



2018 Check Sample Survey

Appendix B: Technical Appendix

Federal Reserve Bank of Atlanta
Retail Payments Risk Forum Working Paper 20-1, Appendix B

Abstract: The technical appendix provides details about survey design, sampling, and analysis used for the 2018 Check Sample Survey (CSS) and accompanies the 2018 CSS report, data tables (appendix A), remotely created checks augmentation (appendix C), and check interrogation forms (appendix D). All data resources are available for download at frbatlanta.org.

JEL classification: E42

Key words: U.S. consumer check use, U.S. business check use, paper checks, personal checks, business checks, Federal Reserve Payments Study, Check Sample Survey

This work is a collaborative effort of staff in the Retail Payments Risk Forum at the Federal Reserve Bank of Atlanta and the GCI Analytics office of McKinsey & Company.

This version: July 30, 2020

Introduction

Section 1 of this report describes the survey objectives. Section 2 discusses the sampling plan, and Section 3, the weighting procedure. Section 4 explains how data was collected, and Section 5 explains how that data was used to categorize the items sampled according to payer, payee, and purpose. Section 6 describes additional data validity measures the team undertook to assure confidence in the conclusions and findings in this study, and Section 7 includes other notes that provide context for the data.

1. Survey objectives

Since 2001, the Federal Reserve’s Check Sample Survey (CSS) has periodically estimated the percentage shares of forward and return checks by payer, payee, and purpose, providing detail about the uses of paper checks during the time period that payments generally have been shifting from paper to electronic. This year’s CSS continues that work and, using a population and sampling plan implemented for the first time with this 2018 data, sets a benchmark for future data collection and analysis. The CSS continues to contrast business and consumer check usage patterns offering detailed data about the functions and circumstances wherein check persists as a payment instrument in this country. It should prove useful as a planning tool for those who use, process, and collect checks, including financial institutions and other check processors, billers, and even check writers. It may also be helpful for public policymakers.

2. Sample

For the years 2001 through 2016, the CSS relied chiefly—sometimes exclusively—on large commercial banks to provide the check sample used in the review. This effort ran concurrently with the Depository and Financial Institutions Payments Survey and placed an extra burden on some larger depository institutions (DIs). In addition, the data set was limited to the check processing and clearing circumstances particular to them. Using the Federal Reserve’s image archive made it possible to both reduce survey burden and gather a more robust and expansive data set. The 2018 data include a wider range of institutions—large and small credit unions and savings institutions as well as large and small banks.

Population: Federal Reserve forward and return checks

The population of forward checks is comprised of checks paid—that is, all of the forward checks were written as checks and paid as checks and none were converted to ACH transactions. All forward checks and all return checks processed were eligible for sampling.¹

Table B-1: Overall Federal Reserve and sampled volume, February 2018–January 2019

	Forward files	Forward items	Return files	Return items
Federal Reserve volume	1,526,429	4,791,404,251	1,036,635	29,461,853
Sampled volume	5,019	15,478,624	1,005	45,965

¹ Forward checks are items deposited with the Federal Reserve that the Fed subsequently clears and presents to paying depository institutions. Return checks are the reverse. That is, they are items sent to the Federal Reserve that paying depository institutions have chosen not to honor and that the Federal Reserve subsequently “returns” to the original collecting institution (where the checks were first deposited). These return items may include forward checks that the Federal Reserve Bank did not present and were originally collected through other banks or clearinghouses.

The 2018 CSS data set is a random sample of forward and return checks that the Federal Reserve processed from February 1, 2018, to January 31, 2019. The unit of observation is the individual check image.

In terms of representation of the U.S. check population, the Federal Reserve's forward-check processing volumes are estimated to be approximately one-third of total checks paid in the United States; total return-check processing volume is estimated to be about half of the total of U.S. return checks.

Sample selection

In line with previous CSS approaches, we sampled 55,000 forward checks and 10,000 return checks. The final sample sizes in the study were 54,609 forward checks and 9,785 return checks.

We used a three-step process to arrive at the analysis sample for forward checks:

1. Selection of files. From all forward check files sent for processing every business day between February 1, 2018, and January 31, 2019, we randomly selected and set aside 20 cash letter files.
2. Selection of checks from the previously selected files. At the end of each month, we pooled and randomly sampled items from the daily set-asides of forward files, which yielded 21,500 items monthly.²
3. Analysis sample. At the end of the 12-month period, we had collected and stored 258,000 forward items (21,500 X 12). We took a random sample of 55,000 checks and inspected the sample for duplicate items and missing metadata. We removed duplicates or items that were missing metadata, which yielded an analysis sample of 54,970 forward checks.

We used a two-step process to arrive at the analysis sample for return checks:

4. *Selection of files.* From all return check files sent for processing every business day between February 1, 2018, and January 31, 2019, we randomly selected and set aside four return cash letter files.
5. *Analysis sample.* At the end of the 12-month period, we had accumulated 45,965 return checks, and from these took a random sample of 10,000 checks. We removed duplicates or items missing metadata, yielding an analysis sample of 9,915 return checks.

Exclusions. As noted above, we eliminated some items prior to data interrogation because they were duplicates or had incomplete metadata. In total, these were less than two-tenths of a percent of the initial sample. We made other exclusions later during data interrogation for checks that could not be categorized by payer or payee. We deemed inconclusive for one or both of the necessary information elements (payer, payee or both) and so excluded from final counts 0.5 percent of forward checks and 1.1 percent of return checks. In addition, we excluded items when we did not reach a conclusive outcome as required at the third-stage interrogator step. Refer to table B-2.

² We took a monthly sample to minimize operational hardships and reduce long-term storage needs. A total of 5,019 forward files containing 15,478,624 items would have been amassed over the 12-month period had a monthly sample for forward items not been taken.

Table B-2: Sample selection detail as a proportion of Fed volume; exclusion details as a proportion of original sample

	Forward files		Forward checks		Return files		Return checks	
	Number	Distribution	Number	Distribution	Number	Distribution	Number	Distribution
Daily file sample	20	0.33%			4	0.10%		
Initial random sample			258,000	0.01%			45,965	0.16%
Original sample			55,000	0.001%			10,000	0.03%
Pre-interrogation exclusions								
<i>Duplicate items or missing metadata</i>			30	0.05%			85	0.85%
Post-interrogation exclusions								
<i>Inconclusive payer or payee</i>			271	0.49%			113	1.13%
<i>Inconclusive tie-breaker review</i>			90	0.16%			17	0.17%
Total exclusions			391	0.71%			215	2.15%
Total checks categorized								
			54,609	0.001%			9,785	0.03%

3. Sampling weights

It was operationally infeasible to pull a single sample from one large population of all Federal Reserve processed volumes at the end of the stated reference period, as footnote 2 explains. As a result, we did interim sampling. The daily set-asides were constant, regardless of daily volumes and deposit patterns, which can fluctuate throughout both weekly and monthly cycles. We applied weights to the forward and return check samples to better replicate an ideal sampling scenario.

- *Forward sample weighting.* We applied both seasonal and proportional weighting adjustments to the final forward sample:
 - Seasonality: We weighted the forward check sample to align with the seasonal fluctuation in Federal Reserve forward processing volume.

Table B-3: Seasonality forward weighting adjustments

Month	Weights on volume		Weights on value	
	Large files (≥10K)	Small files (<10K)	Large files (≥10K)	Small files (<10K)
Feb-18	0.94	0.94	0.88	0.88
Mar-18	1.06	1.06	1.08	1.08
Apr-18	1.06	1.06	1.15	1.15
May-18	1.08	1.08	1.12	1.12
Jun-18	0.99	0.99	1.18	1.18
Jul-18	1.01	1.01	0.95	0.95
Aug-18	1.01	1.01	1.32	1.32
Sep-18	0.91	0.91	0.84	0.84
Oct-18	1.07	1.07	0.92	0.92
Nov-18	0.93	0.93	0.80	0.80
Dec-18	0.97	0.97	1.00	1.00
Jan-19	0.99	0.99	0.94	0.94

- File size: We applied weights so that the distribution of the sample matched the proportion of checks within large (≥10,000 items/file) and small (<10,000 items/file) forward collection files, consistent with the overall forward check population that the Federal Reserve processes.

Table B-4: File size forward weighting adjustments

Month	Weights on volume		Weights on value	
	Large files (≥10K)	Small files (<10K)	Large files (≥10K)	Small files (<10K)
Feb-18	0.97	1.06	0.93	1.15
Mar-18	0.99	1.02	1.11	0.86
Apr-18	1.06	0.89	1.04	0.91
May-18	1.04	0.92	0.89	1.28
Jun-18	0.90	1.28	1.00	1.00
Jul-18	1.03	0.94	0.99	1.02
Aug-18	0.91	1.26	0.94	1.12
Sep-18	1.08	0.87	0.99	1.02
Oct-18	1.01	0.97	1.17	0.77
Nov-18	1.05	0.90	0.99	1.01
Dec-18	0.97	1.07	1.15	0.79
Jan-19	0.98	1.04	1.20	0.73

- *Return sample weighting.* We weighted the return sample for seasonality only since there were effectively no large return files and therefore no benefit in making the proportional weighting adjustment that we applied to the forward check sample.

Table B-5: Return weighting adjustments

Month	Weights on volume	Weights on value
Feb-18	1.08	0.89
Mar-18	0.99	0.69
Apr-18	0.50	0.74
May-18	0.75	0.84
Jun-18	5.18	3.24
Jul-18	1.09	1.68
Aug-18	1.14	0.87
Sep-18	1.80	1.66
Oct-18	2.50	1.95
Nov-18	2.19	1.74
Dec-18	0.83	1.07
Jan-19	0.46	0.50

4. Data collection

Data collected

The CSS analyzed checks based on the flow of funds. That is, we identified the payer and payee of a check according to who originally initiated the check or transaction (payer) and who ultimately received the check (payee). Two data collection forms (“CSS long form” and “CSS short form”) were used to guide our determinations about check usage (see appendix D).

Table B-6: Data considered for check categorization

Objective data, information observable by interrogator from the image of the check	
Payer-specific (nonexhaustive list; see the note below this table)	<p>Payer name, address and ZIP code</p> <p>Descriptive titles or abbreviations, such as trustee, estate, attorney, Corp., Inc., LLC, LTD</p> <p>Topical words or abbreviations, such as church, insurance, service(s), MD, DDS</p> <p>Key phrases paired with “check”: e.g., cashiers, official, WIC, payroll</p>
Payee-specific (nonexhaustive list)	Same identifiable elements as for payer
Check-specific (front of check)	<p>Date</p> <p>Serial or check number</p> <p>Dollar amount</p> <p>Signature characteristics</p> <p>Handwritten elements, such as driver’s license, state initials, phone number</p> <p>Other discernable items from MICR line including transit number and field identifiers (including the number 6 in position 44 of the MICR line)</p>
Check-specific (back of check)	<p>Same as for payer/payee-specific items noted above, appearing in the endorsement section of check</p> <p>Handwritten endorsement (or not)</p> <p>Endorsement perpendicular or parallel to printing/writing on the face of check</p> <p>9-digit endorsement number</p>
Interrogator’s subjective conclusions	
Payer-specific	Consumer, business, or government
Payee-specific	Same identifiable elements as for payer
If business payee	<p>Bill or POS</p> <p>Payee type (utility, grocery, charity, and so on)</p>
Scanned image data	
Check-specific	<p>Serial numbers</p> <p>Value</p> <p>Payer and payee banks’ transit routing numbers</p> <p>External processing code field</p>
Returns-specific	Return reason codes

Note: See appendix D for the check interrogation forms. The forms provide the complete set of criteria including objective and subjective determinations that were applied to conduct this study.

We did not retain any personally identifiable information (PII) for any checks after final data from interrogators were validated and summarized.

Interrogator procedures

A minimum of two staff members viewed or interrogated each sampled check. The first interrogator used the CSS long form, with 22 objective and four subjective assessments, and the second used the short form, which simply reported four subjective assessments of payer, payee, and purpose. If the subjective conclusions of the two interrogators did not match, a third interrogator completed the short form and was the tiebreaker for disputes. We used this three-interrogator procedure to correct keying errors, improve the confidence in the categorization of each check, and reconcile differences in the categorization.

Questionnaire changes from 2016

The 2018 CSS check interrogation forms included two new questions intended to improve the identification of remotely created checks and counterparty and purpose classifications.

Description of question for 2018	Purpose
<p>Does the number '6' appear directly to the left of the leftmost : symbol?</p> <ul style="list-style-type: none"> • Yes • No 	<p>Intended to increase precision in identification of remotely created checks</p>
<p>If the Payee is a business or government entity, how would you categorize the purpose of the payment?</p> <ul style="list-style-type: none"> • Bill payment or invoice payment • Point of sale payment • NOT a business or government • Cannot Determine 	<p>Intended to improve payee and purpose categorizations</p>

5. Check categorization

We categorized each check according to its payer, payee, and purpose, and used answers to objective and subjective questions described above in the determination. We classified payers and payees (as defined above) as either consumers (an individual, household, or small business) or businesses (including private-sector businesses and nonprofits as well as federal, state, and local governments).

There are five purpose categories, which fall into two general groups:

- Business
 - Bill payments
 - Payments at the walk-in point of sale (POS)
 - Indeterminate
- Consumer
 - Income (B2C)
 - Casual (C2C)

Table B-7 shows how the determination of payee and payer flows into the determination of the purpose classifications. For business payees, discernable evidence from interrogators’ observations drove conclusions about a given check’s purpose; that is, whether the check was used to pay a bill or make a purchase at the point of sale. If reviewers didn’t have consensus or clear evidence to support either of these two conclusions, they recorded the purpose for these checks as “indeterminate.”

Note that purpose categories are oversimplified. For example, “bill” covers a wide range of payments to myriad merchants. Point of sale transactions are similarly wide-ranging. Also, all B2C payments are defined as “income” but, as the table shows, this includes items other than straight payroll. Similarly, C2C payments are described as “casual.” We used broad categories (for example, “bill,” “income,” “casual”) because we could not make precise conclusions about the detailed purpose of a given check consistently and with confidence.

Table B-7: Categorization of purpose, as influenced by payee

Payer	Purpose	Detailed purpose
Business payee (B2B, C2B)		
Business	Bill	May include recurring retail bill (e.g., utilities), nonrecurring retail bill (such as medical), commercial bill payments (such as materials or equipment)
Consumer		
Business	Point of sale (POS)	Payments to a business payee that occur in a retail storefront
Consumer		
Business	Indeterminate	Unable to determine if bill payment was at POS or elsewhere
Consumer		
Consumer payee (B2C, C2C)		
Business	Income	May include payroll, pension, benefits/entitlements, rebate/promotional/refund, expense reimbursement, tax refunds, investment disbursements, bill payments to small businesses not distinguishable from consumers
Consumer	Casual	May include payments to family or friends, to purchase goods and services informally, to share a restaurant check or housing expenses, or to repay casual loan

Payer and payee categorization

As noted above, we identified the payer and payee according to who originated the check (payer) or transaction and who ultimately received the check (payee). In the case of a money order, for example, we did not consider the payer to be the money order vendor, such as Western Union. Similarly, when a bank created a check for an online bill payment, we considered the initiator of the transaction (consumer or business, but not the bank) to be the payer.

Information on the front of a check determined its payer type:

- **Business payers** were typically categorized as business based on the characteristics of the MICR line (for example, a federal government check’s MICR line begins with 000), whether the check was machine-printed or handwritten, and what were the characteristics of the payer name and address. For example, the payer’s name and address may contain such indicators as Inc., Corp., Department of, Accounts Payable, and so forth. Because business and government entities usually have access to the same set of payments mechanisms, we treated them collectively as business entities.

- **Consumer payers** generally include checks without the characteristics of business checks described above. While some small businesses or sole proprietors may use personal checks, we deemed acceptable the risk of miscalculation related to this practice, in part because the payments practices of these microbusinesses are more similar to those of consumers than to those of larger businesses.

Information on the front and back of the check determined payee type. Interrogators used the payee line to identify any obvious signs of a business payee, as described above for payers. The endorsement on the back of the check was also a significant determinant of payee type. Business payees tended to stamp or machine print their endorsements on the back of checks, for example. For a complete list of the objective and subjective characteristics we used to determine payer and payee categories, see the CSS long form in appendix D.

Purpose categorization

As table B-7 shows, identifying the payee is an important first step toward determining the purpose of a check.

For consumer payees, we determined the classification by the payer:

- We classified all payments written by businesses as income. Income includes not only payroll checks but also rebate checks, tax refunds, and stock dividends.
- We classified all payments written by other consumers as casual. This category could include payments to or from sole proprietors or small businesses—for example, rent payments to individual landlords.

For business payees, we determined the classification by further investigation:

- We used type of business organization paid to distinguish bills from POS purchases. For example, for payees that were credit card issuers or utilities, we classified the check as a bill payment. We classified payments to retail stores as POS.
- We used the presence of unique printing or stamps on the checks—such as a driver's license number, store number, terminal number— to classify checks as POS.
- Lockbox endorsements, apparent by their alignment across the length of the check in conjunction with terms such as "absentee" or "absent endorsed" indicated bill payment transactions.

6. Additional data review, final data analysis

The study team completed an extensive review of the initial conclusions made about classifications of payer, payee, and purpose to ensure that determinations were accurate. Aspects of this testing included:

- *Logic and consistency testing.* This review was independent of the initial interrogators. Payer, payee, and purpose classifications that had been made previously were reexamined along three lines as follows:
 - Inconclusive initial categorizations
 - Missing tiebreaker interrogations
 - Counterparty combinations that did not logically align with purpose classifications

We corrected and revised the results for any cases where we found categorizations or classifications to be in error. This step assured confidence in the research findings and also reduced the total number items that had to be excluded from the findings.

- *Audit of images and interrogation results.* Following the more comprehensive testing described just above, independent reviewers not involved in the initial data analysis and report findings pulled images at random. These

independent reviewers compared initial report findings to their own findings as an additional accuracy and validity check of the work.

- In total, we corrected 1,887 forward checks (approximately 3 percent of the forward sample) and 349 return checks (about 3 percent of the return sample) following the two measures described just above. We removed any checks that remained “inconclusive” from the final sample, as table B-2 described.

7. Other notes

- *Metadata priority.* For the first time with this CSS, the analysis relied on metadata exclusively for a wide range of inputs, which greatly diminishes the chance of error when recording objective data elements.³ Metadata produced through Federal Reserve check-processing operations provided details for check amount, check date, paying bank routing number, bank of first deposit routing number, serial number, return reason code, and auxiliary fields.
- *Remotely created checks.* The CSS detailed the approach used to identify and categorize remotely created checks (RCC).⁴ See appendix C, table C-1.
- *Bootstrap analysis.* Bootstrap analysis affirmed that the 2018 CSS sample mirrored the overall Fed check population.

[2018 CSS Report](#)

[Appendix A: Data Tables](#)

[Appendix C: Remotely Created Checks](#)

[Appendix D: Check Interrogation Forms](#)

³ In previous efforts, some metadata were supplied by participating institutions and relied on when possible. Metadata were not consistently available across the full sample of checks and so were not relied on as extensively as in this most recent CSS.

⁴ For example, a detailed breakdown of items are objectively noted as being RCCs, by virtue of having a the number 6 as the electronic product code identifier. In addition, a detailed breakdown of the various types of “signature authorization reference” phrases that are commonly employed by RCC users include an accounting of the instances of each. Of particular note are those items that employed the phrase “This is a bill payment draft.” These items may or may not have been originated by the paying bank but to the extent some were, they are not RCCs, according to the definition.