offer longer term products than the incumbent group and pre-review entrants.

16 Affinity accounts are accounts typically associated with a particular sports club, where some form of donation is made to the club based on the total savings sum across those accounts

Table 2: Average Term Length by Entrant Group

	Term Length (Quarters)
Pre-Review Entrant Average Fixed Product Term Length	4.77
Post-Review Entrant Average Fixed Product Term Length	9.70
Incumbent Average Fixed Product Term Length	6.24

Source: FCA Analysis of Moneyfacts data

Due to the nature of the dataset as a daily record tracking changes to product characteristics, we can infer the date products are taken off the market, and thus calculate the length of time a product is on the market. We conduct a comparison of average duration length by group for products introduced at the start of 2016. This analysis suggests that post-2013 entrants keep their products on the market for shorter periods than incumbents or pre-review entrants.

Table 3: Average duration products held on market (quarters)

Term Length	<1 Year	1 Year	2 Year	3 Year
Pre-review Entrant Average Fixed Product Duration	0.75	1.00	1.00	1.33
Post-Review Entrant Average Fixed Product Duration	0.89	0.94	0.87	0.94
Incumbent Average Fixed Product Duration	2.48	1.05	1.01	1.06

Source: FCA Analysis of Moneyfacts data

An analysis of the account types offered by both entrant groups (Table 4) shows that the majority are in the 'Fixed' account type. To generate a reasonable dataset size for comparisons, we therefore focus our analysis exclusively on 'Fixed' category accounts.

Table 4: Breakdown of Selected Product Offerings by Entrant Group between 2008 and 2017[17]

	Incumbent	Post-review Entrant	Pre-review Entrant
Business Fixed	9428	102	272
Business Variable	2129	18	102
Fixed	161249	1512	925
Variable	50789	52	316

Source: FCA Analysis of Moneyfacts data. Note: That where the interest rate has changed in relation to a fixed term product offered by a firm, we have considered this to be equivalent to a new product offering.

There is reasonable heterogeneity (ie differences) amongst the products in this category. For certain bonds and loyalty accounts, the spread of interest rates is dramatically different from the rest of the product set. To generate a comparator group, we remove these products from the analysis.

We condition the accounts for comparison based on their term length. We round term length (provided in days) to the nearest year and group products accordingly. The below provides some distributional statistics for 1, 2 and 3-year term products.

Table 5: Average Gross AER Interest Rate (%) for Products of 1,2 and 3 Year Term Length

		Q1 2017	Q2 2017	Q3 2017	Q4 2017
1-Year Fixed Term Products	**Pre-Review Entrants**	0.90	0.90	1.73	1.20
	Post-Review Entrants	1.48	1.54	1.72	1.71
	Incumbents	0.93	1.08	1.14	1.23
2-Year Fixed Term Products	**Pre-Review Entrants**	1.00	1.00	1.80	1.30
	Post-Review Entrants	1.63	1.76	1.90	1.97
	Incumbents	1.06	1.32	1.45	1.55
3-Year Fixed Term Products	**Pre-Review Entrants**	1.20	1.20	1.92	1.50
	Post-Review Entrants	1.80	1.89	2.09	2.16
	Incumbents	1.35	1.48	1.68	1.75

Source: FCA Analysis of Moneyfacts data

17 We have excluded certain products from this analysis (such as Instant Access products) due to a lack of offerings from the post-review entrants with which to create a meaningful comparison against the incumbent offerings. The items in the table represent the product categories which have a reasonable number of comparators across incumbents, post and pre-review entrants in terms of the number of products on offer.

We see that across all term lengths the average gross interest rate (mean and median) is consistently higher for the post-review entrant group.[18]

Data Limitations

A major omission of the dataset is the lack of information on non-price characteristics. While we know the interest rate offered on the product, we do not know what additional benefits and/or costs may be associated with it. This means our analysis of deposit taking is purely focused on ‘‘price’’ comparisons.

An important point to note is the low number of data points for pre-review entrants relative to post-review entrants and incumbents when we focus on products offered in 2017 (the time period for which we conduct our analysis).[19] For this reason, we primarily focus on comparisons between post-review entrants and the incumbents across 1, 2 and 3-year fixed term products.

Methodology

We extend this analysis by conducting a regression analysis of the gross AER interest rates, controlling for time effects, firm specific effects, the length of time the product is on market, and the term length of the product.[20] We also condition this analysis on any product dummies we have available, eg current account, branch account, telephone account etc.

Since all variation provided by the explanatory variables is time invariant, we are unable to use techniques such as fixed effects and first difference. However, it is not immediately clear that there are any time-invariant product-specific fixed effects which should remain unexplained after we have controlled for all the product dummy characteristics. Thus, we do not expect an estimation method of Pooled OLS (POLS), our estimation method of choice, to suffer from endogeneity arising from unobserved time-invariant variation. Furthermore, given the clustered nature of Moneyfacts interest rate data we do not believe other time series methods are appropriate.[21]

18 As we would expect, the average gross AER interest rate increases with term length.

19 We select 2017 to allow the post review entrants time to establish themselves in the market to the extent that we have enough data to make valid comparisons against incumbents and pre-review entrants

20 While the cost of funding differs across banks, we are unable to control for this.

21 Ashton, J.K. and Hudson, R.S., 2008. Interest rate clustering in UK financial services markets. Journal of Banking & Finance, 32(7), pp.1393-1403.

Our POLS model is specified as follows

$$y_{ijt} = \sum_{t=1}^{T} \beta d_t + \sum_{j=1}^{J} \gamma f_j + \sum_{k=1}^{K} \delta p_k + Term_i + Duration_i + \varepsilon_{ijt}$$

- y_{ijt} represents Gross AER interest rate for a given fixed cash savings product 'i', from firm 'j', in time period 't'
- α_{ijt} represents a constant
- d_t is a dummy variable which is 1 in time period 't' and zero otherwise
- f_j is a dummy variable which is 1 for firm 'j' and zero otherwise
- p_k is a dummy variable which is 1 for product category 'k' and zero otherwise
- $Term_i$ is the term length for product 'i'
- $Duration_i$ is the duration product 'i' is on the market
- ε_{ijt} is the error term

Results

The results of these comparisons (shown in Table 6) seem to suggest a statistically significant (at the 1% level) higher interest rate offered by the post-review entrants on fixed term savings products. Controlling for quarter, firm, and product-specific effects suggests that on average the post-review entrants offer fixed rate products with a one percentage point higher interest rate.

Table 6: Moneyfacts Pooled OLS Regression Analysis

Model Description:	Baseline	Term Length & Duration	Firm Specific Effects	Firm Specific Effects (cluster)	Time effects	Time & Fixed Effects	Time & Fixed Effects (cluster)	Product Categories	Product Categories (Cluster)
Post-review entrant	0.379***	0.308***	0.518***	0.518***	0.300***	0.685***	0.685***	0.980***	0.980***
	-36.12	-35.57	-22.9	-9.49	-38.96	-33.36	-9.68	-13.96	-4.38
Term length	-	0.000580***	0.000480***	0.000480***	0.000583***	0.000482***	0.000482***	0.000472***	0.000472***
		-71.35	-79.86	-16.09	-74.26	-95.95	-18.12	-96.1	-18.66
Duration on market	-	-0.103***	-0.0789***	-0.0789***	-0.0460***	-0.0664***	-0.0664***	-0.0588***	-0.0588***
		(-12.03)	(-14.63)	(-6.52)	(-5.84)	(-13.97)	(-4.48)	(-13.81)	(-4.58)
Constant	1.497***	1.127***	0.719***	0.719***	1.173***	0.772***	0.772***	0.439***	0.439**
	-189.79	-68.91	-26.81	-12.28	-84.96	-30.81	-8.62	-8.09	-2.47
Observations	6584	6584	6584	6584	6584	6584	6584	6584	6584
Time Dummies	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Firm Dummies	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes
Product Dummies	No	No	No	No	No	No	No	Yes	Yes
R squared	0.0912	0.470	0.828	0.828	0.545	0.890	0.890	0.897	0.897
F	1304.4	2014.9	.	.	1406.4	.	.	.	.

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: FCA Analysis of Moneyfacts data

Robustness

We may be concerned about the presence of heteroscedasticity and serial correlation in our models' error terms, as the presence of either would cause our standard errors and associated statistical significance levels to be incorrect.[22]

To avoid this issue all models are estimated using heteroscedasticity-robust errors. To account for the potential for firm-specific unobserved effects, we also re-estimate models with cluster-robust standard errors, clustering at the firm level as an alternative method of accounting for potential error structures. In all cases adjusting from heteroskedasticity-robust standard errors to cluster-robust standard errors has no effect on the significance of our results.

Given the lack of time-variant dependent variables and the unbalanced nature of our panel, we are unable to use statistical methods (such as a Wooldridge test) to check for the presence of error serial correlation. Given the structure of our panel, the presence of serial correlation, should it exist, is likely to be a minor issue. In particular, given that we are focusing on 4 time periods with a large number of products (over 1,600), we are appealing to asymptotic properties through a large number rather than a large time period, and as such, our inability to formally test for serial correlation should not significantly affect our statistical inference.

Limitations

The results of our regression analysis should not be interpreted as caused by the intervention. That is because post-review entrants may have entered irrespective of the intervention, and because, conditional on entering, their deposit-taking activities may have been unaffected by the intervention.

The purpose of this piece of analysis is to examine the difference in product offerings by entrants versus incumbents, given that these firms have entered the market. As such we do not interpret our results as causality, but rather as a statistical analysis of correlations.

Conclusion

The analysis shows that when accounting for factors such as the term length of the product, the duration for which the product is on the market, which firm is offering the product, and the specific time periods in which a product is offered for, post-review entrants offer significantly higher interest rates (between 30 to 100 basis points) on fixed term savings products. This is consistent with our original hypothesis that the post-review entrants appear to offer consumers a higher interest rate on fixed term cash savings products to attract customers.

22 Note that heteroscedasticity and serial correlation may render our estimates inefficient. In the case of serial correlation there may also be the possibility of introducing bias due to contemporaneous correlation with the error term. We discuss this issue in greater detail below.

Section 4: Mortgages Interest Rates Pooled OLS Analysis

Background

A subset of post-review entrants have business models that involve the origination of regulated mortgage contracts. This annex details the analysis undertaken to compare these mortgages to those of incumbent banks and pre-review entrants.

In the following, we describe the types of mortgage products that post-review entrants have focused on, and examine the pricing of post-review entrants' mortgages, relative to that of incumbent banks and pre-review entrants.

Hypothesis

As a means to attracting customers and gaining market share, we expect post-review entrants to target more specialised markets rather than compete on standard mortgages with incumbents. We would also expect post-review entrants to offer lower charges on mortgages relative to incumbents and older entrants.

Data and Dataset Construction

The dataset is constructed from Mortgages Product Sales Data (PSD001), based on regulatory returns that all home finance providers are required to submit to the FCA. These returns provide transaction-level data on all originations of regulated mortgage contracts[23] in the UK, and includes borrower, property and product characteristics.[24]

We extracted 2017 data covering all banks that engaged in regulated mortgage lending, including pre- and post-review entrants. For each transaction, the extract includes information on a range of loan features, borrower characteristics, and property characteristics.[25] Comparing loans that were originated over the same time period allows us to account for wider market conditions as a driver of differences in product features or pricing. The extracted dataset does not contain transactions by non-bank home finance providers. The primary reason for limiting our sample is that we wish to perform a like-for-like comparison of the different groups of banks.

23 The FCA handbook defines a regulated mortgage contract as a contract which, at the time it is entered into, satisfies the following conditions:

(1) the contract is one where a lender provides credit to an individual or trustees (the 'borrower');
(2) the contract provides for the obligation of the borrower to repay to be secured by a mortgage on land in the EEA; and
(3) at least 40% of that land is used, or is intended to be used, as or in connection with a dwelling.

24 PSD001 captures those porting their mortgage to a new property and those switching to a new lender (external switchers). It does not capture advances and product transfers with the same lender (an internal switch).

25 A full list of data items is available in SUP 16 Annex 21, FCA Handbook.

The extract includes about 760,000 transactions in total, of which 8,000 are by pre-review entrants, and 12,000 are by post-review entrants. That is, post-review entrants originated approximately 1.6% of all banks' mortgages. All transactions amounted to around £150bn in 2017, of which £2.2bn are by pre-review entrants, and £2.3bn are by post-review entrants.

Borrower and product features

Across incumbents, pre-review entrants, and post-review entrants, there is variety in the types of mortgage lending that banks have focused on. In what follows we present selected summary statistics of loans by the different groups.

Compared to both incumbents and pre-review entrants, many loans by post-review entrants had less common features. In particular, 40% of post-review entrants' loans had LIBOR tracker reversion rates, which are generally associated with more specialist lending. As shown in Table 7, loan types that were more frequent among post-review entrants' mortgages include products involving debt consolidation, second charges, government support, shared equity, interest roll-up, and bridging loans.

Table 7: Frequency of selected product features for each group (% of loans with feature for each group)[26]

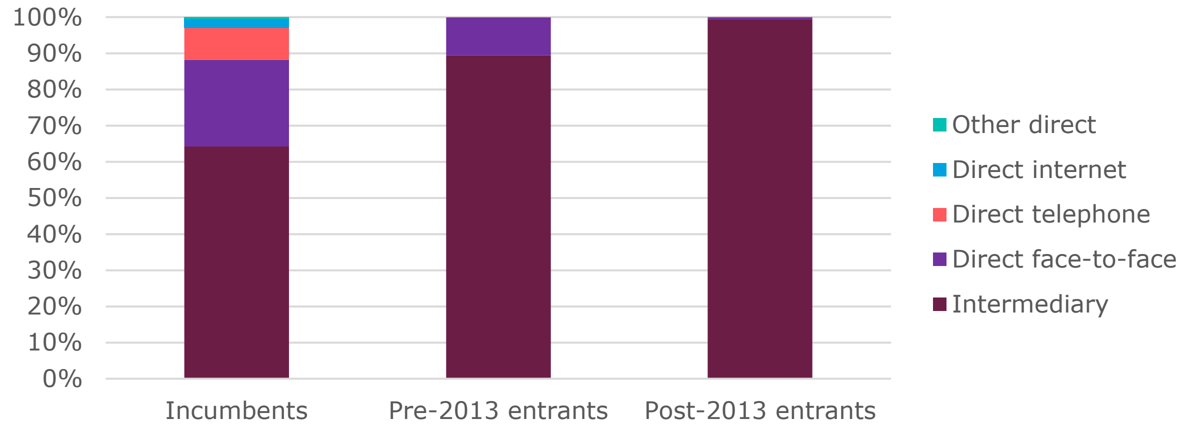
	Post-review entrants	Pre-review entrants	Incumbents
Extra money withdrawn for debt consolidation	18%	3%	7%
Second charge mortgage	13%	5%	1%
Mortgage advanced under a government supported initiative	11%	0%	5%
Shared equity mortgage	11%	0%	4%
Interest roll-up mortgage	10%	0%	0%
Bridging loan	9%	0%	0%

Source: FCA PSD001.

As shown in Figure 8, post-review entrants relied almost exclusively on intermediaries as a sales channel. This is in contrast to incumbents and, to a lesser degree, pre-review entrants, who sold a share of loans directly to consumers.

26 Further details can be found in SUP 16 Annex 21, FCA Handbook. Although it should be noted that Features are neither mutually exclusive nor collectively exhaustive.

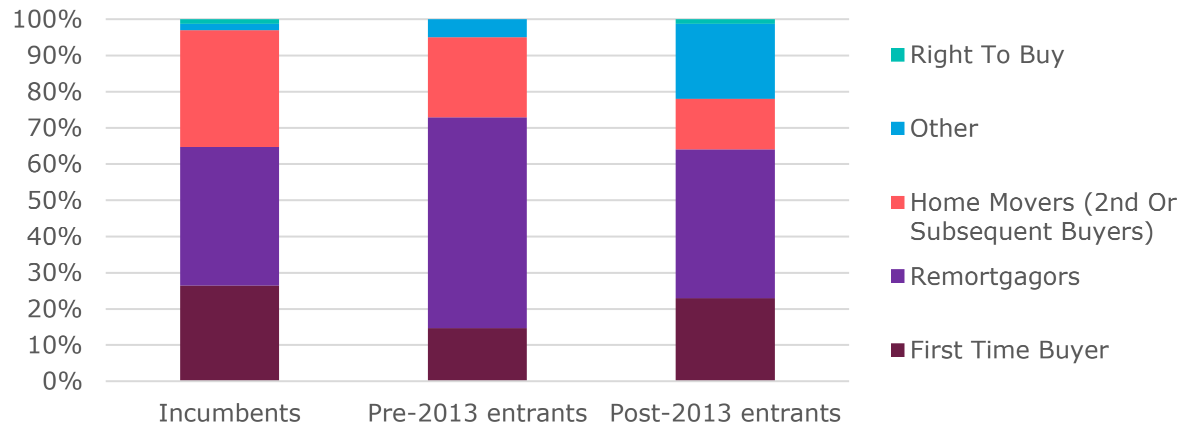
Figure 8: Breakdown of loans by sales channel for each group



Source: FCA PSD001

The distribution of borrower types of post-review entrants has been broadly similar to that of incumbents, with the exception of a stronger focus on 'other' borrower types[27], and less focus on home movers. In contrast to these groups, pre-review entrants focused more heavily on re-mortgagors.

Figure 9: Breakdown of loans by borrower type for each group

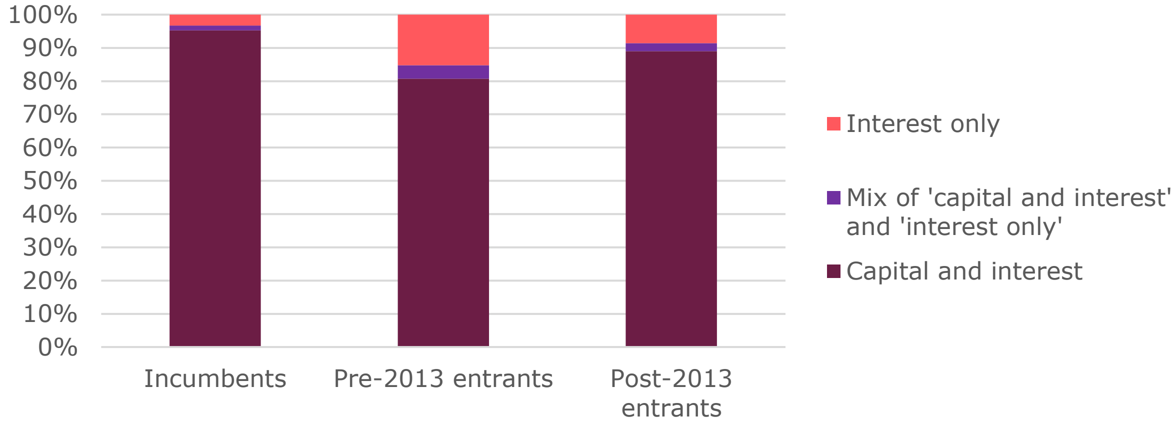


Source: FCA PSD001

With regards to the repayment type, the share of loans by post-review entrants that do not involve full capital repayment is 11%, which is noticeably less than that of pre-review entrants (19%). This is compared to less than 5% by incumbent banks.

27 'Other' is comprised of lifetime mortgages, bridging loans and second charge regulated mortgage contracts that are not for re-mortgage purposes.

Figure 10: Breakdown of loans by repayment type for each group



Source: FCA PSD001

Pricing analysis sample

In what follows, we seek to compare the pricing of regulated mortgages across the three groups of banks. Given our choice of price measure, which is discussed in the Methodology section, we focus on 2-year and 5-year fixed rate contracts. To ensure robust estimation of the regression, we exclude observations with unusual values according to the criteria in Table 8.

Table 8: Overview of criteria for identifying transactions with unusual characteristics

Variable	Excluded if below:	Excluded if above:
Gross income	£0	£500,000
Loan value	£1,000	£5,000,000
Property value	£10,000	£5,000,000
LTV	1	120
LTI	0.1	10
Initial interest rate	0.25%	20%
Borrower age	18	100
Term	2	50
Lender fees	0	£10,000
Intermediary fees	0	£10,000
Fees added to loan	0	£10,000

Source: FCA PSD001

We also exclude loans with LIBOR tracker reversion rates and regulated second charge mortgages. While these types of mortgages make up a significant part of post-review entrants' lending, the limited number of such loans by incumbents and pre-review entrants prevents a robust price comparison for this subset of mortgages.

The table below shows the number of observations that remain in the 2-year and 5-year samples, for each group of banks, compared to the total number of mortgages. The figures in Table 9 imply that our pricing analysis covers around one third of all mortgages by post-review entrants.

Table 9: Number of observations by sample[28]

	2-year fixed sample (cleaned)	5-year fixed sample (cleaned)	Total
Observations	412,927	223,004	763,240
of which pre-review entrants	1,399	5,520	8,340
of which post-review entrants	849	3,332	12,257

Source: FCA analysis of PSD001 data

Methodology

Price measure

Following the approach in Belgibayeva and Majer (2018)[29], we compute a price measure for 2 and 5 year fixed mortgages, based on the Annual Percentage Rate of Charge (APRC).[30] The reason for focusing on the subset of fixed-rate mortgages is that it allows for a close approximation of the actual interest due during the fixed-rate period. Such an approximation is infeasible for variable-rate mortgages in the absence of additional information. That is because the initial interest rate recorded in PSD001 would apply only until the first change of interest rate, the magnitude and timing of which would differ across products and lenders.

The price measure is computed under the assumption that any lender and intermediary fees are rolled-up into the loans.[31] Similarly to Belgibayeva and Majer, it is also assumed that borrowers switch to another product with the same lender or redeem their mortgage at the end of the fixed-rate period, to be able to compare mortgages within a common

28 The number of observations in the two samples does not add up to the total for two reasons. Firstly, some contracts have interest rates that are not fixed, or fixed for a period other than two or five years. Secondly, some observations were removed according to the criteria presented above.

29 Belgibayeva, A. & Majer, T. (2018). 'Six of One…? Choice of intermediary in the UK mortgage market'. *FCA Occasional Paper 35*

30 The general formula for calculating the APRC is defined in MCOB 10.3 in the FCA Handbook, which we adapt to our assumptions and the available data.

31 In contrast to Belgibayeva and Majer, we include intermediary fees in the computation of our cost measure. This is done to account for the full cost of borrowing from post-review entrants, which sell almost all of their loans through intermediaries.

scenario of consumer behaviour. Under these assumptions, the constant monthly payment during the fixed-rate period is computed as follows:

$$P_i = P = \frac{r(A+f)}{1 - \frac{1}{(1+r)^T}}$$

where

- P_i is the payment in month i
- A is the amount of the initial advance
- f is the sum of lender and intermediary fees
- r is the monthly initial interest rate
- T is the mortgage term length in months

Based on these monthly payments, our price measure is implicitly defined as follows:

$$A = \sum_{i=1}^{N} \left(\frac{P_i}{(1+X)^i} \right) + \frac{B_N}{(1+X)^N}$$

where

- N is the length of the fixed-rate period in months
- B_N is the outstanding balance at the end of the fixed-rate period, according to the amortisation schedule
- X is the monthly price measure

The monthly price measure can then be converted to the estimated annual percentage rate of charge $APRC$ as follows:

$$APRC = (1+X)^{12} - 1$$

The resulting annual price measure differs across firm and product groups, as shown in the table below. For two-year fixed loans, post-review entrants have lower estimated APRCs, while pre-review entrants have significantly higher APRCs. For five-year fixed loans, both entrant groups have lower APRCs than incumbents, with post-review entrants having the lowest APRCs.

Table 10: Estimated APRC, by group[32]

		Incumbents	Pre-review entrants	Post-review entrants
Two-year fixed	Mean	2.24%	3.65%	1.95%
	Median	1.96%	2.79%	1.71%
Five-year fixed	Mean	2.44%	2.17%	1.88%
	Median	2.28%	2.03%	1.82%

Source: FCA analysis of PSD001 data

While the statistics above allow us to compare the unconditional pricing of mortgages, they do not account for variation in product and borrower characteristics across incumbents, pre-review entrants, and post-review entrants. The observed differences

[32] As described in the data section, we exclude loans with LIBOR reversion rates and second charge mortgages from the pricing analysis samples. This is done because the limited number of such loans by incumbents and pre-review entrants prevents a robust price comparison for this subset of mortgages, which tend to have significantly higher APRCs than the figures presented here.

may be driven by such variation in characteristics. In what follows, we compare mortgage pricing across these groups, conditional on the observable characteristics of the loans.

Regression specification

Our approach is to approximate the conditional expectation function of the cost of borrowing with linear regression models, using indicator variables to estimate pricing differences between groups.

To compare the cost of borrowing across groups, conditional on observable characteristics, we estimate three regression specifications.

In the first two specifications, the price measure is regressed directly on observable characteristics recorded in PSD001 data. For the third specification, we generate all possible interaction effects among these observable characteristics[33], and then reduce the dimensionality of these using Principal Component Analysis (PCA).[34]

The advantage of using a full set of interaction terms, as done in the third specification, is that it allows for more complex relationships between characteristics. For example, the relationship between the LTV of a mortgage and its price may depend on whether the borrower has an impaired credit history.

The first model is specified as follows:

$$Price_{ilt} = \alpha + \beta NE_l + \gamma OE_l + \delta X_i + \theta Y_t + \epsilon_{ilt} \qquad \text{(Model 1)}$$

where

- $Price_{ilt}$ is the price of mortgage i, by lender l, in month t
- NE_l is an indicator for lender l being a post-review entrant
- OE_l is an indicator for lender l being a pre-review entrant
- X_i is a set of loan and borrower characteristics[35]
- Y_t is a set of monthly controls.
- ϵ_{ilt} is an unobserved disturbance term

The parameters α, β, γ, δ, and θ are estimated with Ordinary Least Squares (OLS), with the parameters β and γ being the parameters of interest.

The second model specification is similar to the first specification, but includes a set of outward postcode controls Z_i as well:

33 The interaction terms do not include outward postcode indicators to keep the matrix dimensions manageable. Based on the results of the first two models, the postcode indicators seem to make little difference once the other characteristics have been controlled for.

34 PCA is detailed in Jolliffe, I. T. (2002). *Principal Component Analysis, Second Edition.* Springer: New York.

35 In particular, the following characteristics and their squares are included as controls: borrowers' gross income, property value, loan value, loan-to-value (LTV), loan-to-income (LTI), mortgage term, and borrower age. Moreover, dummy variables for the following features and mortgage types are included as controls: joint income basis, reversion rate type, borrower type, dwelling type, employment status, sales channel, borrower with impaired credit history, type of income verification, interest-only mortgage, partial capital repayment mortgage, shared ownership mortgage, high net worth mortgage, cashback mortgage, flexible features mortgage, government-supported mortgage, interest roll-up, shared equity mortgage, buy-to-let mortgage, pension mortgage, lifetime mortgage, shared appreciation mortgage, business loan mortgage, guarantor mortgage, low start mortgage, self-build mortgage, secured overdraft mortgage, contract variation mortgage, indemnity insurance mortgage, endowment mortgage, offset mortgage, savings or investment mortgage, extra money raised for debt consolidation, extra money raised for home improvement.

$$Price_{ilt} = \alpha + \beta NE_l + \gamma OE_l + \delta X_i + \theta Y_t + \mu Z_i + \epsilon_{ilt} \quad \text{(Model 2)}$$

Analogously to the first specification, the parameters α, β, γ, δ, θ and μ are estimated with OLS, with the parameters β and γ being the parameters of interest.

The third specification is a principal components regression.[36] Firstly, a full set of interaction effects between loan and borrower characteristics is generated. Using these interactions and the original characteristics as regressors is infeasible due to multicollinearity. To avoid multicollinearity of this large set of features, we use principal components analysis (PCA), a method for dimensionality reduction, such that the principal components capture more than 90% of the variance of the full set of interactions. Finally, these principal components PC_{it} are used as controls alongside the entrant group indicators as follows:

$$Price_{ilt} = \alpha + \beta NE_l + \gamma OE_l + \delta PC_{it} + \epsilon_{ilt} \quad \text{(Model 3)}$$

Results

Regression estimates

The following table presents the results of estimating the different specifications for the 2-year and 5-year fixed samples. All specifications are estimated with standard errors clustered at the level of the lender l.[37] [38]

36 Principal component regression is detailed in chapter 8 of Jolliffe, I. T. (2002). *Principal Component Analysis.*

37 Abadie, Athey, Imbens, and Wooldridge examine when to use clustered standard errors. In the case at hand, using their terminology, the assignment process of whether a mortgage is 'treated' (ie originated by a bank that was subject to the new policy) is clustered at the bank level. See Abadie, A. et. al. (2017). 'When Should You Adjust Standard Errors for Clustering?'. *NBER Working Paper No. 24003.*

38 ~~Results of using heteroskedasticity-consistent standard errors are reported in the robustness section.~~

Table 11: Results of estimating each specification for both samples

Sample :	2-year fixed			*5-year fixed*		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Post-review entrants	-0.28	-0.28	-0.23	-0.45***	-0.44***	-0.46***
Standard error (clustered)	(0.24)	(0.24)	(0.18)	(0.07)	(0.07)	(0.06)
Pre-review entrants	1.50*	1.50*	1.24**	-0.06	-0.05	-0.01
Standard error (clustered)	(0.85)	(0.85)	(0.63)	(0.09)	(0.08)	(0.05)
Loan features	Yes	Yes	Yes (PCA)	Yes	Yes	Yes (PCA)
Monthly controls	Yes	Yes	Yes (PCA)	Yes	Yes	Yes (PCA)
Postcode controls	No	Yes	No	No	Yes	No
Interactions	No	No	Yes (PCA)	No	No	Yes (PCA)
Adj. R-squared	0.67	0.68	0.67	0.70	0.71	0.72
Observations	415,175	415,175	415,175	231,856	231,856	231,856

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: FCA analysis of PSD001 data

Based on the results of estimating the specifications for each sample, mortgages originated by post-review entrants appear to have lower APRC than incumbents, when controlling for observable characteristics. In particular, the central estimate suggests that post-review entrants’ two-year fixed loans have an estimated APRC that is between 23 and 28 basis points lower than that of incumbents’ loans, given the loans’ observable characteristics. This difference is not statistically significant when using clustered standard errors. However, in this case clustering at the lender level results in a small number of clusters, reducing the statistical power of the hypothesis test. As shown in the robustness section, two-year estimates are significant when instead using heteroskedasiticty-consistent standard errors[39], and significant at a 10% level when excluding outliers.

For five-year fixed loans by post-review entrants, this difference is between 44 and 46 basis points. These estimates are highly significant with clustered standard errors.

39 Due to the behaviour of clustered standard errors when the number of clusters is very small, we present variance estimates based on this alternative estimator for transparency.

The APRCs of pre-review entrants' two-year fixed loans are higher than the other groups, when controlling for observed characteristics. Five-year fixed mortgages by pre-review entrants appear to have a marginally lower APRC than incumbents' mortgages with the same characteristics. However, the magnitude of the estimate decreases as the number of controls increases, and the estimate is not statistically significant.

Robustness

In this section, we present the results of two robustness checks of the regression results reported above. Firstly, we investigate whether statistical significance of parameter estimates changes when using heteroskedasticity-consistent standard errors [40] instead of clustered standard errors. Secondly, while some observations with extreme values have been excluded according to the criteria listed above, it is still possible that individual observations have a strong influence on the parameter estimates. To investigate the extent to which this is the case, we use a metric called Cook's distance[41] to identify and exclude such observations. In particular, we exclude any observations with a distance greater than 1000/N, where N is the total number of observations.[42]

Table 12: Results of robustness checks

Sample:	2 year fixed			5 year fixed		
	Baseline	HC	Excl. outliers	Baseline	HC	Excl. outliers
Post-review entrants	-0.23	-0.23***	-0.26*	-0.46***	-0.46***	-0.47***
	(0.18)	(0.03)	(0.16)	(0.06)	(0.01)	(0.06)
Pre-review entrants	1.24**	1.24***	1.26**	-0.01	-0.01	0.01
	(0.63)	(0.03)	(0.63)	(0.05)	(0.01)	(0.05)
Adj. R-squared	0.67	0.67	0.67	0.72	0.72	0.72
Observations	415,175	415,175	415,038	231,856	231,856	231,751

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: FCA analysis of PSD data. Note: The 'Baseline' column shows the estimates of estimating Model 3. The 'HC' column shows the same estimates when using heteroskedasticity-consistent standard errors to test for statistical significance. The 'Excl. outliers' column shows the results of estimating Model 3 after excluding observations identified to be outliers using Cook's distance.

The robustness checks show that the estimated difference in ARPCs between post-review entrants and incumbents becomes highly significant when using heteroskedasticity-robust standard errors, rather than clustered standard errors. With regards to outliers,

40 See White, H. (1980). 'A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity'. *Econometrica.*

41 See Cook, D. R. (1979). 'Influential Observations in Linear Regression'. *Journal of the American Statistical Association.*

42 A distance cutoff of 1 has been suggested for large N, see: Cook, R. D. & Weisberg, S. (1982). Residuals and Influence in Regression. New York, NY: Chapman & Hall. For the purposes of checking robustness conservatively, we set a lower threshold, in order to exclude a larger number of potentially influential observations.

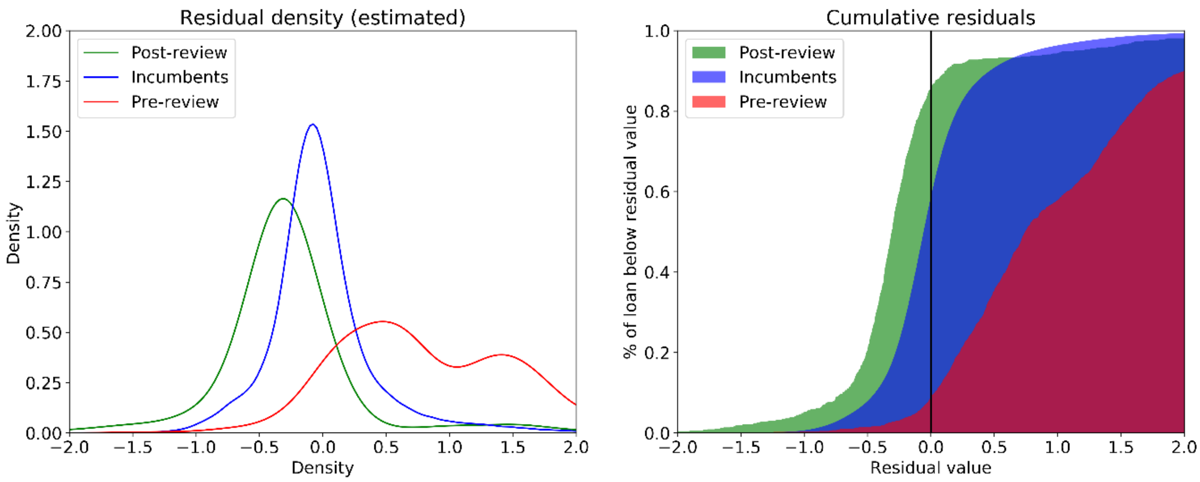
the parameter estimates for both samples are only marginally affected by the exclusion of influential observations.

Price distribution across groups

The estimation results above can be visualised by estimating the regressions without dummy variables for the different groups. The residuals of this regression can then be plotted to compare the groups' distributions of APRC when controlling for loan and borrower characteristics. This also allows to examine each groups' share of loans that have a lower ARPC than observable characteristics suggest.

Figure 11 shows these results for two-year fixed mortgages, reflecting the estimation results. The left chart shows that the distribution of **the regression residuals** post-review entrants is shifted towards lower APRCs than incumbents, and the distribution of pre-review entrants is wider and shifted towards higher APRCs. The right chart of Figure 4 shows the same distributions cumulatively. Nearly 90% of post-review entrants' loans have lower APRCs than observable characteristics suggest.

Figure 11: Distributions of regression residuals by group (2-year fixed)[43]

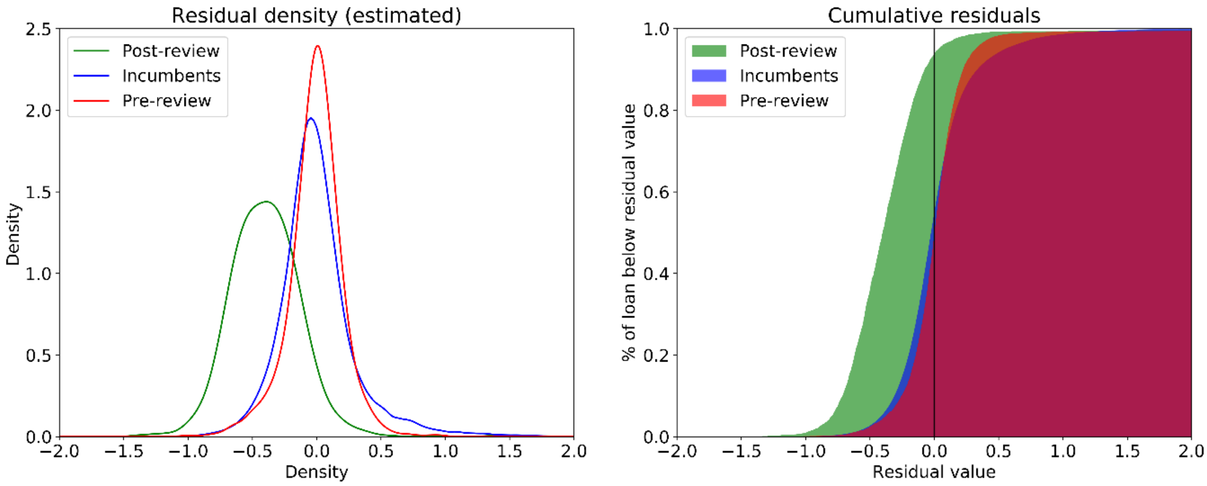


Source: FCA analysis of PSD001 data. Note: The left panel shows the distribution of the regression residuals for each group. The right panel shows the cumulative distribution of regression residuals by group.

Figure 12 shows these distributions for the sample of five-year fixed mortgages. The difference in distributions in the left panel shows that, in line with the result of the regression estimates, post-review entrants' loans appear to have lower APRCs when controlling for observable factors, and pre-review entrants' loans appear similar to incumbents' loans. The right panel shows that around 90% of post-review loans have lower APRCs than observables suggest.

[43] For figure 13 and 14 the left panel shows the distribution of the regression residuals for each group. The right panel shows the cumulative distribution of regression residuals by group.

Figure 12: Distributions of regression residuals by group (5-year fixed)



Source: FCA analysis of PSD data. Note: The left panel shows the distribution of the regression residuals for each group. The right panel shows the cumulative distribution of regression residuals by group.

Quantifying difference in interest payments

In order to quantify the estimated difference in interest payments for a typical loan, we calculate the amortisation schedule of a loan with a loan value, loan term, and APRC equal to the average of loans by post-review entrants. We then calculate a similar amortisation schedule with an APRC that is increased by 23 basis points for two-year fixed loans and 46 basis points for five-year fixed loans.[44] Finally, we evaluate the difference of interest repayments during the first year of this approximately representative mortgage.[45]

Using this methodology for those post-review entrants' loans that are in the pricing samples, we estimate savings to consumers based on observable characteristics. The evidence suggests that consumers that took out loans from post-review entrants in the year 2017 saved around £3m in interest payments during the first year of their mortgage.[46] These savings are relative to the interest they would have paid had they taken out a mortgage with the same characteristics from an incumbent. Estimated reductions in interest payments differ by mortgage term, with an average first-year reduction of £480 per loan for a two-year fixed mortgage, and £940 for a five-year fixed mortgage.

[44] These values are chosen to be equal to the lowest estimated difference to incumbents in the regression approach presented above.

[45] This methodology only approximates the difference in interest payments, due to interest payments depending on loan value, loan term, and interest rate in a non-linear fashion.

[46] This finding is based on the analysis described here, but it is subject to the issues discussed in the Limitations section. Instead of using the central estimate of the regression analysis, this calculation could be carried out for the bounds of the confidence interval around the regression estimate. While the exact bounds of this interval depend on the choice of variance estimator, the interest difference would remain positive and in the region of £3m.

Limitations

As it is not clear to what extent the mortgage lending activities of each entrant may have been affected by the intervention, these results cannot be interpreted as caused by the intervention.[47]

With regards to the pricing of post-review entrants' loans, a potential for selection effects exists to the extent that there are unobserved factors driving differences in APRCs. That is, the borrowers that chose to borrow from post-review entrants may have unobserved characteristics that account for the estimated price difference. For example, the dataset does not contain information on credit scores, though it includes features that are likely to be correlated with credit scores.

Moreover, as discussed above, our price measure is based on assumptions about consumer behaviour, whether or not fees are rolled up into the loan, and about consumer choice after the end of the fixed-rate period

Finally, our analysis does not investigate whether entry has led to changes in the pricing of incumbents. That is, given that entrants have originated only a small share of all banks' loans[48], it is entirely possible that post-review entrants have offered competitively priced products without affecting the wider market conditions noticeably.

Conclusion

Examining the types of mortgages that post-review entrants have originated in 2017, we find that these firms have offered a wide range of products, with more than 40% of their loans having features that are generally associated with specialist lending. Post-review entrants relied almost exclusively on intermediaries as a sales channel, in contrast to incumbents and pre-review entrants who sold a noticeable share of their mortgages directly.

Our pricing analysis focuses on the subset of loans that have two-year and five-year fixed rates, excluding specialist loans. We find that such loans by post-review entrants have lower interest rates than both incumbents and pre-review entrants, when controlling for product and borrower characteristics. We find that two-year fixed rate loans by post-review entrants have APRCs that are, on average, around 25 basis points lower than that of incumbents' loans, given the loans' observable characteristics. Analogously, for five-year fixed rate loans, our analysis suggests post-review entrants' APRCs tend to be around 45 basis points lower on average.

Using the estimates above, we estimate the extent to which borrowers of post-review entrants benefitted during the first year of their mortgage. For the loans analysed, the evidence suggests that borrowers from post-review entrants paid around £3m less in interest during the first year of their loans than they would have done if they had taken the same loan with an incumbent.

47 A special case of this are firms that were mortgage lenders prior to becoming banks. We observe that the mortgage lending volume of these firms increased significantly after becoming authorised. Benefits to customers of these firms can be attributed to the intervention to the extent that the review has facilitated this expansion. It should be noted that most loans of these firms are not in the pricing analysis sample, so the pricing results are robust to the exclusion of these firms.

48 Entrants' share is smaller still when considering the mortgages by non-bank lenders.

12 Endeavour Square, London, E20 1JN
Telephone: +44 (0)20 7066 1000 Website: www.fca.org.uk