

August 2017

Data Point: Frequent Overdrafters



- David Low
- Éva Nagypál
- Leslie Parrish
- Akaki Skhirtladze
- Corey Stone

This is another in a series of occasional publications from the Consumer Financial Protection Bureau's Office of Research. These publications are intended to further the Bureau's objective of providing an evidence-based perspective on consumer financial markets, consumer behavior, and regulations to inform the public discourse.

Table of contents

Table of contents.....	2
1. Introduction.....	3
2. Data.....	7
3. Characteristics of accounts by overdraft and NSF frequency.....	12
4. Differences among frequent overdrafters	19
5. Frequent overdrafter outcomes by opt-in status	28
6. Appendix	36
6.1 Excluding low-activity accounts	36
6.2 Averages versus medians.....	37
6.3 Accounts with fewer than six months of data	44
6.4 Cluster analysis method.....	46

1. Introduction

Consumers with checking accounts sometimes attempt transactions for amounts that exceed their account balances. Financial institutions that offer checking accounts decide whether to allow such transactions to go through and whether to charge fees in connection with them. These decisions depend on a number of factors, including the type of transaction, the financial institution's policies, procedures, and technological systems, and federal regulatory requirements.

In the case of a check or an Automated Clearing House (ACH) transaction, the financial institution may either return a transaction attempt that exceeds a consumer's account balance unpaid for non-sufficient funds (NSF), or process the transaction, in which case an overdraft occurs. In the case of a one-time debit card transaction (also called a point of sale (POS) debit card transaction) or an automated teller machine (ATM) withdrawal, the financial institution may either decline the transaction, in which case the transaction does not proceed, or authorize the transaction, possibly resulting in an overdraft.¹ Whether a financial institution charges an overdraft or NSF fee also varies depending on the circumstances. For check and ACH transactions, financial institutions typically charge fees for either an overdraft or NSF. For one-time card-based (OTCB) transactions (which include both one-time debit card and ATM transactions), financial institutions typically do not charge fees for declined transactions, and

¹ An overdraft may not occur in this circumstance if the consumer deposits sufficient funds into their account after the transaction is authorized but before the transaction settles.

under federal regulation may only charge fees for overdrafts if consumers have consented (or “opted in”) to such fees in advance.²

In a 2014 Data Point published by the Consumer Financial Protection Bureau’s (“CFPB” or “the Bureau”) Office of Research, the authors analyzed consumers’ experiences with overdrafts at a number of large banks.³ The analysis indicated that a small fraction—about eight percent—of consumer accounts were responsible for almost 75 percent of all overdraft fees. That study also compared several outcomes for consumers who opted in and did not opt in. The 2014 Data Point reported that opted-in accounts were three times as likely to have more than 10 overdrafts per year as accounts that were not opted-in. That Data Point also reported that opted-in accounts incurred more than seven times as many overdraft fees as accounts that were not opted-in.

This Data Point focuses on accounts with frequent overdrafts or NSFs.⁴ We define “frequent” overdrafters as accounts with more than 10 overdrafts and NSFs combined in a 12-month period. For comparison, we also define several other groups: “non-overdrafters” (those with no overdrafts or NSFs in a 12-month period), “infrequent” overdrafters (those with three or fewer overdrafts and NSFs combined in a 12-month period), and “occasional” overdrafters (those with more than three and up to 10 overdrafts and NSFs combined in a 12-month period). These

² Regulation E generally requires financial institutions to obtain opt-in from consumers to charge fees for overdrafts on OTCB transactions. The prohibition against charging non-opted-in customers fees for these transactions makes institutions less likely to authorize OTCB transactions that would bring the account balance negative. These requirements were established by the Board of Governors of the Federal Reserve in 2009. The requirements became effective for new accounts on July 1, 2010, and for existing accounts on August 15, 2010. See 12 C.F.R. § 1005.17.

³ CFPB, “Data Point: Checking Account Overdraft,” (July 2014), available at http://files.consumerfinance.gov/f/201407_cfpb_report_data-point_overdrafts.pdf.

⁴ In contrast to the 2014 Data Point, here we focus on both overdrafts and NSFs, since both signify attempts (whether intentional or not) to engage in a transaction that would take a consumer’s account balance negative. While we consistently include NSFs in our definition, for ease of exposition, we often refer to overdraft frequency or frequent overdrafters.

definitions include all overdraft and NSF transactions, whether or not they incur a fee. In some cases, we also discuss “very frequent” overdrafters (accounts with more than 20 overdrafts and NSFs combined in a 12-month period).

For our analysis, we used a de-identified dataset consisting of credit record data joined with transaction- and account-level data from several large banks (“study banks”) that were among those banks previously studied in the 2014 Data Point. This dataset allows us to study differences among the overdraft-frequency groups defined above. We then examine differences among frequent overdrafters, including by opt-in status.⁵

Key findings of this report include:

- Frequent overdrafters account for nine percent of all accounts at the study banks but paid 79 percent of all overdraft and NSF fees.⁶ Very frequent overdrafters account for about five percent of all accounts at the study banks but paid over 63 percent of all overdraft and NSF fees.
- Frequent overdrafters tend to be more credit constrained than non-overdrafters or infrequent overdrafters. Frequent overdrafters generally have lower credit scores and are less likely to have a general purpose credit card. Those that have a general purpose credit card have less available credit on such cards than non- or infrequent overdrafters.⁷

⁵ The banks in our dataset varied in their approach to presenting the opt-in decision to consumers.

⁶ These numbers differ slightly from those in the 2014 Data Point for three reasons. First, our definition of frequent overdrafters includes those with more than 10 overdrafts *and* NSFs *combined* in a year. Second, we now consider both overdraft *and* NSF fees. Third, for this Data Point we analyzed a subset of the accounts studied in the 2014 Data Point.

⁷ A general purpose credit card can be used to purchase goods and services at a wide range of merchants. These cards display one of the main network brands: American Express, Discover, MasterCard, or Visa. General purpose credit cards are accepted at any merchant that has contracted to accept that network brand. In contrast, private label cards—sometimes called “store cards” or “retail cards”—do not carry a network brand. Consumers can use these

- In our full sample, the dollar amount of monthly deposits into a checking account is not strongly correlated with the number of overdrafts or NSF incurred. Once we exclude low-activity accounts, however, we find that overdrafters have lower median deposits than non-overdrafters. The variability in monthly deposits is not strongly correlated with the number of overdrafts or NSF incurred, even after excluding low-activity accounts.
- The account usage characteristics and circumstances of frequent overdrafters vary considerably, which we capture by using a formal method to distinguish six groups of frequent overdrafters. Four groups constituting nearly 70 percent of frequent overdrafters have low end-of-day balances (with medians between \$237 and \$439), low or moderate credit scores (with medians between 532 and 661), and low or moderate monthly deposits (with medians between \$1,516, and \$2,724). Another group constituting 20 percent of frequent overdrafters has low end-of-day balances (median of \$140), low monthly deposits (median of \$1,313) and no credit score. The remaining group, constituting about 11 percent of frequent overdrafters, has higher end-of-day balances (median of \$1,403), significantly higher monthly deposits (median of \$7,828), but only moderate credit scores (median of 635).
- Frequent overdrafters who are opted-in appear similar to frequent overdrafters who are not opted-in along a number of dimensions but differ markedly in the number of overdraft fees they incur. Compared to the median frequent overdrafter that is opted-in to overdraft on OTCB transactions, the median frequent overdrafter that is not opted-in experiences only four fewer overdrafts per year but, accounting for fee reversals, pays 13 fewer overdraft fees per year. Thus, opted-in consumers pay significantly more overdraft fees but incur only slightly more overdrafts than consumers who are not opted-in.

cards only at a particular merchant or affiliated group of merchants. In the tables, we use the phrase “credit cards” to refer only to general purpose credit cards.

2. Data

This Data Point relies on a portion of the data set used in the 2014 Data Point—de-identified account- and transaction-level checking account histories from several large banks obtained through the CFPB’s supervisory role.⁸ The data are representative of the more than 40 million consumer accounts from the study banks.⁹

The transaction-level data contain every transaction that a random sample of the banks’ accountholders undertook during the 18-month period spanning January 2011 through June 2012. Each entry generally includes a transaction’s dollar amount, transaction type, posting date, and posting order.¹⁰ Entries also include the resulting account “ledger balance,” but not the

⁸ Neither the transaction- nor the account-level data contain any directly-identifying personal information. Moreover, because the data used in this analysis are Confidential Supervisory Information, this paper only presents results that are aggregated and does not identify specific institutions or describe the data in ways that could facilitate re-identification.

⁹ While consumer accounts are typically used to manage household finances, some consumers, especially those operating sole proprietorships, may in fact use their consumer account to operate small businesses.

¹⁰ The banks included in this study had a variety of transaction processing policies. While the details of these policies varied, all study banks had either a transaction processing policy approximating high-to-low posting (where debit transactions are posted according to their size, starting with highest transaction amounts), or a transaction processing policy approximating chronological posting (where debit transactions are posted according to the time they were received by the bank).

resulting “available balance.”¹¹ The entries also include a flag indicating whether the transaction resulted in an overdraft or an NSF and, if so, whether a fee was charged.

The account-level data are monthly account summaries for the 30 months spanning January 2010 through June 2012. For each month and each account in the sample, the account summary data include, among other information, the account opening and closing date, the reason for account closure, and the Regulation E opt-in status of the account. The banks also provided the year of birth for the primary holder of each account. In addition, they provided customers’ five-digit zip codes, which allow us to map to the account the median annual household income of the accountholder’s census tract, which we call “neighborhood income.”¹²

We analyze only “active” accounts, defined as accounts that averaged more than one customer-initiated debit transaction per month over the period for which we have data and during which the account was open.

We procured additional data for a randomly selected subset of the active accounts in our sample from a nationwide credit repository. Because our analysis focuses on the small subset of frequently overdrawn accounts, we selected our subsample to oversample from (and therefore provide more reliable results for) these accounts.¹³ We then constructed weights to use with our data so that our results are representative of accounts at the study banks during the study

¹¹ Debits and credits to an account can take place before they settle (*i.e.*, before funds are disbursed between institutions) or post (*i.e.*, the financial institution records the transaction on the account). A consumer’s ledger balance is the net sum of all posted debit and credit transactions against an account. The account’s available balance, in contrast, generally keeps track of outstanding debits not yet settled and credits that have been posted but are not yet cleared and thus are not fully available for withdrawal.

¹² The census tract income data are from the 2013 American Community Survey. The data are five-year estimates that represent information collected from January 1, 2009 through December 31, 2013. Neighborhood income is reported in 2013 inflation-adjusted dollars.

¹³ Twenty percent of the consumers in our subsample are frequent overdrafters compared to less than 10 percent of the original sample.

period. The additional data we procured consist of credit record information including credit scores, number of open credit relationships with lenders (known as “tradelines”), amount of available credit, and credit utilization from December 2010.¹⁴ December 2010 immediately precedes the start of our transaction-level data.

Some accounts could not be matched to the credit repository data. Failed matches could indicate that the accountholders did not have credit records. However, they could also indicate that the matching process itself was imperfect, perhaps because the accountholder’s bank and the credit repository had different or incomplete information that was used to perform the match (such as their address). Other accounts were successfully matched to the credit repository data but those data indicated that the credit repository could not assign the accountholders a value for the credit score we use. We refer to this latter group as “thin files.” Because we do not have credit scores for accountholders in any of these three groups, we collectively refer to them as “no credit score” accounts throughout this Data Point. When we do, the possible reasons for their lack of a credit score in our data should be kept in mind.

Overall, our dataset contains information on roughly 240,000 active accounts, including roughly 48,000 accounts belonging to frequent overdrafters. This dataset is representative of the over 40 million accounts in the study banks we analyze. To characterize annual account activity, we focused on the 12 months from July 2011 through June 2012. Nineteen percent of accounts in our sample have at least six but fewer than 12 months of data for these months. For these accounts we annualize overdraft and NSF counts to make them comparable to accounts

¹⁴ The Bureau did not possess any data that directly identified consumers at any point during or after the procurement of the supplemental credit records. To obtain matching credit records, the Bureau requested participating banks to provide the relevant identifiers associated with the sample accounts to the nationwide credit repository, which used the information to identify and pull the matching credit records. The credit records were then de-identified before being forwarded to the Bureau. The credit scoring algorithm referenced throughout this report is FICO version 8.0.

with all 12 months of data.¹⁵ Fourteen percent of accounts in our sample have fewer than six months of data for the chosen months. Annualizing overdraft and NSF counts is unlikely to be reliable for these accounts, so we exclude them from our main analysis.¹⁶ Some results for these excluded accounts are provided in Table 11 in the Appendix.

Some variable values in the data are outliers, lying well below or above the norm in the data. So that these outliers do not exert undue influence over our results, we replace the values of certain variables with the top (bottom) percentile of that variable if they are above (below) the top (bottom) percentile.¹⁷ Because of the wide distribution of some variable values, in our figures we further replace the values of certain variables with a large number below the top percentile if they are above that number. When we do this, the figures are labeled “topcoded,” and are topcoded at the highest value presented in the figure.¹⁸ This allows the figures to show greater detail for the vast majority of values for these variables but does not otherwise affect the statistics presented.

Because the data are account-level and consumers may hold multiple checking accounts, our findings may differ from those that would be obtained from consumer-level data. There are two main reasons for this. First, transaction-level data for a particular account may reflect only a portion of the consumer’s transactions. For example, average deposits in an account will be

¹⁵ Because we annualize counts, it is possible for accounts to have a non-integer number of overdrafts. For example, an account with nine overdrafts in 10 months of data is considered to have $(9/10)*12 = 10.8$ overdrafts in a year.

¹⁶ The analysis in the 2014 Data Point included accounts with fewer than six months of data. These include accounts that closed within the first six months of the observation period, accounts that opened within the last six months of the observation period, and accounts that were open for six months or longer during the observation period but had fewer than six months of activity.

¹⁷ This process is called “winsorizing” the variables and is a standard statistical technique.

¹⁸ For example, the 99th percentile of end-of-day balances in our data is \$40,421. Therefore, we replace observations of end-of-day balances above \$40,421 with \$40,421 for the numbers reported in this Data Point. Because \$40,421 is far above typical values of end-of-day balances, we further topcode end-of-day balance to \$5,000 in the figures (but not in the tables or in the text of the Data Point).

lower than average deposits in all accounts for consumers who make deposits into multiple checking accounts. Second, the more checking accounts a consumer has with the study banks, the more likely that consumer is to appear in our sample.¹⁹ Therefore, our sample is representative of *checking accounts* at the study banks, not of *consumers* at the study banks. We discuss the implications of our account-level data for the account usage characteristics of consumers throughout this Data Point. When we do, these caveats should be kept in mind.

Finally, as the sample comes from a small number of large banks, it cannot be considered fully representative of the checking account market as a whole. Nevertheless, based on our market observations, we believe it is likely to be similar to what one would observe at any institution that offers similar checking account and overdraft products.

¹⁹ For perspective, of households with a checking account in the 2010 Survey of Consumer Finances (SCF), 55.3 percent had one checking account, 30.9 percent had two checking accounts, 9.6 percent had three checking accounts, and 4.1 percent had four or more checking accounts. Fed. Reserve Sys., *2010 Survey of Consumer Finances*, Full Public Data (last updated 9/4/2014), available at https://www.federalreserve.gov/econres/scf_2010.htm.

3. Characteristics of accounts by overdraft and NSF frequency

This section examines overdraft and NSF behavior across accounts in the data. It also investigates differences in other observable characteristics between non-overdrafters, infrequent overdrafters, occasional overdrafters, and frequent overdrafters.

Overdraft and NSF incidence varies substantially across accounts.²⁰ Table 1 depicts the distribution of accounts and overdraft and NSF fees by groups of varying overdraft and NSF frequency in our sample.²¹ Overdraft and NSF frequency counts include all occurrences regardless of whether fees were assessed. We report findings separately for two groups of frequent overdrafters: (1) those with more than 10 and up to 20 annual overdraft and NSF transactions (“moderately frequent” overdrafters) and (2) those with more than 20 annual

²⁰ A few accounts (0.2 percent) are observed with overdraft or NSF fees but are never observed with an overdraft or NSF transaction. Some of this discrepancy is likely because of temporal mismatch between an overdraft or NSF transaction and the following fee, with the transaction occurring before the start of our data and the fee occurring afterwards. Some of the discrepancy may be due to transactions that were not overdrafts or NSFs incurring overdraft or NSF fees, which may have been refunded later. Part may also be due to transactions that were overdrafts or NSFs not being flagged as such in our data. Given the presence of the fee, we label these accounts as infrequent overdrafters.

²¹ Percentages may not add up to 100 percent here and in the rest of this report due to rounding.

overdraft and NSF transactions (“very frequent” overdrafters). Very frequent overdrafters account for 4.9 percent of accounts but incur 63.3 percent of all overdraft and NSF fees. Together, the two groups of frequent overdrafters account for 9.1 percent of accounts, but pay 78.7 percent of all overdraft and NSF fees.

TABLE 1: DISTRIBUTION OF ACCOUNTS AND OVERDRAFT/NSF FEES BY OVERDRAFT FREQUENCY

	Non- Overdrafters: 0 Annual OD/NSFs	Infrequent: 1 – 3 Annual OD/NSFs	Occasional: >3 – 10 Annual OD/NSFs	Moderately Frequent: >10 – 20 Annual OD/NSFs	Very Frequent: >20 Annual OD/NSFs
Share of all accounts	66.6%	14.5%	9.7%	4.2%	4.9%
Share of all overdraft & NSF with fees	0.0%	6.8%	14.5%	15.4%	63.3%

To examine whether other characteristics vary across these groups, the first part of Table 2 shows medians of selected account and consumer characteristics within overdraft frequency groups.²² Median end-of day balances, consumer age, and account tenure are highest among non-overdrafters and lowest among frequent overdrafters.²³ In contrast, median debit card usage—measured as the number of POS debit card transactions per month—is lowest among

²² To be precise, some reported characteristics are shares (such as the share of accounts matched to credit repository data). Shares are reported in percentage terms. Other characteristics involve only one observation per account (such as the account’s tenure). We report medians for these characteristics. For characteristics that involve multiple observations per account (such as monthly deposits), we first calculate a within-account average for each account (e.g. we calculate average monthly deposits for each account). Then, we take the median of these averages across accounts. Therefore, mentions of “median” in the text refer to the across-account median of the within-account average of these variables. For comparison, we also show the across-account average of the within-account average of the same variables in the Appendix.

²³ The within-account average end-of-day balance is calculated over all business days for which we observe the account regardless of whether transactions take place on that day.

non-overdrafters and highest among frequent overdrafters. The median very frequent overdrafter engages in a POS debit card transaction more than six times as often as the median non-overdrafter.

In our full sample, median monthly deposits do not vary systematically with overdraft frequency.²⁴ However, this may be because some accounts are not the primary financial instrument of their accountholders (possibly because these consumers have another checking account they use more frequently).²⁵ Non-primary checking accounts likely have both low deposits and few, if any, overdrafts. To examine this idea, Table 6 in the Appendix presents the same statistics as Table 2 after excluding “low-activity accounts” (which we define as accounts in the bottom quartile of monthly debit transactions). Once low-activity accounts are excluded, median monthly deposits are negatively correlated with overdraft frequency: median monthly deposits for non-overdrafters are \$3,145, while for moderately frequent overdrafters and very frequent overdrafters they are \$2,276 and \$2,659, respectively. Deposit variability (measured as relative standard deviation) varies little by overdraft frequency group.²⁶ The results support the assertion that low-activity accounts have low monthly deposits because they are predominantly non-primary accounts, and not because they are predominantly the primary accounts of

²⁴ Our definition of “deposits” excludes other account credits, including interest, rewards, and adjustments. Note that income that is not deposited (including income received as cash or on a prepaid card) does not appear in our data. Therefore, average monthly deposits in an account may differ from a consumer’s average monthly take-home income.

²⁵ This is an important concern. Numbers from the 2010 SCF imply that there is only a 33.6 percent chance that a checking account, selected at random out of all consumer checking accounts in the United States, is the only checking account of the accountholder’s household. Fed. Reserve Sys., *2010 Survey of Consumer Finances*, Full Public Data (last updated 9/4/2014), available at https://www.federalreserve.gov/econres/scf_2010.htm.

²⁶ The relative standard deviation of deposits captures the variability of monthly deposits relative to the average level of monthly deposits. Formally, it is equal to the standard deviation of monthly deposits divided by the average of monthly deposits during the observation period. For reference, the relative standard deviation of deposits for an account that receives biweekly deposits in equal amounts throughout the year with no other deposits is roughly 0.17. This variability captures the fact that such an account receives three paychecks in two months and two paychecks in all other months.

consumers with limited financial resources. For example, excluding low-activity accounts increases the median monthly deposits of non-overdrafters by over \$1,000, but increases their median end-of-day balance by less than \$150. This suggests that many low-activity accounts have significant end-of-day balances, despite their low deposits.²⁷

Table 2 also provides measures of neighborhood income, consumer creditworthiness, and consumer credit availability. Neighborhood income and credit scores decrease as overdraft frequency increases. Similarly, the share of consumers who have a general purpose credit card decreases as overdraft frequency increases, as does the median available credit on such cards among those consumers who have general purpose credit cards. The median credit score of frequent overdrafters is low (585 and 563 for moderately frequent and very frequent overdrafters, respectively). Consumers with lower scores generally have more difficulty obtaining new credit. For reference, a credit score below 620 is often considered to be “subprime.”²⁸ These findings suggest that consumers who overdraft frequently have less access to credit.

Because the credit characteristics in Tables 2 and 6 are consumer-level, rather than account-level, they further support the assertion that low-activity accounts are indeed non-primary, rather than primary accounts of consumers with limited financial resources.²⁹ In sharp contrast to the results for monthly deposits, dropping low-activity accounts produces almost no change in median credit scores.

²⁷ Low-activity accounts have median monthly deposits of \$779 and a median end-of-day balance of \$921.

²⁸ There is no universally accepted definition of a subprime credit score. The Bureau’s CARD Act Report defines credit scores below 660 as subprime and credit scores below 620 as deep subprime. See CFPB, “CARD Act Report,” (October 2013), available at http://files.consumerfinance.gov/f/201309_cfpb_card-act-report.pdf.

²⁹ The median credit score of low-activity accounts is 741.

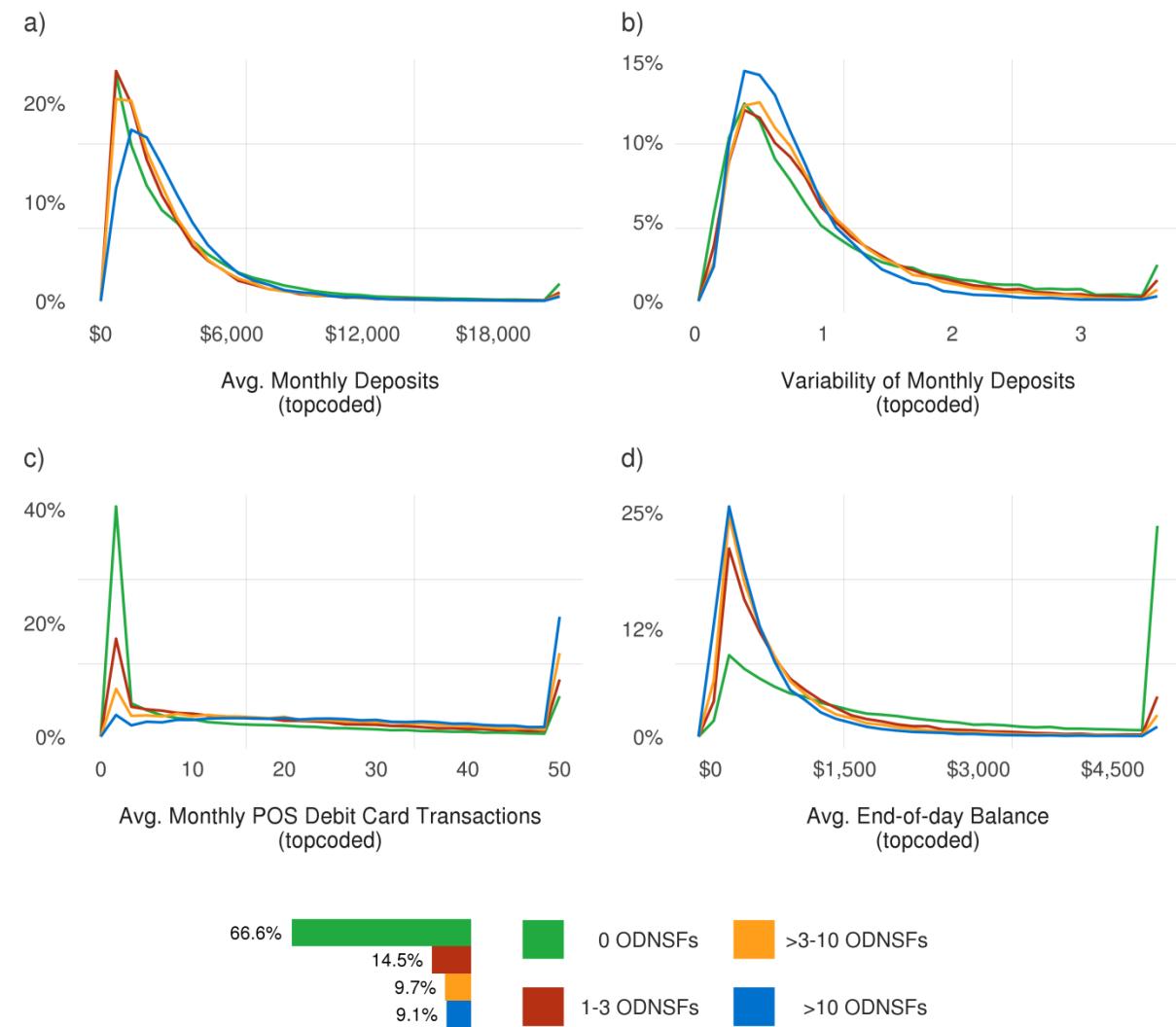
TABLE 2: MEDIAN ACCOUNT AND CONSUMER CHARACTERISTICS BY OVERDRAFT FREQUENCY

	Non- Overdrafters: 0 Annual OD/NSFs	Infrequent: 1 – 3 Annual OD/NSFs	Occasional: >3 – 10 Annual OD/NSFs	Moderately Frequent: >10 – 20 Annual OD/NSFs	Very Frequent: >20 Annual OD/NSFs
End-of-day balance	\$1,585	\$518	\$398	\$345	\$276
Monthly deposits	\$2,093	\$1,726	\$1,816	\$2,050	\$2,554
Variability of monthly deposits	0.64	0.68	0.66	0.62	0.56
Monthly count of ACH deposits	1.2	1.1	1.4	1.7	2.2
Monthly count of POS debit card transactions	4.6	14.6	21.1	25.3	29.1
Months of tenure	63.5	42.5	36.0	33.0	31.5
Age	46.3	39.4	37.1	36.5	37.4
Neighborhood income	\$59,832	\$55,939	\$54,736	\$54,953	\$54,265
Credit score	747	654	610	585	563
Share with credit card	86.6%	72.6%	63.5%	57.0%	48.9%
Available credit on credit cards, if any	\$14,100	\$3,000	\$960	\$521	\$225
Share with thin file	6.3%	10.7%	12.6%	12.7%	12.7%
Share unmatched	12.4%	12.7%	11.4%	8.8%	5.3%

While medians can be informative, they do not reflect within-group variation. To illustrate the variation that exists within overdraft frequency groups, Figure 1 shows the distributions of selected account characteristics by overdraft frequency group. It shows that monthly deposits, deposit variability, average monthly number of debit card transactions, and end-of-day balance

vary considerably *within* each overdraft frequency group. For example, Figure 1c shows that, while frequent overdrafters tend to have more debit card transactions than consumers with no or infrequent overdrafts or NSF, some frequent overdrafters have a low number of such transactions.

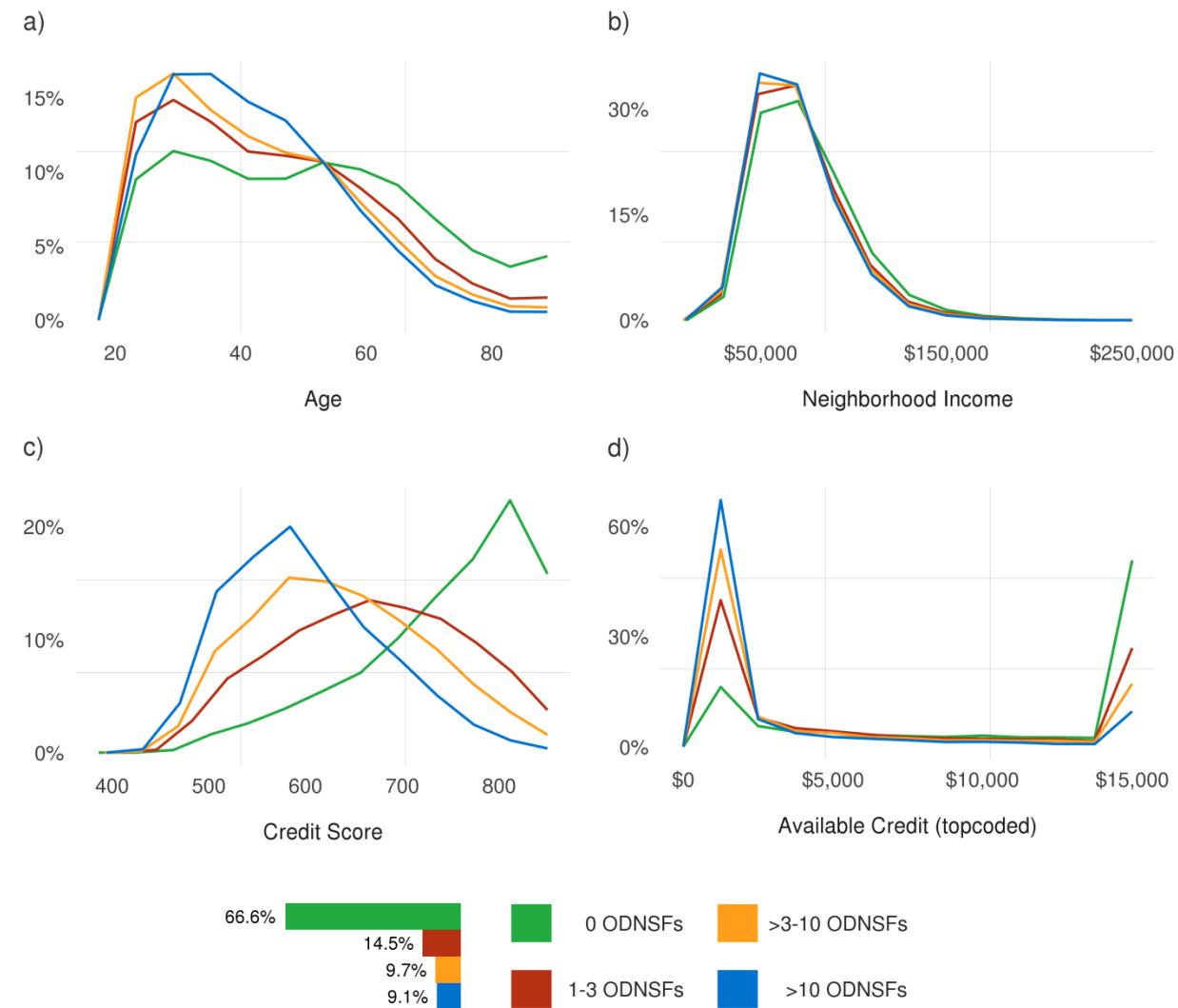
FIGURE 1: DISTRIBUTIONS OF SELECTED ACCOUNT CHARACTERISTICS BY OVERDRAFT FREQUENCY



Note: Figures 1a) – 1d) show the distributions of different account characteristics among accounts with (1) 0 annual OD/NSFs, (2) 1-3 annual OD/NSFs, (3) more than 3 and up to 10 annual OD/NSFs, and (4) more than 10 annual OD/NSFs. The bar graph shows the percent of accounts in each OD/NSF frequency group. Percentages do not add up to 100.0 because of rounding.

Figure 2 depicts distributions of selected consumer characteristics by overdraft frequency group. Age, neighborhood income, credit score, and amount of available credit all vary considerably within each overdraft frequency group.

FIGURE 2: DISTRIBUTIONS OF SELECTED CONSUMER CHARACTERISTICS BY OVERDRAFT FREQUENCY



Note: Figures 2a) – 2d) show the distributions of different consumer characteristics among accounts with (1) 0 annual OD/NSFs, (2) 1-3 annual OD/NSFs, (3) more than 3 and up to 10 annual OD/NSFs, and (4) more than 10 annual OD/NSFs. The bar graph shows the percent of accounts in each OD/NSF frequency group. Percentages do not add up to 100.0 because of rounding.

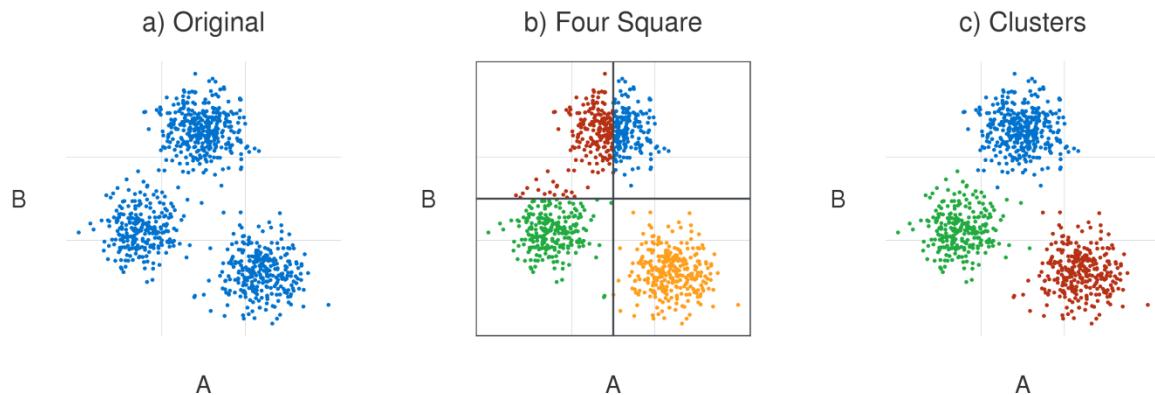
4. Differences among frequent overdrafters

The previous section depicted considerable variability in account usage, account characteristics and credit availability even among frequent overdrafters. This suggests that the circumstances that lead consumers to overdraw their accounts frequently may vary considerably. This section looks at frequent overdrafters in more depth.

We use cluster analysis, a statistical technique used widely in social sciences and marketing, to place consumers into distinct groups. In cluster analysis, a mathematical algorithm finds “clusters” of individuals who are similar to each other and different from individuals in other clusters in certain observable characteristics. The size of each cluster is determined by the data so that the size of the clusters can vary.

A simple version of cluster analysis is depicted in Figure 3. Figure 3a shows how individuals in a group vary along two different variables, A (plotted along the x-axis) and B (plotted along the y-axis). By imposing two thresholds at the mean of A and B respectively, we arrive at a simple “four square” grid grouping that identifies four different segments of individuals depicted in four colors based on their relationship to the respective means (3b). Clustering algorithms provide another approach: they aim to identify clusters of individuals who appear similar to each other and different from others (3c). This approach often yields clusters that display smaller within-group variability of characteristics and may thereby be more useful in defining groups and explaining how they differ.

FIGURE 3: GRAPHICAL REPRESENTATION OF CLUSTERING



Note: Figure 3a is a scatter plot of two arbitrary variables in hypothetical data with three visually obvious clusters. Figure 3b groups these hypothetical data by grouping together observations that lay above and below the respective means. Figure 3c groups the same hypothetical data using cluster analysis.

The first step of implementing cluster analysis is to choose a set of observable characteristics to use in the algorithm. We began with a large number of variables that differ considerably among frequent overdrafters and are strong predictors of overdraft and NSF behavior in our data. From this large number of variables, we chose the following set of four variables because it provides a better statistical predictor of overdraft behavior than any other set of four variables:³⁰

1. The consumer's credit score;
2. The average dollar amount of monthly deposits to the account during the observation period;

³⁰ Specifically, in negative binomial regressions of annual overdraft and NSF count on observable characteristics, the chosen combination of variables provides the highest goodness of fit (measured by McFadden's pseudo-R²) of all possible combinations of four variables.

3. The *variability* in the ratio of the dollar amount of monthly debits to the dollar amount of monthly deposits (“debit-deposit mismatch”), as measured by its relative standard deviation;³¹ and,

4. The share of months during which the consumer incurs at least one overdraft or NSF.

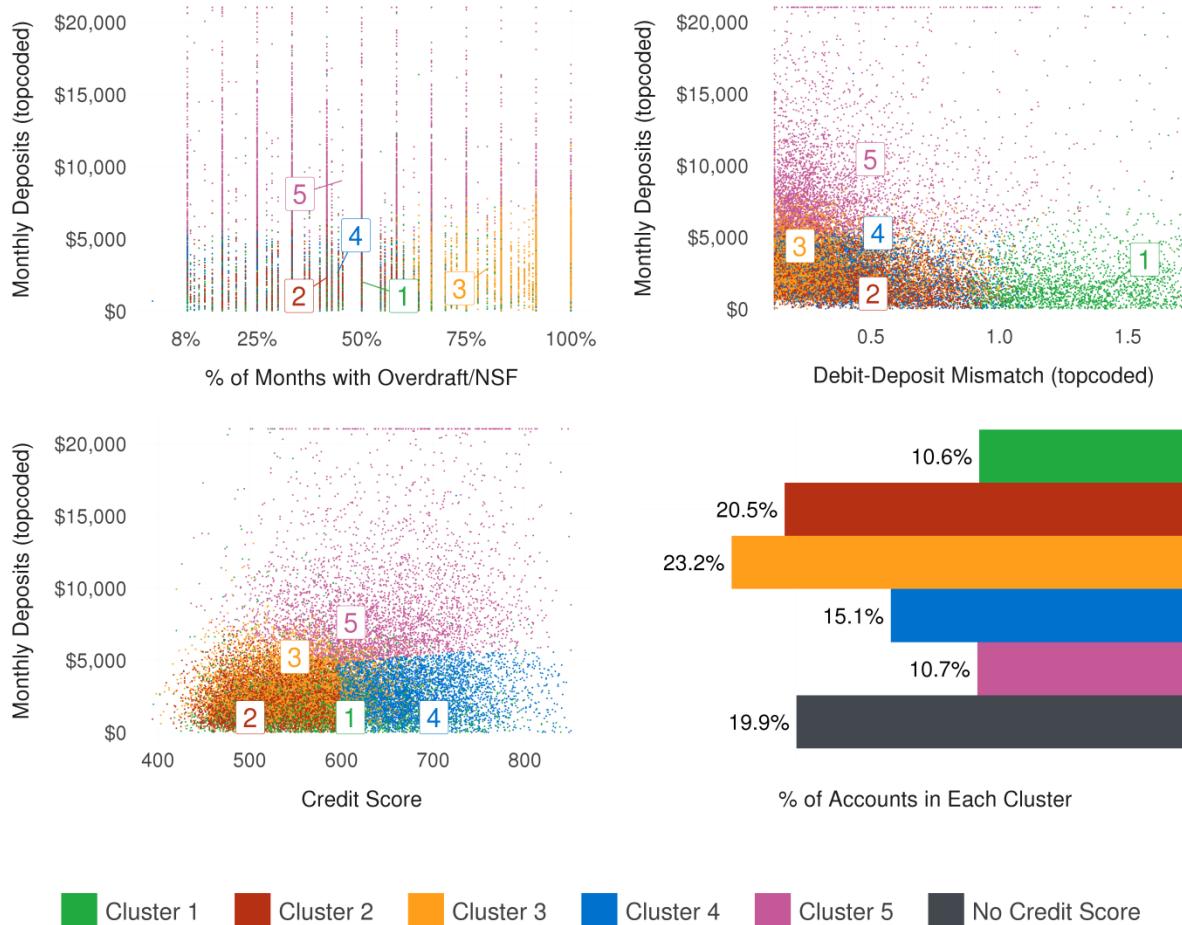
The next step in cluster analysis is to choose a specific algorithm to use to create the clusters, as well as the number of clusters to create. We considered many different possible clusterings of the data, and chose the one that provided the best fit to the data according to a formal ranking methodology.³² The top clustering produced five discrete clusters in the data plus a sixth group of consumers who could not be placed in a cluster because they do not have credit scores. We describe in the Appendix the algorithm used in this clustering, as well as the ranking methodology used to select this clustering.

Figure 4 depicts the five clusters constructed by our algorithm in three scatter plots showing how individual frequent overdrafters vary in different pairs of variables. Dots of five different colors (red, yellow, green, blue, and purple) represent individuals in each of the respective clusters. The numbers 1 through 5 show the average of each cluster along the two dimensions depicted in each scatter plot. Frequent overdrafters without a credit score cannot be placed in a cluster and therefore are not included in these scatter plots. As shown in the bar graph in Figure 4, about 20 percent of frequent overdrafters lack a credit score.

³¹ For example, by this definition a consumer whose dollar amount of monthly debits are always exactly twice the dollar amount of their monthly deposits will have a debit-deposit mismatch of zero (because there is no variability in the ratio of the dollar amount of monthly debits to the dollar amount of monthly deposits). Another consumer whose monthly debits are twice the amount of monthly deposits in six months of the year, and zero during the other six months of the year, will have a debit-deposit mismatch of one.

³² Results for the clustering that provided the second-best fit to the data are provided in the Appendix. These results show that the overall conclusions of our cluster analysis do not depend on the specific clustering generated.

FIGURE 4: CLUSTERING OF FREQUENT OVERDRAFTERS



Note: Figure 4 presents three scatter plots of observable characteristics used to cluster accounts in the data. In each scatter plot, the y-axis is topcoded average monthly deposits. The x-axis of each scatter plot differs. The bar graph shows the percent of frequent overdrafters in each cluster.

Table 3 depicts selected account and consumer characteristics for each of the five frequent overdrafter clusters and for frequent overdrafters without credit scores. For each of these groups, the table also reports overdraft characteristics, such as the share of accounts opted in, the number of annual overdraft and NSF transactions and fees, the share of months with an overdraft or NSF, the length of negative-balance periods (NBP) which indicates how quickly a consumer pays back funds after an overdraft, and the amount of NBP fees that the consumer

incurs.³³ The analysis shows that the five clusters and the no-credit-score group differ considerably from one another.

- Cluster 1: Consumers in this cluster constitute 10.6 percent of frequent overdrafters. In many respects, they appear to be in more financial distress than consumers in any other cluster; Cluster 1 has the lowest median monthly deposits of any cluster, and the highest variability of deposits, median debit-deposit mismatch, median NBP length, and charge-off rate.³⁴ Consumers in Cluster 1 have low credit scores and median end-of-day balances that are comparable to those of Clusters 2 and 3.
- Cluster 2: Consumers in this cluster constitute 20.5 percent of frequent overdrafters. This cluster has the lowest median credit score of any cluster. Unsurprisingly, these low credit scores correlate with low access to credit: Cluster 2 consumers have very low rates of general purpose credit card ownership and very low amounts of available credit on the general purpose credit cards they do have. They also have low median monthly deposits relative to consumers in other clusters. Cluster 2 consumers have a relatively high likelihood (11.4 percent) of having their accounts charged-off. The median length of their negative-balance periods is 7.5 business days, which may indicate that they experience extended periods of illiquidity. The median account tenure of Cluster 2 is 26 months, which is the same as that for Cluster 1 and substantially lower than the median account tenure of every other cluster. The median Cluster 2 consumer incurs a combined 17 overdrafts and NSFs a year, which is equal to or lower than the median for any other group.

³³ NBP fees are charged by some banks when a consumer fails to bring their account balance positive within a specified period of time after the account balance becomes negative. In previous Bureau publications, including the 2014 Data Point, we referred to NBP fees as sustained negative balance (SNB) fees.

³⁴ A charge-off occurs when an account closes while the consumer still owes money to the bank.

- Cluster 3: Consumers in this cluster constitute 23.2 percent of frequent overdrafters. Along several dimensions, these consumers appear similar to those in Cluster 2, although they have higher median monthly deposits. Notably, Cluster 3 consumers overdraft much more frequently and regularly than consumers in Cluster 2 or in any other cluster; the median consumer in this cluster incurs 43 overdrafts and NSFs combined in a year and incurs at least one overdraft or NSF in 10 months of the year. This cluster has a notably lower charge-off rate (3.8 percent) than Cluster 2 and a higher opt-in rate of 37.9 percent.
- Cluster 4: Consumers in this cluster constitute 15.1 percent of frequent overdrafters. Compared to consumers in the first three clusters, consumers in Cluster 4 have notably higher credit scores, account tenure, rates of general purpose credit card ownership, and amounts of available credit on general purpose credit cards. However, their median monthly deposits, variability of monthly deposits, debit-deposit mismatch and age are comparable to consumers in Clusters 2 and 3, while their end-of-day balances are somewhat higher. The charge-off rate of Cluster 4 (3.7 percent) is comparable to that for Cluster 3.
- Cluster 5: Consumers in this cluster constitute 10.7 percent of frequent overdrafters.³⁵ They have by far the highest median monthly deposits of consumers in any cluster. Their median end-of-day balances are roughly five to six times those in Clusters 1, 2, and 3. In addition, Cluster 5 has the shortest median negative-balance episode length (3.3 days) of any cluster. Cluster 5 consumer charge-off rates are the lowest, and their credit scores the second highest, of any cluster. They also have the highest median age.
- No credit score group: Consumers without a credit score constitute 19.9 percent of frequent overdrafters. In most respects they seem to be in even worse financial shape than consumers in Cluster 1. They have lower median monthly deposits, end-of-day

³⁵ Consumer accounts used to run small businesses may be disproportionately represented in Cluster 5, as evidenced by the high number of transactions processed through accounts in Cluster 5.

balances, and neighborhood income. Unsurprisingly, given their lack of a credit score, they have very low rates of general purpose credit card ownership and very low amounts of available credit on the general purpose credit cards they do have.³⁶ These consumers tend to be significantly younger and have shorter account tenure than the other groups.

TABLE 3: MEDIAN ACCOUNT AND CONSUMER CHARACTERISTICS BY CLUSTER

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	No credit score
Share of frequent overdrafters	10.6%	20.5%	23.2%	15.1%	10.7%	19.9%
Charge-off rate	12.9%	11.4%	3.8%	3.7%	0.7%	8.8%
End-of-day balance	\$270	\$295	\$237	\$439	\$1,403	\$140
Monthly deposits	\$1,516	\$2,131	\$2,724	\$2,552	\$7,828	\$1,313
Variability of monthly deposits	1.07	0.61	0.48	0.52	0.44	0.68
Monthly count of ACH deposits	0.8	2.0	2.3	2.1	3.1	1.3
Monthly count of POS debit card transactions	21.5	27.5	27.6	29.0	52.1	22.2
Debit-deposit mismatch	1.53	0.36	0.28	0.35	0.29	0.43

³⁶ Recall, as noted in Section 2, that consumers may lack a credit score in our data because: (1) their account was not matched to credit repository data because they do not have a credit record and hence do not have a credit score; (2) their account was not matched to credit repository data, even though they have a credit record (and potentially a credit score), because the information with which accounts were matched differed between their bank and the credit repository; and (3) their account was matched to credit repository data because they have a credit record, and the data indicated they have no credit score. Our credit repository data for consumers without a credit score therefore come entirely from this third group.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	No credit score
Share opted in	28.5%	24.5%	37.9%	25.2%	32.8%	24.5%
Annual count of overdrafts/NSFs	20	17	43	17	22	20
Annual count of overdraft/NSFs with fee	11	8	21	9	12	8
Annual count of overdraft/NSF fee reversals	0	0	1	0	0	0
Share of months with overdraft/NSF	50.0%	41.7%	83.3%	41.7%	41.7%	54.5%
Negative balance period length (business days)	8.5	7.5	5.9	4.7	3.3	7.9
Monthly negative balance period fees	\$2.06	\$1.17	\$3.89	\$0.00	\$0.00	\$1.94
Months of tenure	26.0	26.0	32.5	53.5	83.5	16.0
Age	36.3	37.3	38.4	36.0	44.3	29.3
Neighborhood income	\$55,960	\$52,163	\$52,467	\$57,151	\$66,565	\$51,119
Credit score	569	532	545	661	635	-
Share with credit card	53.4%	33.8%	40.4%	86.5%	79.5%	6.0%
Available credit on credit cards, if any	\$311	\$41	\$63	\$882	\$1,327	\$60

The cluster analysis suggests that the circumstances leading consumers to overdraft frequently may vary considerably from cluster to cluster. For example, consumers in Cluster 4 have significantly more access to credit than consumers in Clusters 1, 2, or 3, or the no-credit score group. Consumers in Cluster 1 have notably high deposit variability, while consumers in Cluster 5 have notably high deposits.

Frequent overdrafters also vary considerably in the proportion of their monthly deposits they pay in overdraft and NSF fees. A consumer with the median monthly deposits and the median

number of overdraft and NSF fees and reversals in Clusters 1, 2, 3, 4, and 5, and the no-credit-score group, would pay 2.1, 1.1, 2.1, 1.0, 0.4, and 1.7 percent, respectively, of their deposits in overdraft and NSF fees (assumed to be \$34, the median amount in the data). These numbers translate into large differences in the overdraft and NSF fee revenue generated by these clusters. Cluster 3 generates 32.1 percent of all overdraft and NSF fee revenue in our data (including that generated by infrequent and occasional overdrafters). The other clusters and the no-credit-score group generate each generate between 7.6 percent and 12.0 percent of all overdraft and NSF fee revenue in our data.

There are also characteristics most frequent overdrafters share. For example, the no-credit-score group and Clusters 1, 2, 3, and 4 (which together constitute nearly 90 percent of frequent overdrafters), all have low median end-of-day balances, low or moderate median monthly deposits, and low or moderate median credit scores (or no credit score at all).

5. Frequent overdrafter outcomes by opt-in status

As discussed above, financial institutions typically charge fees for check and ACH transactions that overdraw the consumer's account. However, fee decisions regarding one-time card-based (OTCB) transactions are restricted by federal regulation. Specifically, in 2009, the Federal Reserve Board amended Regulation E to require that institutions wishing to charge a fee for overdrafts on OTCB transactions obtain affirmative consent; a consumer who does not provide affirmative consent is deemed to have not opted in.³⁷ Thus, while an institution may authorize OTCB transactions that result in a negative balance on accounts that have not opted in, the institution may not assess overdraft fees for paying these transactions.³⁸

Our market observations suggest that one consequence of this regime is that institutions typically decline OTCB transactions for accounts that are not opted-in and have an insufficient balance at the time of authorization. In contrast, consumers who have opted in are more likely to have OTCB transactions authorized against a negative balance. A consumer can still avert an overdraft on such a transaction by making a sufficiently large deposit prior to the settlement of

³⁷ The amendment was effective July 1, 2010 for new customers and August 15, 2010 for existing customers. See 12 C.F.R. § 1005.17(c)(1). While the Regulation E opt-in rule does not apply to recurring debit card transactions, these transactions are relatively infrequent, making up just one percent of all debit card transactions in our data.

³⁸ See 12 C.F.R. § 1005.17(b)(1).

the OTCB transaction. If such an “intervening deposit” occurs, the OTCB transaction authorizes negative and settles positive.

The reverse can also occur: consumers may engage in OTCB transactions that authorize positive and settle negative. This most commonly happens as the result of an “intervening debit transaction” that the institution authorizes or pays between the authorization and settlement of the OTCB transaction and that sufficiently decreases the account balance.³⁹ “Authorize positive/settle negative” transactions are most likely to occur in accounts that regularly have very low balances and that transact frequently. They can occur even when consumers check their available balances and find that their accounts have sufficient funds to cover a transaction at the time of authorization. They may occur both to consumers who have opted in to overdraft on OTCB transactions and to consumers who have not opted in.

Since the Regulation E opt-in rule prohibits institutions from charging an overdraft fee on an OTCB transaction unless the consumer has opted in, consumers who have not opted in may not be charged overdraft fees on OTCB transactions. Thus, only opted-in consumers can be charged a fee on an authorize positive/settle negative transaction.⁴⁰ Whether they in fact are charged

³⁹ Authorize positive/settle negative transactions may also occur if the authorization amount is less than the settlement amount, such as when a restaurant bill is authorized for an amount that does not include a tip. They may also occur if a deposit that had been credited to an account is returned unpaid between the time of authorization and settlement or under other circumstances (e.g., an account adjustment, fee assessment, etc.).

⁴⁰ For example, suppose a consumer at a financial institution that uses available balance to assess fees has an available balance of \$100 when they engage in a POS debit card transaction for \$45. This leaves the consumer with a positive available balance of \$55 at the time the transaction is authorized. The next day, a check for \$80 from their account is processed, but no other transactions occur, leaving them with an available balance of -\$25. Because the check posts to a negative available balance, the institution is likely to charge the consumer an overdraft fee for the check transaction, regardless of their opt-in status. At the beginning of the following day, the consumer's debit card transaction settles. Because their available balance is negative at the time the debit card transaction settles, this is an “authorize positive/settle negative” transaction. If the consumer has not opted in, they cannot be charged an overdraft fee on the debit card transaction. If they have opted in, the consumer is likely to be charged an overdraft fee on the debit card transaction. Note that in this scenario, the opted-in consumer is likely to be charged two overdraft fees even though their initial balance is sufficient to cover either of the two transactions.

fees on these transactions can be greatly influenced by financial institutions' policies, such as whether to use balance at authorization or settlement for fee decisioning or whether to base fee decisions on ledger or available balance.⁴¹

Our data set does not generally allow us to observe declined transactions or the available account balance at the time an OTCB authorization request is made.⁴² Thus, we cannot generally observe if a transaction authorizes negative but settles positive, nor can we observe if a transaction authorizes positive but settles negative.

However, our data set does allow us to observe the opt-in status of accounts under Regulation E. It also enables us to identify overdraft transactions by transaction type (e.g., when an overdraft has resulted from an OTCB transaction or check/ACH transaction) and to identify when an OTCB transaction results in an overdraft fee being charged.

Of the frequent overdrafters in our data set, 30.5 percent are opted-in to overdraft on OTCB transactions. While this is 2.5 times higher than the 12.2 percent of other accounts that have opted in, a majority of frequent overdrafters have not opted in.

Table 4 depicts median characteristics of frequent overdrafters who are opted-in and not opted-in. Most characteristics are very similar, including credit score, end-of day balances, variability

⁴¹ Some institutions assess overdraft and NSF fees based on the account's ledger balance at the time of settlement while others base these fees on the account's available balance. The available balance does not include deposited funds that have not yet been made available and does include transactions that have been authorized but not yet settled. Ledger balances are typically determined at the end of each business day; available balances may be updated more frequently.

⁴² We observe the ledger balance of accounts following each posted transaction. However, financial institutions typically determine whether to authorize an OTCB based upon the available balance at the time the customer attempts the transaction.

of monthly deposits, and monthly debit card use.⁴³ For example, the median credit scores of opted-in consumers and non-opted-in consumers are within two points of each other, and the median end-of-day balances are within \$8.

TABLE 4: MEDIAN ACCOUNT AND CONSUMER CHARACTERISTICS OF FREQUENT OVERDRAFTERS BY OPT-IN STATUS

	Not Opted In	Opted In
End-of-day balance	\$304	\$312
Monthly deposits	\$2,211	\$2,643
Variability of monthly deposits	0.60	0.55
Monthly count of ACH deposits	2.0	2.1
Monthly count of POS debit card transactions	28.0	26.2
Credit score	573	571
Annual count of overdrafts (excluding NSFIs)	18.0	22.0
Annual count of overdrafts with fee (excluding NSFIs)	5.0	19.0
Annual count of overdraft fee reversals	0.0	1.0

⁴³ As discussed earlier, we have no way to identify which OTCB transactions authorize negative for opted-in accounts in our data. However, in our data opted-in frequent overdrafters do not successfully engage in more debit card transactions than non-opted-in frequent overdrafters. This suggests that debit card transaction attempts against a negative balance are not common enough to be quantitatively significant.

	Not Opted In	Opted In
Annual count of NSFs	0.0	0.0
Annual count of NSFsWith fee	0.0	0.0
Annual count of overdrafts with fee on debit cards	0.0	12.0
Annual count of overdrafts without fee on debit cards	12.0	1.0
Monthly count of recurring debit card transactions	0.1	0.0
Monthly count of other debit transactions	8.2	6.0

However, there are considerable differences in overdraft outcomes between frequent overdrafters who are opted-in and those who are not.⁴⁴ At the median, opted-in frequent overdrafters overdraft 22 percent more often than frequent overdrafters who are not opted in (22 vs. 18 times over twelve months). In contrast, the median opted-in frequent overdrafter pays 260 percent more overdraft fees than the median frequent overdrafter who is not opted in (19 fees with one reversal vs. five fees with no reversals).⁴⁵ Assuming an overdraft fee of \$34, this means that compared to the median non-opted-in frequent overdrafter, the median opted-in frequent overdrafter engages in four more overdraft transactions and pays \$442 more in

⁴⁴ Here we focus on only overdrafts, and not NSFs, unlike elsewhere in this Data Point. The reason is that, while opt-in status may determine whether or not a financial institution authorizes a transaction with a negative balance, opt-in status only affects fees on OTCB transactions and declination of these transactions does not result in an NSF fee. Results for NSFs are shown separately from overdrafts in Table 4.

⁴⁵ Due to differences in banks' policies (including the transaction processing policies mentioned earlier), these ratios are somewhat different among the different banks, ranging from 15 to 50 percent for the first ratio and 161 to 425 percent for the second ratio.

overdraft fees, even after accounting for fee reversals. A large part of this difference appears to be due to differences in overdraft fees on OTCB transactions. The median opted-in frequent overdrafter incurs 12 overdrafts with fees on debit card transactions per year and one overdraft without a fee on debit card transactions.⁴⁶ The median frequent overdrafter who is not opted-in incurs no overdrafts with fees on debit card transactions but 12 overdrafts without fees on debit card transactions. This difference is most likely the result of authorize positive/settle negative transactions. If so, this would mean that the decision whether to opt in has substantial consequences for consumers over and above simply increasing the likelihood that a transaction will be authorized and result in overdraft fees when there are insufficient funds at the point of authorization.⁴⁷

Table 5 depicts median characteristics by opt-in status for each of the five clusters and the no credit score group identified in Section 4. As indicated in Table 3, these six groups have similar opt-in rates. In addition, all six groups display similar disparities in the number of total overdraft fees paid by opted-in and non-opted-in members within each group.

⁴⁶ An overdraft on an OTCB transaction may still not result in a fee for an opted-in consumer due to *de minimis* policies, whereby an institution automatically waives fees when the amount of the overdraft transaction or the amount by which the transaction takes the account negative is within some threshold amount, or due to caps on the number of fees charged per day. Note that discretionary fee reversals show up explicitly as a fee being charged then later reversed in our data, so these discretionary reversals do not explain why overdraft transactions do not incur a fee in the first place.

⁴⁷ This conclusion also seems to hold for infrequent and occasional overdrafters. Pooling these two groups together, the median infrequent or occasional overdrafter who is not opted-in experiences two overdrafts: one overdraft with a fee, zero overdrafts with a fee on a debit card, and one overdraft without a fee on a debit card per year. The median infrequent or occasional overdrafter who is opted-in experiences three overdrafts: two overdrafts with a fee, one overdraft with a fee on a debit card, and zero overdrafts without a fee on a debit card per year.

TABLE 5: MEDIAN ACCOUNT AND CONSUMER CHARACTERISTICS OF FREQUENT OVERDRAFTER CLUSTERS BY OPT-IN STATUS

	Cluster 1: Not Opted In	Cluster 1: Opted In	Cluster 2: Not Opted In	Cluster 2: Opted In	Cluster 3: Not Opted In	Cluster 3: Opted In	Cluster 4: Not Opted In	Cluster 4: Opted In	Cluster 5: Not Opted In	Cluster 5: Opted In	No credit score: Not Opted In	No credit score: Opted In
End-of-day balance	\$290	\$232	\$294	\$301	\$241	\$229	\$433	\$459	\$1,436	\$1,344	\$141	\$136
Monthly deposits	\$1,488	\$1,608	\$2,113	\$2,220	\$2,598	\$2,942	\$2,514	\$2,707	\$7,813	\$7,883	\$1,253	\$1,540
Variability of monthly deposits	1.08	1.03	0.61	0.58	0.48	0.46	0.52	0.49	0.45	0.43	0.69	0.64
Monthly count of ACH deposits	0.8	0.8	2.0	2.1	2.3	2.3	2.1	2.1	3.0	3.1	1.3	1.4
Monthly count of POS debit card transactions	22.8	19.0	28.1	26.1	28.8	25.6	29.6	27.3	52.0	52.1	22.9	19.9
Debit-deposit mismatch	1.5	1.6	0.4	0.4	0.3	0.3	0.4	0.3	0.3	0.3	0.4	0.4
Annual count of overdrafts (excluding NSF)	16.0	18.0	15.0	15.0	37.0	40.0	15.0	16.0	19.0	22.0	17.3	20.6
Annual count of overdrafts with fee (excluding NSF)	4.4	15.6	4.0	13.5	11.0	37.0	5.0	14.0	7.0	20.0	3.4	18.0
Annual count of overdraft fee reversals	0.0	1.2	0.0	1.0	0.0	2.0	0.0	1.0	0.0	1.0	0.0	1.3

	Cluster 1: Not Opted In	Cluster 1: Opted In	Cluster 2: Not Opted In	Cluster 2: Opted In	Cluster 3: Not Opted In	Cluster 3: Opted In	Cluster 4: Not Opted In	Cluster 4: Opted In	Cluster 5: Not Opted In	Cluster 5: Opted In	No credit score: Not Opted In	No credit score: Opted In
Annual count of overdrafts with fee on debit cards	0.0	10.9	0.0	9.0	1.0	22.0	0.0	9.0	0.0	10.0	0.0	13.7
Annual count of overdrafts without fee on debit cards	10.9	1.0	10.0	1.0	22.0	2.0	10.0	1.0	11.0	1.0	12.0	2.0
Monthly count of recurring debit card transactions	0.1	0.0	0.2	0.0	0.2	0.0	0.2	0.0	0.7	0.0	0.0	0.0
Monthly count of check debit transactions	0.6	0.8	0.7	1.0	1.0	1.3	1.4	1.7	5.5	5.8	0.1	0.3
Monthly count of ACH Debit transactions	1.8	1.8	1.7	2.0	2.3	3.0	3.7	4.4	9.7	9.9	0.5	0.8
Credit score	572	562	532	533	543	549	662	659	639	627	-	-

6. Appendix

6.1 Excluding low-activity accounts

Table 6 is identical to Table 2, except it reports results excluding accounts in the bottom quartile of monthly debit transactions. The goal of Table 6 is to exclude accounts that are likely to be non-primary accounts, not to exclude accounts for consumers with limited financial resources. Results show this works well; for example, compared to non-overdrafters that are not excluded, excluded non-overdrafters have lower deposits but comparable credit scores.

TABLE 6: MEDIAN ACCOUNT AND CONSUMER CHARACTERISTICS BY OVERDRAFT FREQUENCY (TOP 75% MOST ACTIVE ACCOUNTS)

	Non- Overdrafters: 0 Annual OD/NSFs	Infrequent: 1 – 3 Annual OD/NSFs	Occasional: >3 – 10 Annual OD/NSFs	Moderately Frequent: >10 – 20 Annual OD/NSFs	Very Frequent: >20 Annual OD/NSFs
End-of-day balance	\$1,724	\$603	\$462	\$390	\$295
Monthly deposits	\$3,145	\$2,197	\$2,136	\$2,276	\$2,659
Variability of monthly deposits	0.52	0.59	0.60	0.59	0.55
Monthly count of ACH deposits	2.1	1.8	1.8	1.9	2.2
Monthly count of POS debit card transactions	15.7	21.3	25.3	28.2	30.4
Months of tenure	75.5	45.5	38.5	34.5	32.0

	Non- Overdrafters: 0 Annual OD/NSFs	Infrequent: 1 – 3 Annual OD/NSFs	Occasional: >3 – 10 Annual OD/NSFs	Moderately Frequent: >10 – 20 Annual OD/NSFs	Very Frequent: >20 Annual OD/NSFs
Age	45.3	38.3	36.3	36.3	37.3
Neighborhood income	\$59,916	\$55,960	\$55,146	\$55,351	\$54,652
Credit score	742	651	612	588	564
Share with credit card	86.7%	72.5%	64.2%	57.9%	49.3%
Available credit on credit cards, if any	\$13,500	\$2,395	\$897	\$526	\$225
Share with thin file	5.8%	10.2%	11.9%	11.9%	12.1%
Share unmatched	9.3%	11.5%	10.4%	8.3%	5.1%

6.2 Averages versus medians

Tables 7-10 are identical to Tables 2-5, except they report averages instead of medians. Average deposits, balances, and neighborhood income are typically higher than the respective medians because the distribution of these variables is right-skewed.⁴⁸

⁴⁸ Recall that the data are winsorized, as described in Section 2.

TABLE 7: AVERAGE ACCOUNT AND CONSUMER CHARACTERISTICS BY OVERDRAFT FREQUENCY

	Non- Overdrafters: 0 Annual OD/NSFs	Infrequent: 1 – 3 Annual OD/NSFs	Occasional: >3 – 10 Annual OD/NSFs	Moderately Frequent: >10 – 20 Annual OD/NSFs	Very Frequent: >20 Annual OD/NSFs
End-of-day balance	\$4,681	\$1,316	\$899	\$704	\$533
Monthly deposits	\$3,433	\$2,698	\$2,691	\$2,845	\$3,312
Variability of monthly deposits	0.94	0.89	0.83	0.76	0.67
Monthly count of ACH deposits	1.6	1.5	1.7	1.8	2.3
Monthly count of POS debit card transactions	13.8	20.5	26.1	30.1	34.2
Months of tenure	104.6	70.9	61.5	57.4	55.7
Age	47.4	41.6	39.4	38.9	39.3
Neighborhood income	\$65,214	\$60,858	\$59,414	\$59,044	\$58,661
Credit score	726	653	617	594	572
Share with credit card	86.6%	72.6%	63.5%	57.0%	48.9%
Available credit on credit cards, if any	\$21,794	\$11,864	\$7,670	\$5,044	\$3,443
Share with thin file	6.3%	10.7%	12.6%	12.7%	12.7%
Share unmatched	12.4%	12.7%	11.4%	8.8%	5.3%

TABLE 8: AVERAGE ACCOUNT AND CONSUMER CHARACTERISTICS BY CLUSTER

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	No credit score
Share of frequent overdrafters	10.6%	20.5%	23.2%	15.1%	10.7%	19.9%
Charge-off rate	12.9%	11.4%	3.8%	3.7%	0.7%	8.8%
End-of-day balance	\$604	\$402	\$303	\$586	\$2,238	\$327
Monthly deposits	\$2,021	\$2,315	\$2,958	\$2,626	\$9,108	\$1,794
Variability of monthly deposits	1.22	0.68	0.55	0.62	0.57	0.79
Monthly count of ACH deposits	1.1	2.0	2.5	2.0	3.2	1.6
Monthly count of POS debit card transactions	25.8	31.5	31.3	31.7	55.3	26.3
Debit-deposit mismatch	1.77	0.31	0.32	0.40	0.40	0.62
Share opted in	28.5%	24.5%	37.9%	25.2%	32.8%	24.5%
Annual count of overdrafts/NSFs	29	22	55	22	33	29
Annual count of overdraft/NSFs with fee	16	11	31	11	19	14
Annual count of overdraft/NSF fee reversals	1	1	2	1	1	1
Share of months with overdraft/NSF	50.5%	42.1%	80.7%	45.1%	46.2%	55.1%
Negative balance period length (business days)	12.3	11.8	7.3	10.4	12.2	11.1
Monthly negative balance period fees	\$4.02	\$2.88	\$6.55	\$1.99	\$1.08	\$4.33
Months of tenure	46.8	45.8	52.8	78.2	106.8	32.1
Age	38.1	39.3	40.7	39.5	45.4	33.9
Neighborhood income	\$60,435	\$56,079	\$56,150	\$60,696	\$70,555	\$55,931

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	No credit score
Credit score	577	530	546	669	639	-
Share with credit card	53.4%	33.8%	40.4%	86.5%	79.5%	6.0%
Available credit on credit cards, if any	\$3,787	\$734	\$1,074	\$5,557	\$7,253	\$1,121

TABLE 9: AVERAGE ACCOUNT AND CONSUMER CHARACTERISTICS OF FREQUENT OVERDRAFTERS, BY OPT-IN STATUS

	Not Opted In	Opted In
End-of-day balance	\$615	\$607
Monthly deposits	\$2,977	\$3,385
Variability of monthly deposits	0.73	0.68
Monthly count of ACH deposits	2.0	2.2
Monthly count of POS debit card transactions	32.5	31.8
Credit score	583	579
Annual count of overdrafts (excluding NSFIs)	26.1	33.8
Annual count of overdrafts with fee (excluding NSFIs)	7.7	29.1
Annual count of overdraft fee reversals	0.7	2.5
Annual count of NSFIs	4.1	4.2
Annual count of NSFIs with fee	3.6	3.8
Annual count of overdrafts with fee on debit cards	1.5	19.3
Annual count of overdrafts without fee on debit cards	17.1	3.6
Monthly count of recurring debit card transactions	0.6	0.3
Monthly count of other debit transactions	11.1	8.6

TABLE 10: AVERAGE ACCOUNT AND CONSUMER CHARACTERISTICS OF FREQUENT OVERDRAFTER CLUSTERS, BY OPT-IN STATUS

	Cluster 1: Not Opted In	Cluster 1: Opted In	Cluster 2: Not Opted In	Cluster 2: Opted In	Cluster 3: Not Opted In	Cluster 3: Opted In	Cluster 4: Not Opted In	Cluster 4: Opted In	Cluster 5: Not Opted In	Cluster 5: Opted In	No credit score: Not Opted In	No credit score: Opted In
End-of-day balance	\$618	\$568	\$402	\$400	\$311	\$290	\$586	\$586	\$2,336	\$2,038	\$305	\$392
Monthly deposits	\$1,968	\$2,152	\$2,290	\$2,392	\$2,850	\$3,136	\$2,587	\$2,743	\$9,176	\$8,969	\$1,695	\$2,099
Variability of monthly deposits	1.24	1.19	0.69	0.67	0.56	0.53	0.63	0.60	0.57	0.56	0.80	0.74
Monthly count of ACH deposits	1.1	1.2	2.0	2.0	2.5	2.5	2.0	2.0	3.2	3.4	1.6	1.7
Monthly count of POS debit card transactions	26.4	24.2	31.9	30.0	32.0	30.1	32.0	30.8	55.2	55.6	26.7	25.1
Debit-deposit mismatch	1.8	1.8	0.3	0.4	0.3	0.3	0.4	0.4	0.4	0.4	0.6	0.6
Annual count of overdrafts (excluding NSF)	21.5	26.0	17.2	18.6	46.5	52.5	18.8	20.8	26.6	35.1	24.1	31.0
Annual count of overdrafts with fee (excluding NSF)	6.2	22.1	4.8	16.1	13.8	45.8	6.5	18.2	9.4	29.2	5.9	25.9
Annual count of overdraft fee reversals	0.6	2.4	0.6	2.4	0.9	3.3	0.7	2.1	0.7	1.5	0.6	3.1

	Cluster 1: Not Opted In	Cluster 1: Opted In	Cluster 2: Not Opted In	Cluster 2: Opted In	Cluster 3: Not Opted In	Cluster 3: Opted In	Cluster 4: Not Opted In	Cluster 4: Opted In	Cluster 5: Not Opted In	Cluster 5: Opted In	No credit score: Not Opted In	No credit score: Opted In
Annual count of overdrafts with fee on debit cards	1.3	15.4	1.0	11.1	2.7	30.1	0.9	11.2	1.1	16.1	1.6	19.9
Annual count of overdrafts without fee on debit cards	14.3	3.1	11.6	1.9	30.5	5.0	11.5	2.0	15.7	4.1	17.0	4.2
Monthly count of recurring debit card transactions	0.5	0.2	0.5	0.2	0.6	0.2	0.6	0.3	1.2	0.6	0.4	0.2
Monthly count of check debit transactions	1.6	1.6	1.6	1.9	2.0	2.2	2.6	2.8	6.9	7.1	0.8	1.1
Monthly count of ACH Debit transactions	3.1	3.0	2.7	3.0	3.5	4.0	4.9	5.5	11.3	11.2	1.3	1.8
Credit score	581	567	530	530	544	549	669	667	642	633	-	-

6.3 Accounts with fewer than six months of data

Table 11 compares the sample of accounts we analyze in the main body of the Data Point—those with at least six months of data—to accounts that do not have at least six months of data. It splits accounts with less than six months of data into three groups: (1) those that opened less than six months before the end of the sample period; (2) those that closed less than six months after the beginning of the sample period; and (3) those that were open for more than six months during the sample period but had fewer than six months with a transaction. Compared to accounts we kept in the sample, these excluded accounts show low and variable deposits and low end-of-day balances. The consumers associated with these short-tenure accounts have lower credit scores. They are also less likely to hold a general purpose credit card and have less available credit on the general purpose credit cards they do have.

Table 11 also compares the two groups of consumers without a credit score in our data—those that are matched to credit repository data but do not have a credit score (“thin file”), and those that were not matched to credit repository data (“no match”)—to consumers who do have a credit score. In general, “no match” consumers appear to be economically less established than “thin file” consumers, with lower monthly deposits, account tenure, and debit card transactions, although they have a slightly higher median end-of-day balance. Both groups of consumers without credit scores appear worse off financially than consumers with credit scores. Given the median ages of the groups, these statistics are consistent with consumers in these different groups being at different stages of their lives.

TABLE 11: SELECTED MEDIAN ACCOUNT AND CONSUMER CHARACTERISTICS BY TENURE STATUS

	Less than 6 months opened	Less than 6 months closed	Less than 6 months inactive	6 or more months, thin file	6 or more months, unmatched	6 or more months, with credit score
End-of-day balance	\$259	\$120	\$215	\$335	\$396	\$1,170
Monthly deposits	\$788	\$541	\$241	\$1,129	\$734	\$2,474
Variability of monthly deposits	0.85	1.20	1.47	0.72	0.94	0.60
Monthly count of ACH deposits	0.0	0.1	0.0	1.0	0.2	1.8
Monthly count of POS debit card transactions	6.5	3.6	1.3	12.3	7.8	9.8
Monthly count of recurring debit card transactions	0	0	0	0	0	0
Monthly count of other debit transactions	1	2	1	3	2	8
Months of tenure	1.0	13.0	23.0	26.5	16.0	65.0
Age	31.0	35.0	34.3	37.3	24.4	45.4
Neighborhood income	\$52,836	\$52,372	\$57,748	\$50,683	\$55,176	\$59,414
Credit score	611	610	690	-	-	710
Share with credit card	60.2%	61.3%	76.1%	12.2%	-	82.3%
Available credit on credit cards, if any	\$2,878	\$3,112	\$8,137	\$607	-	\$10,020

6.4 Cluster analysis method

To select a clustering approach, we began by considering a number of algorithms commonly used in the literature.⁴⁹ We then used each algorithm to divide the data into two, three, four, and five clusters. We did not consider clusterings with more than five clusters in order to keep the variation we study manageable. This process generated 16 clusterings we considered for discussion. To choose one, we scored each clustering according to five commonly used cluster scoring indices that measure mathematically how well the clustering fits the data.⁵⁰ We then ranked the clusterings according to the average value of these five numbers.

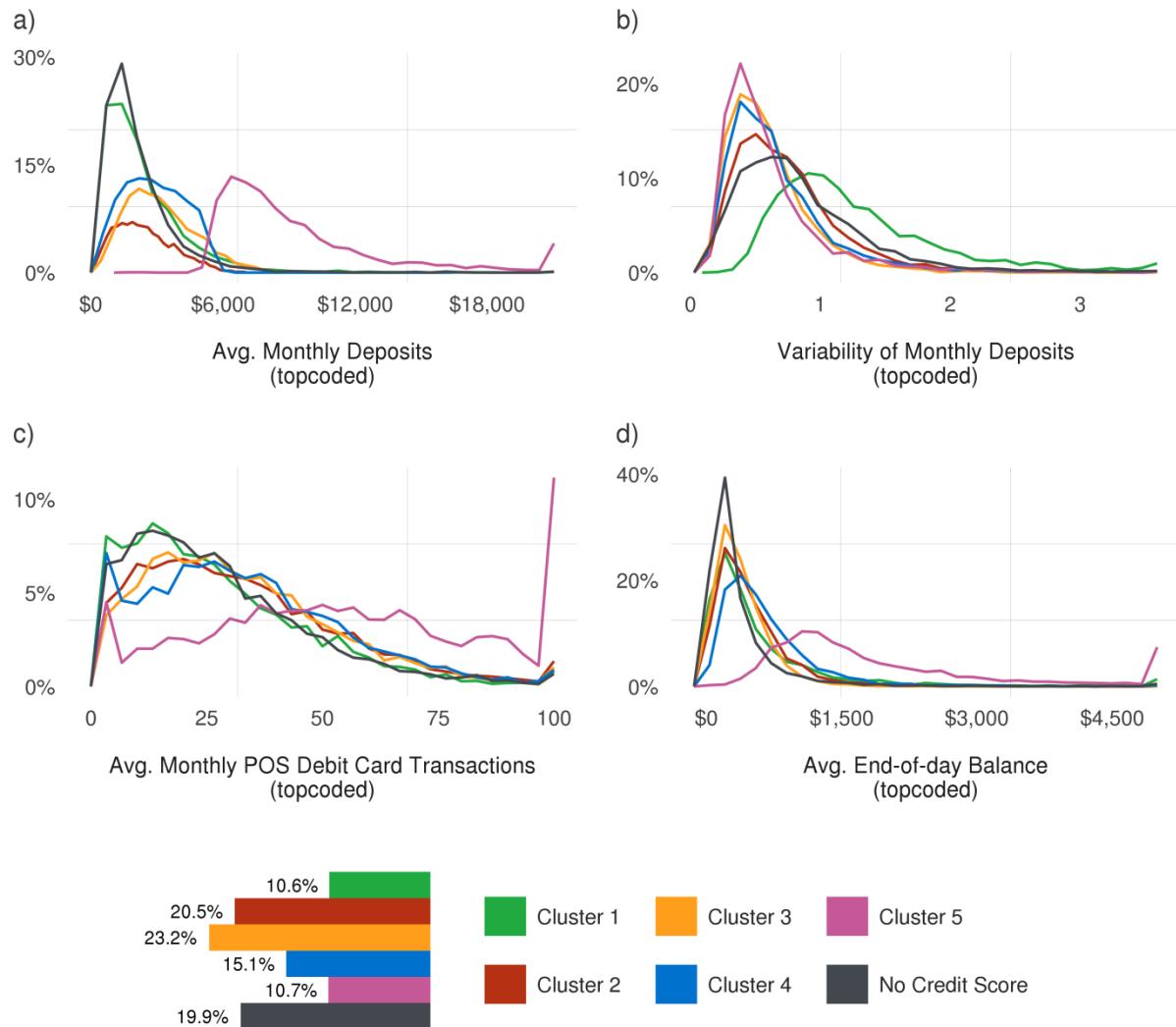
The clustering algorithm that fits the data best according to this approach is called k -means clustering. Given a set of variables, this approach groups observations into k clusters by minimizing the sum of squared deviations between observations in the same cluster. The algorithm starts with an initial grouping and reassigns observations across groups to achieve its minimization objective. For each choice of k , we ran ten thousand randomized trials of the algorithm to discover the initial grouping that maximized the index of the resultant clustering. The best-performing clustering had $k=5$.

In order to provide more information about the best-performing clustering beyond that given by Figure 4 and Tables 3 and 8, Figures 5 and 6 display the distributions of selected variables by cluster. They are very similar to Figures 1 and 2, except they present these distributions by cluster, not overdraft frequency group.

⁴⁹ The algorithms we considered were the “ k -means” clustering algorithm and the “ward”, “average”, and “complete” variants of the “hierarchical” clustering algorithm.

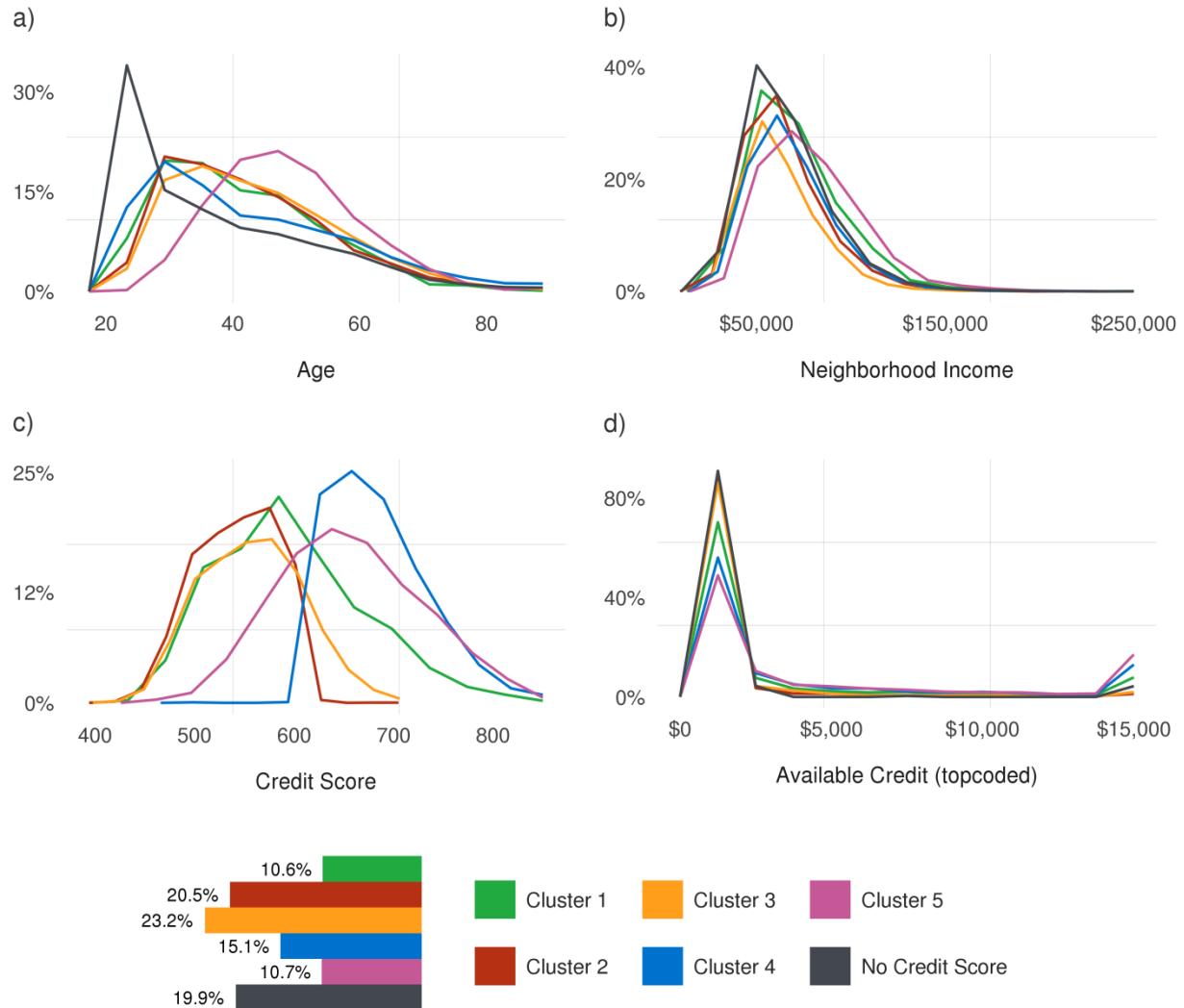
⁵⁰ The indices we used are the Silhouette, Scatter Density, Davies-Bouldin, Xie-Beni, and Calinski-Harabasz indices.

FIGURE 5: DISTRIBUTIONS OF SELECTED VARIABLES, BY CLUSTER



Note: Figures 5a) – 5d) show the distributions of different account characteristics among accounts in distinct groups. The bar graph shows the percent of frequent overdrafters in each group.

FIGURE 6: DISTRIBUTIONS OF SELECTED VARIABLES, BY CLUSTER



Note: Figures 6a) – 6d) show the distributions of different account characteristics among accounts in distinct groups. The bar graph shows the percent of frequent overdrafters in each group.

Although we used a rigorous approach that was deeply informed by the literature to choose our clustering, other rigorous and informed approaches may have chosen a different clustering. In order to investigate the sensitivity of our results to the clustering we chose, Table 12 presents selected median characteristics of frequent overdrafters by cluster, when we instead use the second-best clustering method according to our methodology. This clustering was produced by the *k*-means algorithm dividing the data into four clusters. The algorithm seems to keep the

previous Clusters 1, 3, and 5 largely intact as the new Clusters 1, 3, and 4, respectively. However, most accounts in the previous Clusters 2 and 4 now seem to be in the new Cluster 2.

TABLE 12: MEDIAN ACCOUNT AND CONSUMER CHARACTERISTICS BY CLUSTER WHEN USING FOUR CLUSTERS

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	No credit score
Share of frequent overdrafters	11.5%	27.4%	27.7%	13.6%	19.9%
Charge-off rate	13.0%	9.0%	4.3%	0.7%	8.8%
End-of-day balance	\$268	\$330	\$246	\$1,254	\$140
Monthly deposits	\$1,497	\$2,127	\$2,671	\$6,967	\$1,313
Variability of monthly deposits	1.06	0.58	0.48	0.44	0.68
Monthly count of ACH deposits	0.8	2.0	2.3	2.8	1.3
Monthly count of POS debit card transactions	21.5	27.6	27.6	48.3	22.2
Debit-deposit mismatch	1.47	0.36	0.28	0.29	0.43
Share opted in	28.3%	24.1%	36.3%	31.2%	24.5%
Annual count of overdrafts/NSFs	20	16	39	20	20
Annual count of overdraft/NSFs with fee	11	8	19	11	8
Annual count of overdraft/NSF fee reversals	0	0	1	0	0
Share of months with overdraft/NSF	50.0%	41.7%	75.0%	41.7%	54.5%
Negative balance period length (business days)	8.5	6.6	5.9	3.3	7.9
Monthly negative balance period fees	\$2.06	\$0.88	\$3.53	\$0.00	\$1.94

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	No credit score
Months of tenure	26.5	35.5	31.5	83.5	16.0
Age	36.3	36.3	38.3	43.3	29.3
Neighborhood income	\$55,906	\$54,092	\$52,401	\$64,912	\$51,119
Credit score	571	577	541	654	-
Share with credit card	54.3%	55.9%	39.9%	82.2%	6.0%
Available credit on credit cards, if any	\$347	\$270	\$72	\$1,640	\$60