

## PAYDAY LENDING MARKET INVESTIGATION

### Customer- and transaction-level descriptive statistics—methodology note

#### Introduction

1. This note provides a brief overview of the process by which customer- and transaction-level data was collected from payday lenders and the steps taken to clean this information. The purpose of the note is to provide further detail of the methodology used by the Competition Commission (CC) in producing its summary statistics regarding payday customers and their loans. It should be considered alongside the ‘Customers and their loans—presentation based on analysis of the transaction data’ slide pack.
2. Please note that additional methodology details are provided in the notes section under the relevant slides.

#### Information requested

3. In August 2013, the CC requested customer- and transaction-level data from 11 major payday loan companies, including the largest lenders and a mix of online and high street lenders.
4. The data requested covered all payday loans issued by each lender in the period 1 January 2012 to 31 August 2013. The parties were asked to provide the data in three batches:
  - (a) **Table A: Customer-level data**—data relating to any customer who took out a loan in the reference period. The information included: details of the most recent loan (eg date, channel of purchase) and details of previous transactions (eg

number and value of loans in the reference period, date of the first ever loan taken with the lender etc).<sup>1</sup>

(b) **Table B, part 1: Loan details**—data relating to any loan taken in the reference period, for example product type, channel of purchase, loan date and value, interest and other charges, and details of any risk scores used to decide whether or not to approve the application. Table B, part 1 also included customer demographics such as date of birth, gender and income.

(c) **Table B, part 2: Loan status**—repayment information relating to each loan included in Table B, part 1, eg whether the loan was repaid in full, whether it was rolled over, the value of any repayments made on the loan before/on and after the original due date. The information about loan status was to be provided as of 1 October 2013.

5. We assigned each loan and customer a unique ID so that the information in the three tables could be merged together, allowing the creation of a single dataset containing loan and customer-level data for the reference period.

## Key definitions

6. For the purpose of the data requests, the following key definitions were used:
  - (a) **Payday** loan—unsecured loan product taken out for 12 months or less, where the minimum amount that can be borrowed is £1,000 or less, regardless of whether the loan is repaid as a lump sum or in instalments.
  - (b) A **loan** was defined as a fixed-sum loan agreement or an advance made to customers under a running account credit agreement. Rollovers/extensions were not considered separate loans.

---

<sup>1</sup> Table A also provides contact details of each customer. However, this information was mainly used in the customer survey rather than in our analysis.

- (c) A loan was defined as having been **rolled over** if the loan (or part of the loan) was not repaid on the date originally agreed, but where the borrower was not considered to have defaulted as a further agreement to extend the repayment period was entered into between the borrower and the lender. The customer may have paid all outstanding and unpaid fees, finance charges or interest at the time the rollover was executed but in all cases, some, or all, of the loan principal was carried forward after the date of the rollover. It was not considered a rollover if a lender and customer agreed to keep a loan agreement open and unpaid on the due date, AND the borrower was not considered to have defaulted, AND the customer was not charged further interest, fees or charges for the additional time.
- (d) **Principal**—the loan amount received by the customer.
- (e) **Repayment date**—the original date on which a loan was repayable, as agreed at the outset of the loan. For instalment products we refer to the date of the final repayment as agreed at the outset of the loan.
- (f) **Total amount repayable**—total amount due on the repayment date *as agreed at the time the loan was taken out*.
- (g) A **broker/lead generator**—any credit business involved in the effecting of introductions of individuals desiring to obtain credit to consumer credit businesses.
- (h) A loan was considered to have been **repaid in full** if there was no outstanding balance on this loan as of 1 October 2013.

## Data cleaning

7. As soon as the data was submitted, the CC began a thorough data checking and cleaning exercise to identify and correct any significant anomalies in the data and ensure consistency across datasets provided by different lenders.

8. Data checking involved identifying any missing or unusual values; duplicate observations; and any discrepancies between different datasets provided by the same lender (eg Table A with Table B) or between different variables in the same dataset. We performed the same set of checks on each lender's datasets.
9. Where a potential error affected a significant proportion of observations in the dataset, we raised it with the party in question, offering the opportunity either to correct the data or, if it turned out not to be an error but an accurate reflection of the lender's database, to explain why the discrepancy might have occurred. A number of variables/datasets were corrected as a result of these queries.
10. To summarize the most significant discrepancies remaining in the data subsequent to this cleaning (in addition to missing information for certain lenders):
  - (a) In a number of instances we observed discrepancies between Tables A and B. Since Table A provides aggregated information about loans taken by each customer and Table B provides detailed information about each loan, we considered the information in the Table B data to be more reliable and therefore used aggregated information from Table B when assessing the number of loans taken out by customers in the period covered by the sample.
  - (b) Credit limit—a number of lenders were unable to provide information about the maximum amount available to the customer at the time the loan was taken out. As a result the analysis is based on closest proxies where they were available. As certain products or lenders are excluded, the figures are not representative of the whole sample.
  - (c) Rollover and top-up fees—some lenders were not able to separate the rollover interest/fees from interest/fees on top-ups. The implications of this for our analysis of rollover interest/fees are noted in the slide pack.

11. Once all updated versions of the datasets were submitted, steps were taken to ensure that the format of variables, their names and labels were consistent across the different datasets. Finally, the different lenders' datasets were appended together to enable analysis of the set of companies as a whole.