

RETAIL BANKING MARKET INVESTIGATION

Proposed approach for comparing actual and perceived behaviour of personal current account customers

1. The purpose of this paper is to set out the Competition and Market Authority's (CMA) proposed approach to comparing actual and perceived behaviour of personal current account customers.
2. The aim of the analysis is to help us assess whether consumers are engaged in the market, and whether they are aware of their account usage. This is particularly relevant to Theory of Harm 1, as set out in the CMA's Issues Statement.¹
3. The proposed analysis would combine anonymised data from a quantitative survey the CMA has commissioned, and anonymised data on customers' current account usage ('transactions data') which the banks² will provide pursuant to section 174 of the Enterprise Act 2002. It would seek to assess differences between customers' perceptions of their account usage and charges, and their actual past behaviour.
4. The main part of the perceived vs actual analysis would consist of testing a number of hypotheses around overdraft usage and charges. Evidence we have seen so far suggests that overdrafts may be a significant source of income for banks, and that a significant proportion of customers use them (almost 40% in 2013³), but that customers' awareness of their usage levels, as well as charging structures and/or level of fees, may be relatively low.
5. In addition to analysis on overdrafts, we would also like to test customers' awareness of whether or not they are receiving interest payments on credit balances.

¹ See the [Issues Statement](#).

² Where we refer to banks we mean banks and building societies.

³ According to the phase 1 report ([Personal current accounts: Market study update](#), dated 18 July 2014, page 116), in 2013, 25.3 million customers used an overdraft, which represents 39% of all active PCAs.

6. The rest of this note is structured as follows:
 - (a) First, we outline the hypotheses we are seeking to test in relation to overdraft usage and charges, and other charges and account usage, and the data we are planning to use in the analysis in relation to each hypothesis.
 - (b) We then outline the customer segments for which we are planning to test the hypotheses.
 - (c) Finally, we summarise the proposed methodology for testing the hypotheses.
2. If parties have comments, please send them to 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

Null hypotheses and data

7. We are seeking to test six hypotheses with the actual vs perceived analysis. These are described below alongside the data which would be used to test each hypothesis.
 - A. Does customers' awareness of their arranged overdraft limits match the limits actually provided by their banks?***
8. The hypothesis we are seeking to test is that customers are unaware of their arranged overdraft limits. This could demonstrate lack of engagement, and it may lead to use of unarranged overdrafts.
9. From the transaction data provided by banks we can obtain the actual value of the arranged overdraft limit for each customer as at the end of 2014.

10. From the customer survey we can obtain the perceived value of the arranged overdraft limit for each customer as of February 2015. The relevant survey question for the analysis is:
 - *Ask respondents who state they have arranged overdraft:* What is the authorised overdraft limit on your main current account?
11. We can then calculate the difference between perceived and actual value of the limit for each customer. We are interested in both the actual sign of the differences (whether customers underestimate or overestimate their limits), as well as the absolute values of the differences, as these will reveal the extent to which customers are wrong in either direction.
12. A potential drawback with this approach relates to the timings of actual and perceived values, in that the actual value is taken as of 31 December 2014, while the perceived value is taken in February 2015, and it may be the case that the actual value has changed for some customers between 31 December 2014 and February 2015. We believe that this may be the case for only a small set of surveyed customers (as we believe that overdraft limits do not change often), and can seek data from the banks to confirm this. It may then be possible to drop from the analysis those customers whose arranged overdraft limits changed between 31 December 2014 and February 2015.

B. Does customers' perception of their usage of overdrafts match their recent behaviour?

13. The hypothesis we are seeking to test is that customers underestimate their overdraft usage, which may then adversely affect their decisions on whether or not to switch banks, or use alternatives to overdrafts. Within this hypothesis, we can test the following three measures:
 - (a) Do customers underestimate the number of months per year in overdraft?
 - (b) Do customers underestimate the number of months per year in unarranged overdraft?
 - (c) Do customers underestimate the number of days per month in overdraft?
14. From the transactions data provided by banks we can obtain the actual value of number of months in 2014 that each customer was overdrawn (in unarranged overdraft) and average number of days per month that a customer was overdrawn in 2014. The analysis would use the following customer-level data from banks:
 - (a) The number of days in arranged overdraft for each month in 2014.

- (b) The number of days in unarranged overdraft for each month in 2014.
15. Data on the number of days in arranged and unarranged overdraft for all 12 months in 2014 has been requested from the five largest banking groups,⁴ from here on referred to as Group 1. Other banks, from here on referred to as Group 2⁵, are providing data for the last three months of 2014.
 16. For Group 1 banks (and their brands) we can then add the number of days in arranged overdraft and number of days in unarranged overdraft to obtain the number of days in overdraft for each month in 2014, and calculate the average number of days in overdraft per month when the customer was in overdraft. We can also calculate the number of months in 2014 that each customer was in overdraft (either arranged or unarranged, or only unarranged).
 17. It is worth noting that the data will differ for switchers and non-switchers, in that for non-switchers we have data for all 12 months, while for switchers we may have anything between one and 12 months. We will need to consider the impact of this, though importantly both the actual and perceived data for switchers will relate to a similar number of months.⁶ This also applies to hypotheses C and D below.
 18. For Group 2 banks, we can calculate the average number of days in overdraft across the last three months of 2014 by adding the number of days in arranged and unarranged overdraft for each month and averaging across months when in overdraft, as for Group 1. Using Group 1's data (where we can calculate both an average across all 12 months of 2014 and the last three months of 2014), we can then check whether the average number of days differs when calculated across three months vs 12 months. If it does not, then we can use the average across three months for Group 2 banks as a proxy for average across 12 months, and undertake the same analysis on the number of days in overdraft for Group 2 banks as for Group 1. If the average across the last three months of 2014 is not a good proxy for the average across all 12 months of 2014, we may not be able to undertake this analysis for customers of Group 2 banks.
 19. We are also not able to undertake the analysis of number of months in overdraft (unarranged overdraft) in 2014 for Group 2 banks. We believe that this would not have a significant impact on the analysis, as we believe that

⁴ Barclays, HSBC, Lloyds Banking Group, RBS and Santander.

⁵ Allied Irish Bank, Bank of Ireland, Clydesdale, Co-operative Bank, Danske Bank, Metro Bank, Nationwide and TSB.

⁶ The actual data will relate to the x number of months in 2014 that the respondent has been with the new bank, while the perceived data will relate to the x number of months in 2014, plus possibly January 2015.

any biases in customer behaviour (eg unawareness of actual number of days in overdraft) are unlikely to differ significantly between customers of Group 1 and Group 2 banks.⁷

20. From the customer survey we can obtain the perceived value of the number of months (in the last 12 months) in overdraft (unarranged overdraft) and perceived average number of days in overdraft per month for each customer. The relevant survey questions for the analysis are:

- *Ask all:* Have you gone overdrawn on your main current account at any time in the last 12 months?
- *Ask all who have been overdrawn:* In how many months in the last year were you overdrawn? Please give your best estimate.
- *Ask all who have authorised overdraft and have been overdrawn:* Were there any months when you exceeded your authorised overdraft limit? *IF YES:* In how many months did you exceed the limit?
- *Ask all who have been overdrawn in the last 12 months:* Thinking about the x months that you have been overdrawn in the last year, how many days in the month were you usually overdrawn? Please give me your best estimate.

21. We can then calculate the following differences between actual and perceived for each customer:

- (a) The number of months per year in overdraft – the difference between the perceived number of months in overdraft in the 12 months prior to the survey⁸ and the actual number of months in overdraft in 2014 (at least for customers of Group 1 banks).
- (b) The number of months per year in unarranged overdraft – the difference between the perceived number of months in unarranged overdraft in the 12 months prior to the survey⁹ and the actual number of months in unarranged overdraft in 2014 (at least for customers of Group 1 banks).
- (c) The average number of days per month in overdraft – the difference between the perceived number of days usually overdrawn in the months

⁷ We can also test this hypothesis directly using data for particular brands owned by Group 1 (eg Halifax, First Direct, etc), for which we will have data.

⁸ Assuming zero where respondents state they have not been in overdraft in the last 12 months in G4.

⁹ Assuming zero where respondents state they have not been in unarranged overdraft in the last 12 months in G4 if they do not have arranged overdraft, or in G6 if they have arranged overdraft.

when overdrawn in the 12 months prior to the survey¹⁰ and the actual average number of days overdrawn in the months when overdrawn in 2014.

22. We are interested in both the actual sign of the differences (whether customers underestimate or overestimate their usage), as well as the absolute values of the differences, as these will reveal the extent to which customers are wrong in either direction.
23. It is worth noting that though we have slightly different time periods for the actual (12 months in 2014) and perceived (most likely February to December 2014 and January 2015) data, this should not materially affect the analysis as we have 12 months in both cases, and can thus capture any cyclicity within the year. We can also explore obtaining data for January 2015 from the banks for the key variables so that we can compare the same months within the actual and perceived. This applies also to hypotheses C and D below.

C. Does customers' perception of the costs paid for overdrafts match their recent charges?

24. We are seeking to test the hypothesis that customers are (to varying degrees) unaware of the precise amount of charges they are paying for their overdrafts.
25. From the transaction data provided by banks we can obtain the actual value of overdraft charges paid by customers for each month in 2014. The analysis would use the following customer-level data from banks:
 - (a) Total interest paid for overdraft.
 - (b) Total value of non-interest charges paid in the use of the overdraft.¹¹
26. For Group 1 banks (and their brands) we can then add the total interest paid for overdraft and total value of non-interest charges in each month when the customer was in overdraft to obtain the monthly overdraft charges paid by each customer in each month of 2014 when in overdraft. We can then calculate an actual average monthly overdraft charge for each customer.
27. For Group 2 banks we can calculate the monthly overdraft charges paid by each customer in the last three months of 2014 when the customer was in overdraft. As above (see paragraph 15) we can try to use the average across up to three months (and drop customers of Group 2 banks who only used

¹⁰ Assuming zero where respondents state they have not been in overdraft in the last 12 months in G4.

¹¹ Data on overdraft charges for all 12 months in 2014 has only been requested from Group 1 banks. Group 2 banks are providing data for the last three months of 2014.

overdraft in one month in Q4 2014) for Group 2 banks as a proxy for average across up to 12 months, and undertake the same analysis on overdraft charges for Group 2 banks as for Group 1. Alternatively, we may have to omit data for Group 2 banks.

28. From the customer survey we can obtain the perceived value of the average monthly overdraft charges paid by customers. The relevant survey question for the analysis is:

- *Ask all who have been overdrawn in last 12 months:* Thinking about the x months that you were overdrawn, what was the average amount that you were charged for your overdraft in a month, whether authorised or not?

29. We can then calculate the difference between perceived and actual average monthly overdraft charges for each customer. We are interested in both the actual sign of the differences (whether customers underestimate or overestimate their usage), as well as the absolute values of the differences, as these will reveal the extent to which customers are wrong in either direction.

D. Does customers' perception of their average debit balances match their recent average debit balances?

30. The hypothesis we are seeking to test is that customers are unaware of their average debit balances. This would help us in measuring customers' awareness of their overdraft usage, and therefore their ability to assess alternative offers.

31. From the transaction data provided by banks we can obtain the actual value of average debit balances in 2014 for each customer. The analysis would use the following customer-level data from banks:

- Average overdraft balance for each month in 2014.¹²

32. For Group 1 banks we can then calculate an average across the monthly debit balances for each customer for the months in which the customer was in debit. We may not be able to undertake the analysis for (all) customers of Group 2 banks.

¹² Data on overdraft balances for all 12 months in 2014 has only been requested from Group 1 banks. Group 2 banks are providing data for the last three months of 2014.

33. From the customer survey we can obtain ranges of the perceived value of the average monthly debit balances for each customer. The relevant survey question for the analysis is:

- *Ask all who have been overdrawn in last 12 months:* Thinking about the x months that you have been overdrawn in the last year, what is the average amount you were overdrawn in a month?¹³

34. We can then compare the perceived values with the actuals for each customer, evaluating whether the actual value falls within the range of the perceived amounts. We are interested in both the actual sign of the differences (whether customers underestimate or overestimate their usage), as well as the absolute values of the differences, as these will reveal the extent to which customers are wrong in either direction. However, as we are only able to compare the perceived range with the actual value, we are likely to get a lower number of customers who are unaware of their usage than if we have more precise perceived values. This means that the analysis of this hypothesis is likely to be a conservative estimate of the true state.

E. Does customers' awareness of charging structures applying to their overdrafts match those applying to their accounts?

35. We are seeking to test the hypothesis that customers are (to varying degrees) unaware of the charging structures which apply to overdrafts.

36. We have requested data on pricing structures of PCA products from the banks, including specified types of charges applying and the amounts of those charges.

¹³ Responses will be grouped as follows:

1. £50 or less
2. More than £51 but less than £100
3. Less than £100
4. More than £100 but less than £200
5. More than £200 but less than £500
6. More than £500 but less than £1,000
7. More than £1,000 but less than £2,000
8. More than £2,000 but less than £5,000
9. £5,000 or more.

37. From the customer survey we can obtain the perceived charging structures. The relevant survey questions for the analysis are:
- *Ask all who have been overdrawn in last 12 months:* Which of these are included in what you pay for your authorised overdraft?
 - Interest.
 - Daily fee.
 - Monthly fee.
 - A fee for each amount paid from your account.
 - A fee for each payment cancelled by the bank.
 - Annual fee.
 - Notification fee (charged for notifying the consumer of a charge).
 - *Ask all who have been overdrawn in last 12 months:* Which of these are included in what you pay when you exceed your overdraft limit/are overdrawn?
 - Interest.
 - Daily fee.
 - Monthly fee.
 - A fee for each amount paid from your account.
 - A fee for each payment cancelled by the bank.
 - Annual fee.
 - Notification fee (charged for notifying the consumer of a charge).
 - Not sure/never overdrawn/never exceed limit.
38. We can then compare to what extent customers are aware of the different types of charges that apply on their current accounts.

F. Awareness of interest payments on credit balances

39. The hypothesis we are seeking to test is that customers may be unaware whether or not they are being paid interest on credit balances.

40. In the Market Questionnaire to the banks we have requested data on pricing structures of PCA products, including interest paid on credit balances, which would be used to calculate the actual values.
41. From the customer survey we can obtain the perceptions of credit interest. The relevant survey question for the analysis is:
- Which of these features, if any, does your main current account have?
 - Pays interest on money in account.
 - Includes insurance, such as for travel or mobile phones.
 - Free overdraft so you don't pay if you are overdrawn.
 - Pays cashback on bills/purchases.
42. We can then assess to what extent customers are aware of the following:
- (a) Are customers with accounts that pay 0% interest on credit balances aware that they are not getting interest? In other words, what proportion of respondents stated (in the survey question) that their account pays interest on money in account, when we know from the Market Questionnaire that this is not the case?
- (b) Are customers with accounts that pay interest on credit balances aware that they are receiving interest? In other words, what proportion of respondents stated (in the survey question) that their account does not pay interest on money in account? We can also test whether this proportion varies with the level of credit interest.

Customer segmentation

43. We would aim to test the above hypotheses both for all respondents, as well as for different segments (depending on sample sizes), including:
- (a) age groups;¹⁴
- (b) recent switchers (ie switched in the last three years) vs non-switchers;¹⁵

¹⁴ Using variable a127: Customer's year of birth from the transaction data, and supplemented where necessary by question K3a: What was your age last birthday? from the customer survey.

¹⁵ Using question F3: Have you personally changed your main current account in the last three years – this could be from one bank to another bank, or just from one account to another within the same bank? from the customer survey.

- (c) considerers (in the last three years) vs non-considerers;¹⁶
- (d) heavy usage of overdrafts vs light usage – these categories would need to be defined based on some or all of: number of months in overdraft, number of days per month in overdraft, monthly overdraft balances and monthly overdraft charges (the relevant data is discussed under hypotheses B, C and D above); and
- (e) holding one PCA vs multiple PCAs.¹⁷

44. We may explore further customer segmentation once we have analysed the transaction data.

Methodology

45. The hypotheses outlined above would need to be tested. There are different approaches to testing the hypotheses, from simply reading over the data and descriptive statistics, to statistical tests (such as the T-test, Z-test or Analysis of Variance test (ANOVA)). Our proposed approaches are outlined below.

Descriptive statistics

- 46. We are planning to produce descriptive statistics (mean and variance) of both the perceived and actual data points, and produce scatter plots of perceived vs actual. These will give us information on the extent to which customers underestimate and overestimate the various figures.
- 47. We will also produce charts showing the distributions of the differences between perceived and actual, and check for biases (eg towards underestimating usage) in consumers' perceptions.
- 48. We may also produce tables summarising the proportion of respondents whose perceptions do not match the actual (possibly allowing a small amount of variation, eg 5% or 10%), and the proportion whose perceptions are more than 25%, 50%, 75% or 100% different from the actual.

¹⁶ Using question F1: In the last three years have you looked around to compare different current accounts? from the customer survey.

¹⁷ Using question B3: Do you have any current accounts with anyone else apart from [bank], either sole or joint accounts? from the customer survey.

Statistical tests

49. It is also possible to run statistical tests on the data to check whether observed differences are statistically significant. We will explore which of the T-test and Z-test (both commonly used statistical tests) may be more appropriate in this analysis, alongside the ANOVA where more than two sets of observations need to be compared.