Searching and switching in retail banking

Stefano Callari, Alexander Moore, Laura Rovegno[*]
Competition and Markets Authority[†]

Pasquale Schiraldi
London School of Economics

Abstract

This paper investigates the factors that drive customers to search for, and switch to, a new bank account. The research was conducted as part of the CMA's official market inquiry into the retail banking market, and exploits a unique dataset combining transactions data on retail banking customers with their responses to a detailed telephone survey. We estimate a bivariate recursive probit model in which the decisions to search and to switch are modelled jointly, and the decision to search directly influences the decision to switch. We model the decisions to search and switch as a function of the expected costs – captured by demographic variables-, benefits of engagement and trigger factors. We find that higher education, financial literacy and confidence in the use of internet all significantly increase the probability of search, but results are not significant for switching once we control for searching. The same is true for customers that have experience the closure of a local branch. We also find that customers with high levels of credit balances are more likely to both search and switch, while overdraft users are less likely to switch.

Keywords: Banking; competition; discrete choice modelling

JEL Codes: G21, G28, C35

---

[*] Corresponding author; Competition and Markets Authority. Email: laura.rovegno@cma.gsi.gov.uk
[†] Although the authors worked on certain aspects of the retail banking market investigation at the UK Competition and Markets Authority (CMA), the views and opinions expressed in this paper are the sole responsibility of the authors and do not necessarily reflect those of the CMA or the inquiry group.

1. Introduction

This paper investigates the factors that drive customers to search for, and switch to, a new bank account. The research was conducted as part of the CMA's official market inquiry, and exploits a unique dataset on retail banking customers. In particular, we have transactions data from the accounts of 3,767 PCA customers in 2014, together with their responses to a detailed telephone survey. We know whether each customer searched for, and/or switched to, a new bank account within the last 12 months, and we quantify the importance of both qualitative and quantitative factors on these decisions. We do so using a bivariate recursive probit model in which the two decisions are modelled jointly, and the decision to search directly influences the decision to switch. That is, the decision to switch bank account is a function of both exogenous explanatory variables *and* the endogenous decision to search. This reflects the fact that 75 percent of those that switched account had searched beforehand.

We model the decisions to search and switch as a function of the expected costs and benefits of engagement, as well as a number of "trigger factors" and (real or perceived) barriers to switching. Expected costs are largely captured by demographic variables, and we find that higher education, financial literacy and confidence in the use of internet all significantly increase the probability of searching. We also find that customers with high levels of credit are more likely to search and switch, which is likely to be associated with higher potential monetary gains from switching.[3] We also find that some trigger factors impact on the propensity to search for and switch to a new bank account. In particular, we find that the closure of a local branch increases significantly the probability of searching, while a change in working status increases both the propensity to search and switch. Also, we find that customers who hold their PCA with a bank belonging to one of the largest banking groups are less likely to search and switch.

Our results show that once we control for searching, many of the drivers considered are not significant on the decision to switch, suggesting that the decision to search is an essential factor in the ultimate decision to switch. Moreover, our results show that searching increases the probability of switching by 12 percentage points on average. Even controlling for the decision to search however, there remain barriers to switching. In particular, older customers and overdraft users are significantly less likely to switch. Many heavy overdraft users face genuine uncertainty over the acceptance and timing of an overdraft approval. For lighter overdraft users however, barriers largely stem from informational constraints; overdraft charges are particularly complex to compare across banks due to both the complexity of charging structures and customers' difficulties in understanding their own usage (CMA 2016).

Why are searching and switching so important in the retail banking market? A lack of searching and switching prevents new entrants from expanding organically, reducing the competitive constraint on the dominant incumbents (CMA 2016). A lack of searching and switching by overdraft users for example has resulted in limited price competition on arranged and unarranged overdraft fees (CMA 2016). Further, previous research on household utilities has found that switching alone is not sufficient to reduce prices, or improve quality, for inactive consumers. Waddams Price et al (2013) for example find that "passive customers" are heavily influenced by direct marketing, and so even the switchers in that market tend not to achieve the best market outcomes. As in this paper, it is therefore important to distinguish between searching and switching decisions and analyse the drivers of both.

[3] This is consistent with the CMA's analysis on gains from switching which found that customers with higher credit balances and those who use overdraft present higher potential monetary gains from switching (CMA, 2016).

The paper is structured as follows. Section 2 presents the empirical framework and the rationale for our bivariate recursive probit model. Section 3 discusses the data and choice of variables. Section 4 presents the results and Section 5 presents a discussion and conclusion.

2. Empirical framework

To accurately estimate the determinants of searching and switching decisions, our empirical framework should take into account the fact that the two decisions may be jointly determined. In particular, it is highly likely that for many customers, searching is a prerequisite for switching and the result of their search determines whether or not they will switch. This is reflected in the data: 75 percent of those that switched had initially searched, and those that searched were far more likely to switch – 14 percent of searchers ultimately switched, compared to just 3 percent of the overall population.

To capture this inter-dependency, we model the consumer's decisions to search and switch using a bivariate recursive probit model (Greene 1998), in which:

$$Pr(Search_i|X_i) = f(\beta' X_i) \quad (1)$$
$$Pr(Switch_i|Z_i\ , Search_i) = f(\delta' Z_i, Search_i) \quad (2)$$

where X_i are the variables that influence the decision to search, and Z_i are the variables that influence the decision to switch. (Section 3 provides full details on the variables included in X_i and Z_i.) Our model therefore allows the decision to search to influence the decision to switch, but not vice-versa. This assumption reflects the fact that in reality such decisions are sequential; a consumer typically switches *after* searching.

The full recursive bivariate probit model is given by:

$$S_{1i}^* = \beta' X_i + \varepsilon_{1i} \qquad Search_i = 1\ if\ S_{1i}^* > 0, \qquad 0\ otherwise \quad (3)$$
$$S_{2i}^* = \delta' Z_i + \gamma Search_i + \varepsilon_{2i} \qquad Switch_i = 1\ if\ S_{2i}^* > 0, \qquad 0\ otherwise \quad (4)$$

where the error terms (ε_1 and ε_2) are jointly normally distributed with

$$(\varepsilon_{1i}, \varepsilon_{2i}|X_i, Z_i) \sim N\left(\begin{matrix} E(\varepsilon_{1i}) \\ E(\varepsilon_{2i}) \end{matrix} \ , \ \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right) \quad (5)$$

As demonstrated in Greene (1998), consistent and efficient estimation of the parameters in (3) and (4) can be achieved through maximum likelihood, exactly as in the standard bivariate probit case. That is, for estimation purposes, we can simply "ignore" the simultaneity issue caused by the inclusion of $Search_i$ in the equation for $Switch_i$. The joint probabilities that enter the log-likelihood equation are given by:

$$\begin{aligned} \Pr(Search_i = 1, Switch_i = 1) &= \Phi(\beta' X_i, \delta' Z_i + \gamma\ , \rho) \\ \Pr(Search_i = 1, Switch_i = 0) &= \Phi(\beta' X_i, -\delta' Z_i - \gamma, -\rho\ ) \\ \Pr(Search_i = 0, Switch_i = 1) &= \Phi(-\beta' X_i, \delta' Z_i, -\rho\ ) \\ \Pr(Search_i = 0, Switch_i = 0) &= \Phi(-\beta' X_i, -\delta' Z_i, \rho\ ) \end{aligned} \quad (6)$$

where $\Phi(.)$ denotes the cumulative distribution function of a bivariate normal distribution.

Although the recursive structure of the model does not affect estimation, it does have substantial implications for the calculation and interpretation of marginal effects. Indeed, any variable that is present in both the searching and switching equations now has two effects on the decision to switch – a *direct* effect, and an *indirect* effect via search. The recursive model enables us to estimate both of these effects.[4]

Compared to a univariate model, the bivariate recursive model therefore allows for a deeper understanding of the mechanisms driving a consumer's decision to switch bank account. Consider the impact of a local branch closure for example, a variable that is present in both the searching and switching equations. It is plausible that a branch closure induces certain consumers to search, which in turn increases the probability of switching. In the univariate model for switching, these two effects are conflated, and so the estimated impact of branch closure on switching may be large and significant. The bivariate recursive model, in contrast, disaggregates the two effects: we may find that the *direct* effect of a branch closure on switching is insignificant, but there is a significant *indirect* effect through increased search. The univariate estimates can therefore be misleading, and might lead us to over-estimate the impact of branch closure on switching.

We account for the use of survey data in the empirical specification. For estimation this is straightforward, as we can rely on pre-existing computer routines.[5] The computation of standard errors for the marginal effects is non-standard however. We compute marginal effects as discussed in Dong et al. (2010) and estimate associated standard errors by bootstrapping following Roa and Yue (1988).

3. Data and variables

3.1 Data sources

We exploit a unique database of retail banking customers, combining details on account transactions sourced directly from 13 major banking groups, with the results of a telephone survey conducted by GfK. In particular, as part of the CMA's market investigation, the 13 banking groups were asked to provide a full list of customer accounts, from which GfK selected a stratified random sample of 120,000 accounts.[6] The CMA received detailed transactions data for each of these accounts from the banks, and each account holder was requested by GfK to participate in a detailed telephone survey. The final sample consists of 4,549 customers that participated in the survey; account information is from the final quarter of 2014, and interviews took place between February and March 2015.[7]

3.2 Variables

[4] The mathematical expressions for all the marginal effects are available upon request to the authors.

[5] In particular, we estimate the model using the statistical package Stata and its in-built svy commands. This command allows to account for both the sampling probability and the stratification of the survey sample.

[6] A sample of 120,000 was selected in order to achieve a target of 5,000 telephone interviews. The 13 banking groups consist of the UK's four largest groups (Lloyds, RBS, Barclays, HSBC), plus nine smaller groups (Santander, Nationwide, TSB, Co-op, NAB, Metro, Danske, Allied Irish Bank, Bank of Ireland). In total, the sample includes information on customers from 21 banks belonging to the 13 banking groups.

[7] From the initial sample of 120,000, almost 25,000 were not contacted for interview due to incorrect or missing information. The overall response rate to the survey was 6 percent. Detailed technical details on sample selection and data preparation are provided in the GfK NOP PCA banking survey technical report, available online at https://www.gov.uk/cma-cases/review-of-banking-for-small-and-medium-sized-businesses-smes-in-the-uk

We define $Search_i$ as a dummy variable equal to 1 if a customer responded that they had searched for a new current account in the last 12 months. Analogously, $Switch_i$ is a dummy equal to 1 if a customer responded that they had switched their main current account to a different bank in the last 12 months. We exclude customers that had searched or switched in the previous two to three years, as well as those who had switched accounts within the same bank.[8] The first group is excluded to ensure that information provided during the interview is reliable, and the second group is excluded because it is unclear to what extent "internal switchers" were engaged with the market and took an active decision to change their account.[9] Our reference group – those for whom $Search_i$ and/or $Switch_i$ equals 0 – is therefore the group of customers who have not searched/switched at any point in the last three years, and who have not switched accounts within the same bank.

Our explanatory variables are collected into six groups: customer demographics, use of internet, monetary features, identity of the bank of origin, and trigger factors. We outline the variables here and their expected impact on searching and switching. Table A1 in the Appendix provides a full list of the explanatory variables, together with their definitions. In general, we do not have any *ex ante* reason to constrain any of the variables to influence only searching and not switching, or vice-versa. We therefore include each of the explanatory variables in the equations for both searching and switching. For the system to be well identified however, we require at least one variable to be included only in the searching equation (see Dong et al 2010). To select which variable to drop, we performed a preliminary univariate probit analysis with alternative specifications and tested which drivers had a significant effect on searching and switching overall. We exclude from each equation drivers that appear as not significant across alternative specifications.

Customer demographics consist of gender, income, age, level of education and financial literacy. The primary use of demographic variables is to proxy for the costs of search (Giulietti et al 2005). Customers with a higher level of education and a high degree of financial literacy may find it easier to compare competing alternatives, and identify the best option available to them.

We use three measures related to the use on internet: access and confidence in the use of internet as reported by respondents in our customer survey, and two indicators of whether the customer uses online banking products – internet banking and mobile apps. Access and proficiency in the use of internet is likely to be associated with lower costs of searching and switching. A large amount of information on personal bank accounts is available on line, constituting an easily accessible source to gather information and make comparisons. In addition, many banks also offer the facility to open an account or even switch accounts on their websites. Additionally, customers that regularly interact with their bank using the internet, either through internet banking or mobile apps, may also be more likely to use this tool to search and switch across banks.

For monetary features we use the account transactions data to look at the customer's number of transactions, their use of overdrafts and their credit balance. Monetary features are likely associated with both potential gains from searching and switching, and potential barriers. Customers holding higher credit balances for example will receive higher gains from bank accounts offering better levels of credit interest, which increases the incentive to search and switch. In contrast, those

[8] These exclusions, together with other exclusions due to missing data, reduce are sample to 3,676 customers.

[9] A share of internal users may be customers that engaged with the market and decided that the best product for them was offered by their current bank, and hence switched internally. However, this group is also likely to include customers that were upgraded by their bank or just took on a particular offer they received from their bank without engaging with the market. Given the difficulty interpreting this group, we did not include them in the analysis.

customers with large numbers of transactions may be deterred by the risk of disruption to their account. Even conditional on search, such customers may therefore face (perceived) barriers to switching. Overdraft users may also face barriers to switching. As argued in the introduction, some heavy overdraft users will have a limited choice set as many providers will not provide the required overdraft facility. The complexity of overdraft charges and the difficulty of comparing alternative providers may act as a barrier to lighter overdraft users (CMA 2016).

We also consider the identity of the bank of origin.[10] In particular, we look at whether holding an account with a bank belonging to one of the larger banking groups has an impact on searching and switching behaviour.

Finally, we consider both supply and demand trigger factors. Supply trigger factors are changes in the conditions or services a bank provides that may lead customers to search for alternative options. In particular, we focus on one trigger factor associated with reduced level of service, the closure of a local branch. Demand or individual trigger factors are life events that could lead customers to demand different services from their bank account, or prompt inactive customers to participate in the market. We include moving house, changing work status and changing relationship status as trigger factors that could prompt customers to search or switch. If these life events impact on the services customers demand from their PCA, we would expect searching and switching to be higher among customers who experienced such events.

Summary statistics are provided in Table 1.

[10] For customers that switched accounts, the bank of origin is the bank they switched from. For customers that did not switch, it is their current bank.

Table 1: Summary statistics

	Mean (standard deviation in parenthesis)		
	All customers	Searchers	Switchers
Dependent variables			
Search	0.1997		
	(0.0001)		
Switch	0.0376		
	(0.0000)		
Customer demographics			
Female	49.05%	43.77%	40.61%
	(0.0001)	(0.0004)	(0.0016)
Low income	53.68%	44.2%	49.18%
	(0.0001)	(0.0004)	(0.0016)
Aged 35 to 54	35.03%	33.04%	35.08%
	(0.0001)	(0.0004)	(0.0015)
Aged 55 to 64	20.77%	18.4%	19.29%
	(0.0001)	(0.0003)	(0.0013)
Aged 65 or above	17.7%	21.24%	13.98%
	(0.0001)	(0.0003)	(0.0007)
Degree	40.97%	50.88%	42.27%
	(0.0001)	(0.0005)	(0.0016)
Financial literacy	56.42%	68.39%	62.17%
	(0.0001)	(0.0004)	(0.0016)
Use of internet			
Internet confidence	76.89%	89.88%	85.42%
	(0.0001)	(0.0002)	(0.0009)
No internet banking	35.87%	24.64%	27.04%
	(0.0001)	(0.0004)	(0.0015)
No bank mobile app	64.37%	61.2%	55.06%
	(0.0001)	(0.0005)	(0.0016)
Monetary features			
Number of transactions	39.336	38.8261	34.9842
	(0.2923)	(1.2343)	(2.839)
Overdraft user	30.43%	27.62%	20.94%
	(0.0001)	(0.0004)	(0.001)
High credit balance	24.42%	32.23%	32.61%
	(0.0001)	(0.0004)	(0.0014)
Bank of origin			
Large bank	76.36%	70.23%	67.63%
	(0.0001)	(0.0003)	(0.0013)
Trigger factors			
Local branch closure	6.75%	9.04%	11.74%
	(0.0000)	(0.0002)	(0.0008)

Moved house	13.99%	16.39%	20.06%
	(0.0000)	(0.0002)	(0.0009)
Changed work status	14.42%	17.08%	16.73%
	(0.0000)	(0.0003)	(0.0007)
Changed relationship status	8.06%	6.82%	9.47%
	(0.0000)	(0.0001)	(0.0006)
Observations (unweighted)	3676	895	339

Note: all statistics are calculated using sampling weights provided by GfK.

4. Results

Tables 2 and 3 present the results of estimating the bivariate probit model of searching and switching. In table 2 we present the model coefficients, while in table 3 we show the corresponding marginal effects. For switching, we report the overall marginal effects, that is, the sum of the direct effect of each driver on switching conditional on searching plus the indirect effects through switching. The coefficients reported in table 2 correspond to the direct effects only, and therefore, allow us to identify whether overall effects are due to a direct effect on switching, an indirect effect through searching or both.

As expected, having searched has a strong effect on switching. Customers who report having looked around for a new PCA are 12 percentage points more likely to switch than those that did not.

In what concerns customer demographics, we find that women are on average less likely to switch than men but not necessarily to search. Low income customers are 4 percentage points less likely to search, which translates into a reduction in the probability of switching of 1.3 percentage points. This effect on switching is due uniquely to an indirect effect through searching. In other words, low income customers are not less likely to switch conditional on having searched. Regarding age, we find that the 55 to 64 age bracket are 6 percentage points more likely to search with respect to the reference group of 34 or younger, but are less likely to switch than them. No statistically significant difference is found for any of the other age groups.

As expected, the level of education and the confidence in the use of internet both impact significantly on the probability of searching. Customers that hold a degree or have a higher level of financial literacy are 3 and 5 percentage points more likely to search, respectively. The use of internet appears as a particularly important factor in whether customers search or not. Customers who report being confident in the use of internet are 13 percentage points more likely to search than those that do not use it or are not confident in its use. Moreover, customers who use internet banking are also more likely to search. Surprisingly however, none of these factors has an effect on the probability of switching. Only customers who use banking mobile apps are more likely switch.

In what concerns monetary features, customers with higher number of transactions are less likely to both search and switch, suggesting that a more intensive use of the PCA may be associated with higher perceived costs of switching. Customers holding high credit balances are more likely to both search and switch, consistent with higher expected monetary gains for this group of customers. However, as table 2 shows, the latter effect is the result of the indirect effect on switching through searching; high credit balances have no direct effect on switching conditional on searching. Overdraft users are less likely to switch than non-users, while no effect is found for switching. Although, overdraft users could potentially realise higher monetary gains by switching accounts, this does not translate in higher level of engagement. At the same time, the fact that overdraft users are less likely to switch suggests that these group of customers face barriers to switching. [11]

Customers whose bank of origin belongs to one of the four largest banking groups are less likely to search. They are also less likely to switch, although this is the result of an indirect effect through

[11] Information on overdraft usage comes from customers' bank of destination and reflects usage after switching or not. Therefore, the observed lower level of overdraft usage of switchers may be driven partly by switchers who have not yet been able to secure an overdraft facility with their new bank.

searching. Conditional on searching, holding an account with one of these banks has no effect on switching.

Regarding trigger factors, the closure of a local branch has a large positive effect on searching of 8 percentage points, while no statistically significant effect if found for switching. Also, changing work status has a positive effect on searching and indirectly on switching. We also tested whether other life events, such as moving house or changing relationship status, had an impact on searching and switching but did not find a statistically significant effect.[12]

[12] For brevity, these results are not reported.

Table 2: Searching and switching bivariate probit model, estimated coefficients

	(1) Searching	(2) Switching
Searching		1.541***
		(0.526)
Customer demographics		
Female	-0.067	-0.159*
	(0.0618)	(0.0867)
Income below £24,000	-0.157**	-0.095
	(0.0732)	(0.104)
Aged 35 to 54	-0.020	-0.097
	(0.0811)	(0.104)
Aged 55 to 64	0.222**	-0.386***
	(0.0960)	(0.149)
Aged 65 or above	0.115	-0.277**
	(0.108)	(0.140)
Degree	0.114*	-0.167*
	(0.0660)	(0.0909)
Financial literacy	0.192***	-0.021
	(0.0669)	(0.0956)
Use of internet		
Internet confidence	0.560***	-0.164
	(0.102)	(0.139)
No internet banking	-0.183**	
	(0.0798)	
No bank mobile app		-0.210**
		(0.0955)
Monetary features		
Number of transactions	-0.003**	-0.004**
	(0.00136)	(0.00187)
Overdraft user	-0.038	-0.208**
	(0.0734)	(0.102)
High credit balance	0.195**	0.117
	(0.0778)	(0.110)
Bank of origin		
Large bank	-0.133**	-0.101
	(0.0657)	(0.101)
Tigger factors		
Local branch closed	0.304***	0.100
	(0.112)	(0.158)
Changed work status	0.189**	
	(0.0838)	
Searching		1.541***
		(0.526)
Constant	-1.250***	-1.504***
	(0.161)	(0.221)
Rho	-0.207	
	(0.305)	
F-statistics	9.426	
Observations	3,502	

*** p<0.01, ** p<0.05, * p<0.1. Bootstrapped standard errors in parentheses.

Table 3: Searching and switching model, marginal effects

	(1) Searching	(2) Switching
Searching		0.116***
		(0.0135)
Customer demographics		
Female	-0.017	-0.013**
	(0.0166)	(0.00578)
Income below £24,000	-0.041**	-0.013*
	(0.0194)	(0.00684)
Aged 35 to 54	-0.005	-0.007
	(0.0203)	(0.00699)
Aged 55 to 64	0.062**	-0.015**
	(0.0270)	(0.00709)
Aged 65 or above	0.031	-0.013
	(0.0290)	(0.00792)
Degree	0.030*	-0.006
	(0.0171)	(0.00589)
Financial literacy	0.050***	0.006
	(0.0166)	(0.00579)
Use of internet		
Internet confidence	0.127***	0.010
	(0.0194)	(0.00716)
No internet banking	-0.047**	-0.007
	(0.0201)	(0.00426)
No bank mobile app		-0.014**
		(0.0068)
Monetary features		
Number of transactions	-0.001*	-0.000**
	(0.0004)	(0.000125)
Overdraft user	-0.010	-0.014**
	(0.0185)	(0.00562)
High credit balance	0.053**	0.017**
	(0.0223)	(0.00847)
Bank of origin		
Large bank	-0.036*	-0.013*
	(0.0185)	(0.00667)
Tigger factors		
Local branch closed	0.088**	0.022
	(0.0358)	(0.0154)
Changed work status	0.052**	0.008*
	(0.0261)	(0.00461)
Rho	-0.207	
	(0.305)	
F-statistics	9.426	
Observations	3,502	

*** p<0.01, ** p<0.05, * p<0.1. Bootstrapped standard errors in parentheses.

Note: Reported marginal effects for switching are the overall marginal effects, comprising of both the direct and indirect (through searching) effects.

5. Conclusion

This paper investigates the factors that drive customers to search for, and switch to, a new bank account. The decisions to search and switch are modelled as a function of drivers related to potential benefits of switching, expected costs or barriers to searching and switching, as well as supply and demand trigger factors. Our model captures the relationship between the decisions to search and switch, and shows that searching is one of the main drivers of switching increasing the probability of switching by 12 percentage points on average. Moreover, our results show that, once we control for searching, many of the drivers considered are not found to have a significant impact on switching, suggesting that the decision to search is an essential factor in the ultimate decision to switch.

Our results show that some customers may be facing cost of searching that discourage them to engage with the market. In particular, we find that customers with higher levels of education and financial literacy – who are likely to find it easier to compare and chose among alternatives – are more likely to search. Also, confidence in the use of internet, as well as the use of internet banking in particular, seem particular relevant drivers of searching. The use of internet is likely to be particularly important in reducing searching costs given that a large amount of information on personal bank accounts is available online.

However, these results do not carry over to switching, and a different set of drivers appear relevant to explain switching behaviour. Overdraft users are found to be less likely to switch, suggesting that overdraft usage may constitute a barrier to switching for some customers. Additionally, we also find that a higher level of account activity – which could be associated with higher perceived barriers to switching – is also found to lower the probability of switching.

In what concerns pull factors, we find that high credit balance holders are more likely to search, which is likely to be associated with higher potential gains from switching for this group. This however does not have a direct effect on switching once searching behaviour has been accounted for, suggesting that the main decision for this group is whether to search or not.

We also find evidence of the impact of supply side push factors on searching. In particular, the closure of a local branch of the customer's bank appears as a factor driving some customers to search. We also find evidence of demand trigger factors affecting searching. In particular, a change in work status increases the probability of searching.

References

Competition and Markets Authority (2016) "Retail banking market investigation, Final report"

Dong, Fengxia, David A. Hennessy and Helen H. Jensen (2010) "Contract and exit decisions in finisher hog production", *American Journal of Agricultural Economics* 92 (3) 667-684.

Giulietti, Monica, Catherine Waddams Price and Michael Waterson (2005) "Consumer choice and competition policy: a study of UK energy markets", *The Economic Journal* 115 (506) 949-968.

Greene, William H. (1998) "Gender economics courses in liberal arts colleges: Further results", *Journal of Economic Education* 29 (4) 291-300.

Roa, J., and K. Yue (1988) "Resampling inference with complex survey data", Journal of the American Statistical Association, vol. 83, pp. 231-241.

Waddams Price, Catherine, Catherine Webster and Minyan Zhu (2013) "Searching and switching: empirical estimates of consumer behaviour in regulated markets", *Centre for Competition Policy Working Paper 13/11*.

Appendix

Table A1: Variable definitions and sources

Variable	Definition	Source
Dependent variables		
Search	Dummy equals 1 if customer reported having looked around for a new PCA in the 12 months prior to the survey, 0 otherwise.	Customer survey
Switch	Dummy equals 1 if customer reported having switched their main current account to a different bank in the 12 months prior to the survey, 0 otherwise.	Customer survey
Customer demographics		
Age	Difference between 2015 and the customer's year of birth.	Transactions data
Female	Dummy equals 1 if customer is female, 0 otherwise.	Customer survey
Low income	Average monthly total value of payments and transfers into the account.	Transactions data
Degree	Dummy equals 1 if the highest level of education achieved is university degree or higher, 0 otherwise.	Customer survey
Financial literacy	Dummy equals 1 if customer answer correctly financial literacy question in survey, 0 otherwise. Respondents were asked to compute the total amount to be paid back for a loan of £500 and an interact rate of 10%. We considered as 'right' responses both £50 and £550.	Customer survey
Use of the internet		
Internet confidence	Dummy equals 1 if customers report having to the internet and being fairly or very confident in the use of internet, 0 otherwise.	Customer survey
No internet banking	Dummy equals 1 if customer reports to never use internet banking, 0 otherwise.	Customer survey
No bank mobile app	Dummy equals 1 if customer reports to never use bank's mobile app, 0 otherwise.	Customer survey
Monetary features		
Number of transactions	Monthly average number of credits and debits in the customer's PCA.	Transactions data
Overdraft user	Dummy equals 1 if either the customers' monthly average overdraft balance or the number of days in overdraft are positive, 0 otherwise.	Transactions data
High credit balance	Dummy equals 1 if customer's monthly average credit balance is within the top 25% of the overall distribution of average monthly credit balances in the transaction data, 0 otherwise.	Transactions data

Bank of origin		
Large bank	Dummy equals 1 if customer's bank of origin belongs to one of the largest banking groups - Barclays, Lloyds Banking Group (Bank of Scotland, Halifax and Lloyds Bank), RBS group (Natwest, RBS, Ulster), and HSBC group. For customers that have switched bank accounts, this is the bank they switched from, as reported in their response to the customer survey. For non-switchers it is the same as their current bank as indicated in the transactions data.	Customer survey and transactions data
Trigger factors		
Local branch closure	Dummy equals 1 if customer reports having experienced the closure of a local branch of their bank of origin in the 12 months prior to the survey, 0 otherwise.	Customer survey
Moved house	Dummy equals 1 if customer reports having move houses in the 12 months prior to the survey, 0 otherwise.	Customer survey
Changed work status	Dummy equals 1 if customer reports having started or stopped working in the 12 months prior to the survey, 0 otherwise.	Customer survey
Changed relationship status	Dummy equals 1 if customer reports having changed their relationship status, 0 otherwise. Changes in relationship status include getting married, start living with someone else, getting divorced, separating or widowing.	Customer survey