

# RETAIL BANKING MARKET INVESTIGATION

## Proposed econometric analysis of the determinants of searching and switching for personal current accounts

### *Introduction*

1. In this paper we describe an econometric analysis we are proposing to carry out to study the factors that drive some customers' decisions to search and switch personal current accounts (PCAs), while the other customers remain inactive. This analysis will inform part of our understanding of customer behaviour in relation to PCAs.
2. This paper is structured as follows. In the next section, we present a general discussion of the aims of the model and the information we expect to obtain from the analysis. We then explain the proposed methodology, including a description of the econometric model, the data and econometric methods we are planning to use to estimate it. The limitations of the analysis and how we plan to address them are discussed in the relevant sections of the paper.
3. If parties have comments, please send them to [retailbanking@cma.gsi.gov.uk](mailto:retailbanking@cma.gsi.gov.uk) or write to the team at the address below. **The deadline for comments is midday Friday 20 March 2015.**

Retail banking team  
Competition and Markets Authority  
Victoria House  
Southampton Row  
London  
WC1B 4AD

### *Overview of the method and key questions to address*

4. The objective of this analysis is to study the factors that lead some customers to search and switch PCAs, while the majority of customers stay inactive.

5. The basis of the analysis is a comparison between searchers and non-searchers, and switchers and non-switchers, with the aim of understanding the relevant differences between these different groups of customers.
6. This comparison should allow us to infer what are the main drivers and deterrents to customers' engagement in relation to PCAs. So, for example, we will be able to observe if PCA customers who use overdrafts more frequently are more or less likely to search and switch. But importantly, the use of regression analysis will allow us to evaluate this while 'holding fixed' other factors, such as age, level of education, etc.
7. We have identified a number of potential drivers of and deterrents to customers' decisions to search and switch. These can be grouped in the following categories:
  - (a) **Monetary and quality gains.** Customers that search and switch to a new PCA may do so because they expect to get a better deal as a result. This includes not only monetary gains, such as reduced overdraft charges or higher interest rates, but also quality considerations, such as better customer service or access to a local branch.
  - (b) **Push factors.** Customers might engage with the market following certain 'events' that change their relationship with their bank. These 'push factors' could include changes to the services offered by the bank, such as the closure of a local branch or an increase in overdraft charges. They could also include 'life events' or changes in customers' personal circumstances that may lead them to demand different services from their PCA or change the type of PCA they have. For example, a couple that starts living together may decide to use a joint account and, as a result, look around for a new one. Therefore, we might expect customers that have experienced these events to be more likely to search and switch.
  - (c) **Costs of searching and switching.** Customers might be expected to decide whether to search and switch by comparing their expected gains to their expectation regarding how costly the process is going to be for them.<sup>1</sup> Customers may not find it convenient to allocate time to shopping around if the perceived costs significantly outweigh any benefits from

---

<sup>1</sup> See, for example: Paul Klemperer (2005), *Network Effects and Switching Costs: two short essays for the new New Palgrave*, in 'The New Palgrave Dictionary of Economics', S. N. Durlauf and L. E. Blume (ed.), Palgrave Macmillan; and Catherine Waddams Price, Catherine Webster, Minyan Zhu (2013), *Searching and Switching: Empirical estimates of consumer behaviour in regulated markets*, CCP Working Paper 13-11, Centre for Competition Policy, University of East Anglia.

switching. These costs may concern an expectation around the time it would take or how difficult they perceive the process to be.

(d) **Individual characteristics.** The model will include demographic and other customer characteristics, such as income, age and level of education. Some of these characteristics are likely to be associated with the expected gains and costs of search and switching. For example, someone with a higher level of education or better access to the internet could potentially require less time to identify a good deal and, at the same time, be more likely to find the best available option to them. At the same time, demographic characteristics will allow us to identify whether there are groups of customers that are particularly inactive in the market.

8. A detailed description of how we will measure each of these factors is presented in the data section below.

## **Methodology**

### *Model specification*

9. The decisions to search and switch can be modelled as follows:

$$Pr(Searching_i/X_i) = f(\beta_i'X_i)$$

$$Pr(Switching_i/X_i) = f(\delta_i'X_i)$$

10. The equations above indicate that the probability that a customer  $i$  searches/switches (denoted  $Pr(Searching/Switching)$ ) is a function of a set of drivers and deterrents  $X_i$ .
11. In practice, the model is estimated using a sample of customers that searched/switched and others that have not. The model will allow us then to compare the differences between these two groups, and how these differences contribute to the probability of being among one group of customers or the other.<sup>2</sup>
12. The set of drivers and deterrents  $X_i$  are the factors we discussed above, namely, monetary and quality gains, push factors, cost of searching and switching, and other customer characteristics. In the next section, we present

---

<sup>2</sup> Although there is likely to be overlap between searchers and switchers, we expect to find some switchers that did not search for the new PCA. Therefore, the analysis will not consider searching as a necessary previous step to switching.

a detailed description of the specific variables we intend to include to measure these factors.

#### *Description of variables and data sources*

13. In this section, we present a discussion on how we propose to construct the variables to populate the switching and searching models. The selection of variables to include aims to capture the most important relevant factors associated with customers' decision to search and switch, while at the same time keeping the model tractable.
14. We propose to estimate the model by combining anonymised data from the customer survey with anonymised data on customers' current account usage obtained from banks pursuant to section 174 of the Enterprise Act 2002.

#### *Definition of searchers and switchers*

15. The dependent variables for the two models will be based on customers' responses to relevant survey questions as follows:
  - (a) Searchers are customers that indicate they have looked around for a new PCA in the last 12 months.
  - (b) Switchers are customers that indicate they have switched their main current account in the last 12 months.
16. The survey question concerning switching asks customers to indicate whether they have switched a PCA to another bank or within the same bank. As a starting point we intend to focus the analysis on external switching since we are interested in understanding the drivers of consumers' willingness to shop around and switch between banks. However, we do not rule out extending the analysis to customers that have switched within banks.

#### *Monetary gains*

17. We intend to combine transactional data on customer usage and pricing information to construct a measure of the potential gains from switching for each customer prior to switching. For example, suppose a customer A that has switched from bank X to bank Y uses on average £100 of overdraft per month. We will calculate how much customer A would typically be charged by bank X for that level of overdraft, and compare this to the average and the best offer elsewhere in the market.

18. Information on the bank a customer has switched from will come from the survey. For people that have not switched banks, we will consider the level of charges made by their current bank.
19. To make the analysis comparable between switchers and non-switchers, all pricing information will be taken as of 1 January 2014. We will combine this with information on usage obtained from the PCA transaction data request for the last three months of 2014.
20. We cannot use information on usage for the entire year as, for switchers, current banks will only hold information since they switched. Therefore, year averages would be calculated for a different set of months for switchers and non-switchers, and therefore would be an unsuitable measure for comparing these customer groups. Focusing on the last three months of data will minimise this problem, while at the same time still providing a representative measure of customer usage. For customers that switched after November 2014, we will not have data for all three months. To account for this, we plan to perform a robustness check excluding these customers.
21. In order to keep the model simple, we will focus on the following features of PCAs:
  - (a) Overdraft costs (including interests and charges).
  - (b) Credit interest payments.
  - (c) Cashback.
  - (d) Foreign ATM withdrawals and debit card payments.<sup>3</sup>
22. We will complement this information with survey information regarding customers' level of satisfaction with the charges associated with their PCA, as follows:
  - (a) Switchers: satisfaction with charges of the bank they switched from.
  - (b) Non-switchers: satisfaction with charges of their current bank.
23. The aim of including this latter variable is to account for customers' heterogeneity regarding their perceptions around PCA costs.
24. Our ability to calculate monetary gains as described in the previous paragraphs will depend on data availability. In particular, it relies on whether

---

<sup>3</sup> This will focus only on transactions done in Europe since data for transactions outside Europe is not available.

we can match customer usage with banks' pricing information. If this is not feasible, we will consider alternative measures to proxy for price differences between banks and how to associate them to customer usage. For example, a possibility would be to compare banks' standard PCAs focusing on the four dimensions listed above, and interact with customer usage. This would be a less exact measure of the actual potential gains at a customer level, but would still provide a bank level measure of monetary differences across PCAs offered by different banks.

### *Quality gains*

25. Quality is particularly difficult to measure since there are many dimensions to the quality of service a particular customer may experience. On the basis of parties' comments and in order to keep the model tractable, we have selected three dimensions that appear particularly relevant in customers' decisions to search and switch:
  - (a) Branch offer.
  - (b) Customer service.
  - (c) Bank errors.
26. Regarding the branch offer, we will construct a series of measures capturing local branch services and branch network. We will use a customer's home postcode to define their local area.<sup>4</sup> Within this local area, we will calculate the following indicators of the branch offer:
  - (a) Indicator variable of whether the customer's bank has a branch in the local area.
  - (b) Number of other banks that have branches in the local area.
  - (c) Indicator variable whether the customer's bank has a local branch that opens on weekends or is open long hours during the week.
  - (d) Indicator variable whether at least one other bank has a local branch that opens on weekends or is open long hours during the week.
27. Not all customers may visit a bank close to where they live, but may prefer to visit branches in another location, for example, in the area where they work.

---

<sup>4</sup> We will define local areas as the one mile radius around the customer's home postcode.

To capture this effect we will construct the following indicator of a bank's branch network:

- (a) Number of branches each customer's bank has in the city or region.
  - (b) Average and maximum number of branches of other banks in the city or region.
28. By customer's bank we mean the current bank for non-switchers, while for switchers we will consider the bank they have switched from. As with pricing information, to make the information comparable and relevant to switching decisions during our period of analysis, we will consider all branch information as of 1 January 2014.
29. Not all customers may value the branch offer to the same degree. To account for this, we will experiment interacting these indicators with a survey question asking how important it is for them to have a branch that is conveniently located.
30. For the other two dimensions of quality, customer service and bank errors, we will use information from two survey questions asking customers their level of satisfaction with their bank on:
- (a) staff and customer service; and
  - (b) quality and speed of handling problems.
31. The second question aims to measure the fact that customers may switch following an error by the bank. Parties have pointed out that what matters most for consumers is not whether the bank makes mistakes but how it deals with them. The question has been drafted so as to capture the latter.
32. As with other variables, for non-switchers we consider their level of satisfaction with the current bank, while for switchers we will use information on the level of satisfaction with their previous bank.

*Push factors*

33. We will construct two variables capturing push factors associated with bank changes:
- (a) Closure of a local branch.
  - (b) Changes in prices of a customer's PCA.

34. The information on branch closure will come from the survey where we have asked customers to indicate whether a branch that they frequented closed in the last 12 months. Changes in costs will be calculated using pricing information submitted by the banks. For switchers, these variables will refer to the bank the customer switched from and not their current bank.
35. Another event that we have considered is an error by the bank. This aspect is captured by customer's satisfaction indicator regarding 'handling of problems', as indicated above.
36. As for push factors associated with life events, we have asked respondents to the survey to indicate whether, in the past 12 months, they have:
  - (a) moved house;
  - (b) got married/started living with someone else;
  - (c) got divorced/separated/widowed; or
  - (d) started or stopped working.
37. We will also control for whether the customer has turned 18 in the last 12 months using information on customers' year of birth.

*Cost of searching and switching*

38. Customers' perceptions regarding the cost associated with searching and switching are difficult to measure. We will use a series of different measures to capture the main factors affecting these costs.
39. Information on customers' expectations of the difficulty of searching and switching will come from the customer survey. In particular, we have asked customers that have not searched or switched, how difficult they think the process would be if they were to search/switch. As for searchers/switchers, we asked them about their expectations before they actually searched/switched. These survey questions include the following dimensions of searching and switching:
  - (a) Gathering information on account features and charges.
  - (b) Understanding the options available.
  - (c) Making comparisons between current accounts.
  - (d) The actual process of moving accounts once the decision to switch has been made.



40. However, as we discussed above, survey questions about expectations before searching/switching may be biased. For this reason, we will also include information on customer characteristics that are associated with higher or lower costs of searching and switching. These measures combine information gathered from the survey and bank data, and include:
  - (a) financial literacy (assessed by a survey question which asks respondents to calculate a simple total debt computation);
  - (b) confidence in the use of the internet (survey question);
  - (c) highest level of education achieved (survey question);
  - (d) working status (survey question); and
  - (e) per hour individual income (based on banks' transactional data assuming a 40-hour work week).
41. The first three indicators are associated with the difficulty in assessing and processing information on PCAs, while the last two provide measures of the opportunity cost of time.
42. Additionally, we will also include information on:
  - (a) the number of direct debits and standing orders; and
  - (b) overdraft usage.
43. The intuition behind this is that a customer who has more direct debits or standing orders may perceive switching accounts to be more difficult and time consuming, and hence be less willing to switch. As for overdrafts, a customer who uses overdrafts more frequently may have a level of arranged overdraft with their bank and hence be less likely to switch given that their overdraft limit will not be transferred directly to the new bank.
44. As we discuss further in the last section of this methodology paper, all transaction data will come from a customer's current bank. This presents a problem for switchers, since we will not observe the number of direct debits and standing orders or level of overdraft usage in their bank of origin. However, although these may not always match, there is no reason to think that switching would lead to significant changes in the number of direct debits or standing orders. In spite of this limitation and the unavailability of alternative measures, we consider that this still represents a good measure to account for customers' heterogeneity in the use of PCAs.

### *Other individual characteristics*

45. Additionally to the customer demographic characteristics indicated above, we will also include gender and age. There is not a clear interpretation of the effects of these variables in terms of gains or costs from switching, but previous studies<sup>5</sup> have shown that switching behaviour differs between genders and age groups and, therefore, are important to control for in the analysis.

### *Estimation method*

46. This section discusses our current thinking regarding the specific econometric methods we are going to use to estimate the searching and switching model.
47. Given that the dependent variables in our model are binary (take values zero or one indicating whether or not the customer has searched or switched), as a starting point we will estimate the models using standard binary response models, such as Logit or Probit. These models are typically used when the dependent variable is binary given that a linear model predicts probabilities outside the [0,1] interval. As is standard in the literature, we will also estimate a linear probability model as a robustness check.
48. The fact that the dataset will contain a small fraction of switchers could create some challenges for the analysis. To control for this, we will perform robustness checks using techniques that account for the fact that one of the possible outcomes is rare, such as complementary log-log regression. In spite of this, we expect to have a sample that will contain around 1,000 switchers, which would provide a rich enough dataset for standard binary response models to perform well.
49. A related issue concerns the fact that switchers have been oversampled for the purpose of the survey. We are considering how this sampling may affect the results of the model. A standard way to account for this would be to use weights when estimating the models to account for this. We are exploring this issue and the best way to construct such weights.
50. Related to this is the fact that some customers may not respond to the survey or part of it, and this may not be random but associated with customers' characteristics. We are exploring the possibility of performing an analysis of non-responses that would allow us to control for this in the analysis.

---

<sup>5</sup> For example, Kiser, E. K. (2002). Predicting household switching behaviour and switching costs at depository institutions. *Review of Industrial Organization*, 20(4), 349-365.

51. As a first step, we will model separately the decisions to search and switch. However, for many customers searching is a pre-requisite for switching, and the result of their searching efforts determines whether they switch or not. For this reason, we are also exploring the possibility of estimating a model that links the two. Some of the methods we are currently considering include seemingly unrelated regression in the case of the linear probability model, and using a recursive bivariate probit model.
52. It is likely that effects of different drivers and deterrence may vary between different customer groups. We plan to account for this customer heterogeneity by introducing interactions between the relevant variables in the analysis. An alternative would be to apply a Random Coefficient Model. However, we believe that a simpler approach is better suited for our purposes.