Research Data Reports

FEDERAL RESERVE BANK OF BOSTON™

No. 17-5

The 2012 Diary of Consumer Payment Choice: Technical Appendix

Marco Angrisani, Kevin Foster, and Marcin Hitczenko

Abstract:

This document serves as the technical appendix to the 2012 Diary of Consumer Payment Choice administered by the RAND Corporation. The Diary of Consumer Payment Choice (DCPC) is a study designed primarily to collect data on financial transactions over a three-day period by consumers over the age of 18 in the United States. In this data report, we detail the technical aspects of the survey design, implementation, and analysis.

Keywords: survey design, samples election, raking, survey cleaning, poststratification estimates

JEL Codes: D12, D14, E4

Marco Angrisani is an associate economist at the University of Southern California Dornsife Center for Economic and Social Research. Kevin Foster is a survey methodologist and Marcin Hitczenko is a statistician; both are members of the Consumer Payments Research Center in the research department of the Federal Reserve Bank of Boston. Their respective e-mail addresses are marco.angrisani@usc.edu, kevin.foster@bos.frb.org, and marcin.hitczenko@bos.frb.org.

This paper, which may be revised, is available on the website of the Federal Reserve Bank of Boston at http://www.bostonfed.org/economic/wp/index.htm.

The Diary of Consumer Payment Choice is a product of the Consumer Payments Research Center (CPRC), housed in the research department at the Federal Reserve Bank of Boston.

The authors thank their colleagues and management in the CPRC and the Boston Fed research department. In addition, we thank the management and staff at CESR and the RAND Corporation. From the Boston Fed: Tamás Briglevics, Sean Connolly, Claire Greene, Vikram Jambulapati, Adam Karabatakis, Suzanne Lorant, William Murdock, Scott Schuh, Oz Shy, Joanna Stavins, and Bob Triest. From CESR and the RAND Corporation: Tania Gursche, Arie Kapteyn, Bart Orriens, and Bas Weerman. Finally, the authors acknowledge John Sabelhaus and the staff of the Survey of Consumer Finances at the Federal Reserve Board of Governors for their advice and mentorship. Geoff Gerdes and May Liu from the Board also shared advice and knowledge.

This version: September 2017

CONSUMER PAYMENTS Research Center

Contents

1 Introduction

This document serves as the technical appendix for the 2012 Diary of Consumer Payment Choice (DCPC), a survey designed and sponsored by the Consumer Payment Research Center (CPRC) at the Federal Reserve Bank of Boston (Boston Fed), along with the Cash Product Office at the Federal Reserve Bank of San Francisco and the Federal Reserve Bank of Richmond. The DCPC is the second payments survey developed by the CPRC, the first being the Survey of Consumer Payment Choice (SCPC). Much like the process used for the 2012 SCPC, the programming, respondent recruitment, management and data collection for the 2012 DCPC was contracted to the RAND Corporation (RAND). The SCPC has been fielded annually since 2008. Annual reports, data, and tools for data use are available on the SCPC website.[1] Similarities and distinctions between the data product of the SCPC and DCPC are discussed throughout this document.

All materials related to the DCPC, including the public use dataset and the questionnaires, are available on the DCPC website.[2] In particular, the "Guide to the 2012 Diary of Consumer Payment Choice" provides details for every variable in the DCPC dataset. Research papers using the DCPC data include Samphantharak, Schuh, and Townsend (2017), which provides a thorough discussion of how the SCPC and DCPC relate to one another and to other household financial surveys. Additionally, a paper discussing the use of the DCPC to estimate cash use and a report summarizing other key findings in the DCPC are forthcoming (Greene, O'Brien, and Schuh 2017; Schuh and Stavins 2017).

The organization of this work follows the natural, chronological progression of considerations involved in conducting and analyzing a survey. We begin by establishing the context and goals of the survey in Section 2. In Section 3, we briefly discuss two DCPC pilot studies conducted in 2010 and 2011, with a focus on a few key findings relating to survey methodology that influenced the implementation of the 2012 DCPC. Section 4 details the sample selection strategy for the DCPC and presents statistics relating to diary response and completion. Section 5 delineates the methodology used to generate sample weights that can be used to generate estimates for the population of U.S. consumers. Section 6 describes the preprocessing done in order to produce the public use dataset. Finally, in Section 7, we present the statistical methodology used for generating population estimates based on the DCPC data.

[1] http://www.bostonfed.org/economic/cprc/SCPC
[2] http://www.bostonfed.org/economic/cprc/DCPC

2 Survey Objective, Goals, and Approach

In this section we describe the overall objectives, goals, and general approach to survey design of the DCPC, especially as they relate to those of the SCPC. In particular, we explain the choices made in selecting the unit of observation and the interview mode. In both cases, the CPRC seeks to comply with best survey practices, subject to constraints imposed by budget and resources.

2.1 Survey Objective and Goals

The DCPC was conceived as a natural extension of the SCPC. The SCPC is a 30-minute online survey fielded at the end of September and in early October that asks respondents about payment instrument preferences, adoption, and frequency of use. The DCPC is also a study of payments, but it focuses on the dynamics of payment instrument choice at the transaction level. It is dubbed a "diary" because respondents record details of their financial transactions, including all payments, on a daily basis over a three-day period around the month of October.

One major contribution of the DCPC is that it offers a new line of insight into payment instrument choice by collecting a different array of information than the SCPC. Perhaps the most useful component unique to the DCPC is knowledge of transaction-specific information, such as merchant and dollar value associated with each payment instrument use. Information combined from the two surveys provides the most insight into a respondent's payment behavior. For example, a complete model of payment instrument choice might reasonably benefit from knowledge of not only the transaction-level characteristics provided by the DCPC but also the consumer's personal preferences and adopted set of payment instruments, collected in the 2012 SCPC but not the 2012 DCPC. Note that the DCPC does not give a full picture of payment instrument adoption, since the lack of use of a payment instrument within the diary period does not imply non-adoption of that instrument.

A second contribution of the DCPC is to gauge the reliability of certain SCPC variables. Certain economic concepts are measured in both surveys, albeit through different data structures, thus requiring alternative procedures to generate population estimates. Most notably, both surveys can be used to determine the frequency of payment instrument use or cash withdrawals. The SCPC asks for a self-reported rate, such as the number of monthly payments made, through recall, while the DCPC has respondents record a portion of each individual's monthly behavior. Each dataset requires different statistical ways of pooling the information in order to generate a population-wide estimate. Comparisons of SCPC-based estimates and DCPC-based estimates provide some insight into the accuracy and efficacy of different data collection and estimation strategies.

2.2 Unit of Observation

The unit of observation and sampling for the DCPC is the consumer day. The DCPC is primarily designed to make population estimates corresponding to those prevailing in the month of October. However, the suspected loss in data quality that would occur due to respondent fatigue or respondents dropping out mean it is impractical to ask respondents to track all daily transactions for an entire month. Indeed, research has found that diary fatigue generally settles in after a few days (Ahmed, Brzozowski, and Crossley 2006; Jonker and Kosse 2009; Schmidt 2011). Partly for this reason, most consumer diaries last between a day (Netherlands) and a week (Germany, France, Austria, and Australia), although the UK Payments Council Survey uses a four-week survey. Of course, even if data quality were not a potential concern, under a fixed budget, the cost of incentives involved in tracking respondents for an entire month precludes fielding the survey to a larger, more diverse sample of respondents. The DCPC relies on three-day diaries in an attempt to balance the quality of data collected with respondent burden.

Just as it is important to sample a broad swath of consumers, it is also important to have data from all parts of the month, as there may be temporal heterogeneity of economic behavior. Overall, it is best to observe a representative mix of consumers uniformly throughout the month. Designing a scheme for collecting data involves not only selecting who should participate but also over which days. Overlapping, three-day waves of respondents are distributed uniformly, but at random, throughout the entire month, for complete temporal coverage.

In order to most easily link the SCPC and the DCPC, we continue to rely on the consumer within the unit of observation, as we have done in the SCPC. This choice means that each respondent reports only his or her own transactions in the DCPC. A natural alternative would be a household, which is the fundamental unit of observation in the Survey of Consumer Finances and the Consumer Expenditure Survey. Beyond a desire to be consistent with the SCPC, the choice of the consumer is appropriate for both logistical and theoretical reasons.

Household surveys, which collect data on all transactions made by anyone in the household, naturally require a greater coordination of survey responses. Ensuring that all household members participate likely takes considerable effort, and facilitation by in-person interviews would drive up data collection costs significantly. Again, under a fixed budget, sampling households rather than consumers would mean that fewer households are observed in the sample, potentially reducing the diversity of observed behavior.

In addition, for many economic concepts covered in the DCPC, we argue that a per-consumer questionnaire is likely to yield more accurate data. Asking one household member to collect information about all transactions from other household members could lead to under-counting, as individuals may not feel comfortable sharing details of all purchases with one another or simply because gathering reliable information requires consistent communication. Transactions of a certain kind, such as small-value cash transactions, may be affected disproportionately, because they are easier to forget. Cash use, in particular, is of great interest to the CPRC, and it is not measured elsewhere.

For other variables, most notably for the payment of bills or other expenses more closely associated with a household than with an individual, the consumer as a unit of observation may not be ideal. Many such payments are scheduled to be automatic and often come out of joint accounts or pooled resources. As a result, it can be difficult to attribute responsibility for such payments, potentially leading to measurement bias in the form of under-counting if the payments are not reported at all or to double-counting if several household members each claim responsibility for the same payment. In such a case, the framework of a household payment might make more sense.

To study intra-household dynamics, the DCPC and the SCPC samples include some respondents from the same household. The presence of households with several sampled members does allow considerable insight into forms of under- and over-counting. Research based on these multi-sample households in the SCPC suggests that survey respondents are more likely to have a higher share of financial responsibility within the household than would be expected if household members were selected at random, and thus respondents may tend to be more likely to make certain types of payments than an average sample of the population (Hitczenko 2015). Treating such a sample as representative of all consumers may lead to overestimation of the number of bills paid. To accurately measure bills, it might be better to ask about the entire household's bill payment behavior. Nevertheless, for consistency within the survey instrument, the DCPC asks respondents to estimate only the number of bills that they physically pay themselves, either by mail, by phone, online, or in person.

2.3 Interview Mode

The DCPC is a computer-assisted web interview (CAWI). This mode of interview fits best with our sampling frame, the internet-based American Life Panel (ALP). All ALP members are given internet access upon recruitment into the panel. The survey instrument is programmed in the MMIC survey system, developed by the RAND Corporation, and compatible with all web browsers.[3]

The nature of the DCPC makes an online survey a natural choice for implementation. In-person interviews, beside being more expensive to conduct, would be difficult to implement because respondents are expected to record certain data, such as cash holdings, at the end of the day. Using a CAWI allows respondents to log on at the end of their day and enter all relevant information. In addition, there is some evidence that sensitive information, as some might consider certain purchase details, are more likely to be reported if administered through an online survey (De Leeuw 2005).

A second alternative, a paper survey, is also not practical, because the survey is designed to ask follow-up questions about transactions based on the information provided about the transaction. For example, the use of cash may lead to a different set of questions than if a credit card were used. This type of skip logic fits easily within the CAWI framework, but would be difficult with a paper survey.

Although official data are collected online, respondents are encouraged to keep track and record details of daily transactions through other means. To this end, the respondents are sent two paper memory aids and a pouch in which they can keep receipts. The larger memory aid is the size of a folded $8.5'' \times 11''$ sheet of paper and includes instructions, examples, and response categories for all transactions types. The second memory aid is the size of a checkbook and only provides space to record basic information about each transaction. The use of the memory aids is not required nor are they collected after the DCPC is complete. Despite this, out of 2,470 respondents, 1,341 (54.3 percent) used the memory aids (426 used the large one only, 662 used the small one only, and 253 used both). However, the statistics on memory aid use, along with the fact that 1,763 respondents (71.4 percent) said they relied, at least partially, on the collection of receipts, suggests that, unlike in the SCPC, much of the data entry in the online questionnaire is not based on recall.

[3]MMIC stands for Multimode Interviewing Capability. More information on MMIC is available at http://www.rand.org/labor/mmic.html.

2.4 Public Use Datasets

Users who are interested in downloading the original, unprocessed datasets can obtain these from the RAND Corporation's website devoted to the American Life Panel.[4] Interested users must create a username and password to download data from the RAND website. These data contain only the survey variables found directly in the survey instrument itself. These survey variables have not been edited or processed. For those interested in using these data, the simplest way of identifying variables is by finding them directly in the DCPC questionnaire, which can be downloaded as a pdf document from the Boston Fed's SCPC website.[5]

The CPRC also offers a processed version of the data, available in Stata, SAS, and CSV formats from the DCPC website.[6] The processed dataset is restricted to the vast majority of respondents who completed all three days of the diary. The two datasets also differ in how they are organized. Each row in the processed dataset represents either an indicator of participation for a particular day or a single transaction (the indicator of participation makes it easy to identify days on which a respondent had no reported transactions). In the original dataset, one row contains information about all transactions made by a particular consumer on a particular day. The reorganized dataset is easier to manipulate and sort according to transaction types. A second difference between the processed dataset and the raw dataset is that certain variables have been edited in the former in an effort to improve the quality of the data. This process is detailed in Section 6.4. Finally, demographic variables collected by RAND's My Household Questionnaire, and not featured in the DCPC itself, are appended in the processed dataset. We recommend referencing the companion document, "Guide to the 2012 Diary of Consumer Payment Choice," which is also available at the DCPC website. The user's guide gives a brief description of each variable, indicating the diary question or questions that define it as well as all response options.

One variable of note is the variable `prim_key`, which serves as the unique identifier for each respondent. This variable is used as the primary key for both the RAND and the Boston Fed datasets and can be used to merge the data with any other dataset from RAND's American Life Panel. In particular, `prim_key` can be used to merge the DCPC dataset with SCPC data.

[4] https://alpdata.rand.org/
[5] http://www.bostonfed.org/economic/cprc/SCPC
[6] http://www.bostonfed.org/economic/cprc/DCPC

3 2010/2011 DCPC Pilots

The 2012 DCPC is the first official version of the diary survey. However, in anticipation of the 2012 DCPC, the CPRC conducted two pilot studies of the diary in 2010 and 2011, with respective sample sizes of 353 and 376. The fundamental structure of the pilot studies was very similar to that of the 2012 version, with three-day periods randomly assigned throughout the month of October. In 2010, the pilot served primarily to test the plausibility of fielding and managing the diary for an entire month. The unique structure of the DCPC, which involves assigning dates up to a month in advance and mailing memory aids throughout the month, presents unique challenges for management of the process and proper data collection. We concluded that a three-day diary over the month of October was manageable and yielded data of appropriate quality. In 2011, based on the success of the 2010 pilot, the focus was on studying the effects of various aspects of diary design on the quality of data. Some analyses were done through formal experiments, while others were based on observational data.

As we are interested in linking the two surveys, we were interested in ascertaining whether the order of surveys affected responses in the DCPC in any way. A particular concern is that the SCPC, which asked respondents to consider their financial and payment habits, might affect behavior, leading to different reported tendencies in the ensuing DCPC. In 2011, all DCPC respondents were asked to take the SCPC as well. The DCPC diary dates were assigned to respondents, but the SCPC was not, with a general release at the end of September. While not an experiment in that treatments were not assigned, 182 respondents took the SCPC before the DCPC, while 194 took the DCPC first. Analysis focused on the distributions of the number of reported payments. A comparison of distributions was done using a Kolmogorov-Smirnov test (Daniel 1990), and we found no significant difference in distributions between groups defined by the order in which the surveys were taken(p-value = 0.54). As a result, in 2012, we did not enforce a schedule by which respondents had to complete the SCPC and DCPC relative to each other, simplifying administration of the surveys considerably.

One official experiment conducted in the 2011 DCPC pilot involved studying response differences between those respondents who had had diary participation experience in the form of the 2010 DCPC, and those who were participating for the first time. To test this, roughly half of respondents in 2011 (203) were recruited from among those who participated in 2010, while the other half (173) were recruited at random from the subset of respondents who were recruited to take the 2011 SCPC, but had never participated in the DCPC. Again, the focus of the analysis was on the number of transactions reported, and again, using the

Kolmogorov-Smirnov test for distributions, we found no evidence of a difference between the two groups (p-value = 0.99). This result gives us confidence that a longitudinal study over years will provide reliable data and will not be diminished by learned behavior.

A second experiment conducted in 2011 focused on the use of memory aids. All respondents were mailed a package with memory aids that can be used to help record transactions throughout the diary period. In 2010, we asked all respondents to use and mail back the memory aids, and we found few discrepancies between the respondents' online entries and the paper versions. As a result, a decision was made not to require the use of the memory aids. However, the 2011 pilot fielded an experiment to see whether requiring individuals to mail back the memory aids, regardless of use, would affect responses. Thus, 81 respondents were asked to mail back the memory aids, while 295 were not. The discrepancy in treatment group sizes comes from the fact that we did not want to ask those who had participated in 2010 and thus had to mail the memory aids back, to not do so in 2011. In order to be able to distinguish the effects of memory aid protocol from the effects of being a new diarist, about half of new diarists were assigned to each treatment group. One potential hypothesis might be that the requirement to mail back the paper diary effectively serves as a strong suggestion to carry and use the memory aid, perhaps leading to more complete data entry. In practice, the Kolmogorov-Smirnov test found that there was no real difference in the number of transactions reported (p-value = 0.81) or the proportion of transaction details filled out (p-value = 0.98) between the two groups. Due to the lack of evidence that sending back memory aids led to more diligent recording, the 2012 DCPC did not require respondents to mail back memory aids.

4 Data Collection

Once the survey instrument is finalized, the collection of data involves two general steps: sample selection and administration of the survey. The strategies and philosophies adopted in each step for the 2012 DCPC are described below. In addition, summary statistics related to survey completion are detailed. Similar expositions focusing on the SCPC (Angrisani, Foster, and Hitczenko 2013; 2014; 2015) can be found on the SCPC website.[7]

[7] http://www.bostonfed.org/economic/cprc/SCPC

4.1 American Life Panel

As with the SCPC, all DCPC respondents are recruited from RAND's American Life Panel (ALP). The ALP commenced in 2003 as a panel of approximately 500 members, with the original intent to study the methodological issues of internet-based surveys among the older population. As a result, until 2006 all recruits into the ALP were over the age of 40. Since then, the ALP has expanded to also include individuals between the ages of 18 and 39 and has grown considerably in size. At the time of the 2012 DCPC sample selection (the end of September 2012), there were roughly 5,000 panelists.

There are several pathways by which individuals are recruited into the ALP, but from a survey methodological point of view these condense into two general strategies. The first strategy involves recruiting volunteers from households that are not yet represented in the ALP. At first, RAND did this by gathering volunteers from other, already-established panels, such as the University of Michigan Internet Panel Cohort (http://www.sca.isr.umich.edu/) and the National Survey Project Cohort (terminated in 2009). In more recent years, potential subjects have also been recruited via address-based sampling directly by RAND. Most notably, in 2011, around 2,000 panel members from ZIP code areas with high percentages of Hispanics and low-income households were added to the ALP (labeled by RAND as the "Vulnerable Population Cohort"). The second strategy involves asking individuals already in the ALP to recommend acquaintances or fellow household members to participate in ALP-distributed surveys. In 2012, members who were lone representatives of their households represented about 75 percent of the ALP cohort.

In its early stages, due to its original research intentions, the ALP was not demographically representative of the U.S. population of adults. The pivot to represent all U.S. adults naturally improved demographic coverage significantly, but certain strata, most noticeably men with lower incomes and lower levels of eduction, remained under-represented. This may be explained partly by nonuniform eagerness to join the ALP across demographic strata. Recruitment rates ranged from around 30 percent to approximately 50 percent, meaning heterogeneity across strata could easily produce a cohort that differs substantially from the pool of invitees. Nevertheless, largely due to targeted efforts by RAND, the overall representativeness relative to the 2012 Current Population Study (CPS) with respect to a variety of demographic variables has been improving. More information about the American Life Panel can be found at the website http://mmic.rand.org/alp.

ALP members remain in the panel unless they formally ask to be removed or stop participating in surveys over a prolonged period of time. At the beginning of each year, RAND contacts all members who did not take any survey for at least a year and removes them from the panel, unless they explicitly declare continued interest in participating. Since inactive members are removed only once a year, the pool of those invited to answer the survey at a given point in time may include inactive members. Nevertheless, the annual attrition rate is roughly 10 percent, so the proportion of such cases is likely to be relatively small.

4.2 DCPC Sample Selection

Sample selection for the 2012 DCPC was a relatively complex process intended to balance several competing goals for the final sample. The budget available to the CPRC in 2012 was allocated generally for the SCPC and the DCPC, so sample selection design must be considered for both surveys jointly. Below, we outline the priorities of the sample selection and provide details of respondent recruitment.

Due to an interest in generating population estimates based on DCPC variables alone, the primary goal of the CPRC, with respect to the DCPC, was to have the set of individuals participating in the DCPC be as representative of the U.S. population as possible. A key tenet of the SCPC recruitment strategy was to invite all past SCPC respondents to participate in the 2012 SCPC (though not necessarily in the 2012 DCPC) in order to preserve longitudinal aspects within the annual SCPC samples. The benefits of a longitudinal panel, namely, the added power associated with tracking trends at the individual level, have been well discussed (Baltagi 2008; Duncan and Kalton 1987; Frees 2004; Lynn 2009). For many research agendas, it is advantageous to base results on a longitudinal panel rather than on a sequence of cross-sectional studies. As discussed in previous sections, the DCPC naturally links with the SCPC. Therefore, the secondary goal was to have as many respondents as possible take both surveys. Doing so not only helps build a joint SCPC-DCPC longitudinal panel in the future, but provides a richer history of payment instrument information from the SCPC in analyzing the 2012 DCPC data. For the same reasons, our third priority was to include as many people as possible who participated in the pilot DCPCs in 2010 and 2011.

Taking into account the associated cost as well as the expected demographic composition of the SCPC panel, we estimated that we could afford around 2,500 respondents who would take both the SCPC and the DCPC in 2012. To prioritize representation, we chose 15 demographic categories with which to match our sample to U.S. proportions as measured by the Current Population Survey (CPS), administered in March. The demographic strata are shown in Table 1. The next step involved selecting enough individuals into each stratum so that the final sample has an expectation of at least 2,500 respondents once non-participation is taken into account. The determination of appropriate recruitment sizes was done by RAND.

Table 1: Strata used in 2012 DCPC sample selection.

Stratum	Race	Age	Income	Stratum	Race	Age	Income
1	White	18−39	<\$30K	10	Non-white	18−39	<\$30K
2	White	18−39	\$30K-\$60K	11	Non-white	18−39	≥\$30K
3	White	18−39	≥\$60K	12	Non-white	40−55	<\$30K
4	White	40−55	<\$30K	13	Non-white	40−55	≥\$60K
5	White	40−55	\$30K-\$60K	14	Non-white	56+	<\$30K
6	White	40−55	≥\$60K	15	Non-white	56+	≥\$30K
7	White	56+	<\$30K				
8	White	56+	\$30K-\$60K				
9	White	56+	≥\$60K				

For each stratum, as many individuals as possible are selected from the pool of respondents who have taken the SCPC in the past. Among these, preference is given to those who have taken the SCPC more recently and then to those with more years of participation. Thus, those who participated in 2009, 2010, and 2011 are the most coveted. Additionally, the preference for continuity translates to greater desirability of someone who participated only in 2011 than the desirability of someone who participated in 2009 and 2010 but not in 2011. If individuals share the same pattern of past participation for the SCPC, we prioritize first individuals who participated in both DCPC pilot studies, then those who participated in just the 2011 DCPC pilot, and, finally, those who participated in just the 2010 DCPC pilot. The remainder of slots were filled by members of the ALP who had never participated in the CPRC surveys. The new respondents mostly filled the strata corresponding to non-white, younger, and lower-income individuals. In fact, many came from a subset of newly recruited panelists representing ZIP codes identified as having higher proportions of Hispanic and low-income residents.

4.2.1 Diary Day Assignment

The 2012 DCPC is fielded predominantly in the month of October. The desire to standardize this response period is three-fold. First, from an analytical point of view, trends from year to year are identified more easily if differences in behavior are not attributable to seasonal behavioral variation. Second, from an economic point of view, the month of October is a reasonably representative month with respect to certain economic variables such as employment or sale volumes; it includes no major holidays and falls between summer and winter. Third, responses from both surveys can be linked more naturally if they correspond to the same period of economic activity.

Not only do we want the number of diarists to be uniformly distributed throughout the month of October but we also want the number of diarists responding on each of the three diary days to be uniform. This makes comparing results across days methodologically simpler and is useful since different diary days collect different types of information in addition to the daily transactions. For example, information about bills is collected on the third diary day, so it would be a mistake to omit fielding the Day 3 module on October 1st, a date on which many bills are likely to be scheduled.

In order to balance the number of individuals on any given day of October who are participating on each of the three diary days, we start diary periods on September 29th and continue through October 31st. By doing so, we ensure that although there are diaries being taken in late September and early November, the number of responses for each diary day for every day in October is the same. This yields 33 diary waves and, with 2,500 respondents, about 76 people per wave. Under ideal sampling and participation conditions, we would expect about 227 individuals on each day in October.

The DCPC recruitment process involved randomly assigning a three-day diary period to each individual prior to the period of fielding the diary. The set of respondents invited to participate in the DCPC received an email with a link to a page explaining the structure of the DCPC as well as the pre-assigned diary dates and were asked to consent to participation. Because the CPRC is only interested in money spent in the United States, the DCPC invitation asked respondents to confirm that they were not traveling abroad during the assigned diary period. In date assignment, an important principle was not to allow respondents to change to different diary periods, as this request may be driven by periods of the month that are relatively more convenient and thus introduce bias in responses. In fact, individuals who were travelling domestically during the diary period were strongly encouraged to participate, and RAND offered to extend the deadline for online data entry for that subgroup.

As a simpler survey to manage, the SCPC does not require special invitation. Instead, it was released directly to potential respondents, meaning notice of a new survey, along with a link, was sent to all selected panelists on September 26th. This set of people includes all those who agreed to take the DCPC. Respondents are allowed to begin the SCPC at any point after receiving the SCPC link.

Because diary day assignment is done in advance (for those taking the diary in late October, almost a month before the survey), reminder emails are sent to individuals several days before the first day of the diary. These reminders include a link to an instruction video that explains the basic goals and methodology of the diary. Finally, RAND mails packages three days before the beginning of the diary period with two versions of paper memory aids. The use of memory aids is not required, but they can be useful and serve as a second reminder of the upcoming diary period.

4.3 Participation

In 2012, the recruitment process yielded 2,547 individuals who participated in the diary in some form. Participation can be defined in a variety of ways, but we define it to mean that the respondent logged on to the online questionnaire module on at least one day. The individuals represent 2,275 households: 2,038 households with one representative, 208 with two, 23 with three, and six with four. Because of high participation rates, the sampling design led to a high degree of overlap with the set of SCPC respondents. All but 52 individuals participated in the 2012 SCPC as well. Table 2 shows how many respondents participated in the DCPC for several multi-year SCPC panels. For example, there are more than 1,100 individuals who participated in the 2012 DCPC for whom we have four years' of SCPC data from 2009 to 2012. Table 3 shows the number of respondents in the 2012 DCPC who also participated in the 2010 or 2011 pilots. A majority of those who participated in 2010 or 2011 were included in the 2012 sample as well, with 143 individuals having taken the DCPC in all three years.

Table 2: Size of various continuous SCPC longitudinal panels and the number of each who took the 2012 DCPC.

	SCPC Panel			
	2009 – 2012	2010 – 2012	2011 – 2012	2012 Only
Size of Panel	1,250	295	116	1,515
# in 2012 DCPC	1,109	215	67	1,021

Source: Authors' calculations.

4.4 Completion

Overall, completion rates for the SCPC and DCPC were high. Completion of a diary day can be measured in a variety ways, but we focus on those who log into the online questionnaire

Table 3: Number of 2012 DCPC respondents by participation history in 2010 and 2011 DCPC pilots.

Participated in 2010 DCPC	Participated in 2011 DCPC	# Individuals	# Who Participated in 2012 DCPC
Yes	Yes	200	143
Yes	No	153	63
No	Yes	175	121
No	No	NA	2,220

Source: Authors' calculations.

and then log off. Of course, it is possible that individuals may have entered information and simply not logged off, so these are conservative estimates of the percentage of respondents who entered diary information. Of the entire sample, 2,474 individuals started three days of the diary (by logging on); all but four of these individuals completed all three days. Only 22 individuals started only one day, and all of them completed it. Of those who logged on only two days, nine finished both days and seven finished only one day.

Figure 1 shows the number of respondents logging on each day over the diary period in the Fall of 2012. Optimal survey design calls for the same number of respondents for each diary day on all days in October. As is apparent from Figure 1, this ideal has not quite been met. Random fluctuations across days are to be expected, but Figure 1 reveals a noticeably larger number of respondents in the second half of the month. Part of the explanation for this may be that respondents forgot to fill out the diary on their given day and did so at a later date. This might also explain the tail into November, as some individuals may have filled out the online diary after receiving reminders from RAND. Figure 2 depicts the distribution of the number of days between the date on which respondents filled out the online diary and the date on which they were assigned. Note that if an individual were to log on at 1 A.M. on the morning after the assigned day, that individual would effectively be completing the diary as intended, but the metric in Figure 2 would record that as a delay of one day. We find that 48 percent of respondents complete all days as assigned, while 79 percent do so within one day. However, there are also cases of significantly longer lags, with close to 4 percent of diary days being recorded 10 days after the assigned date. Information regarding assigned diary days as well as time stamps for first log on and last log off are provided in the dataset, so respondents with excessive lags may be discarded as researchers see fit.

Figure 3 shows the distribution of the number of days between the date when respondents took the SCPC and the first day of the DCPC. The average absolute difference in days is 13,

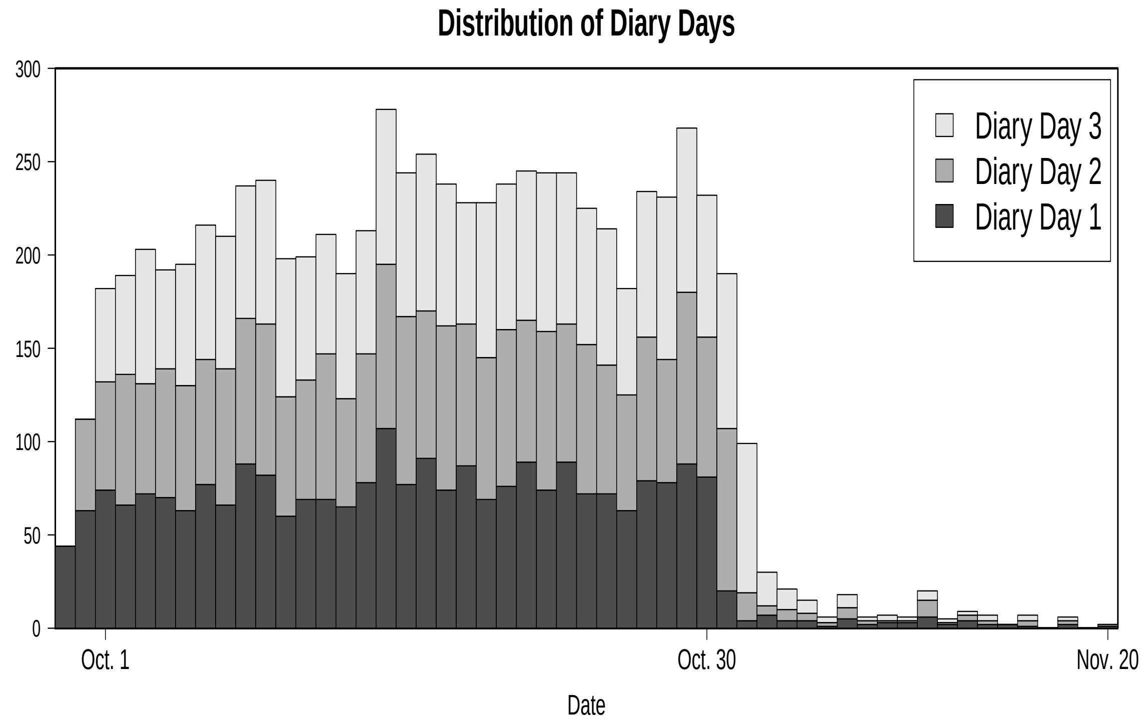

Figure 1: Number of respondents logging on to complete a particular day of diary, by date. Source: Authors' calculations.

and roughly half took the SCPC before the DCPC. Again, while there were some sizeable lags, most people took the two surveys within a few weeks of each other.

Overall, all three days of the diary took an average total of about 35 minutes to complete. The times were measured as the difference between first log on and last log off. Figure 4 shows the distribution of completion times for each day. The distribution of completion times varied across days, with completion times on day 1 and day 3 significantly longer than those on day 2. The additional time to complete day 1 may be partly explained by the fact that there is a learning curve, with respondents becoming more comfortable and efficient at entering data on the latter two days. In addition, day 3 featured questions about bills that were not featured on the other days. Those respondents who took at least 60 minutes to complete the survey might correspond to cases in which respondents stepped away from the computer in the middle of entering data or those in which respondents logged in several times throughout the day.

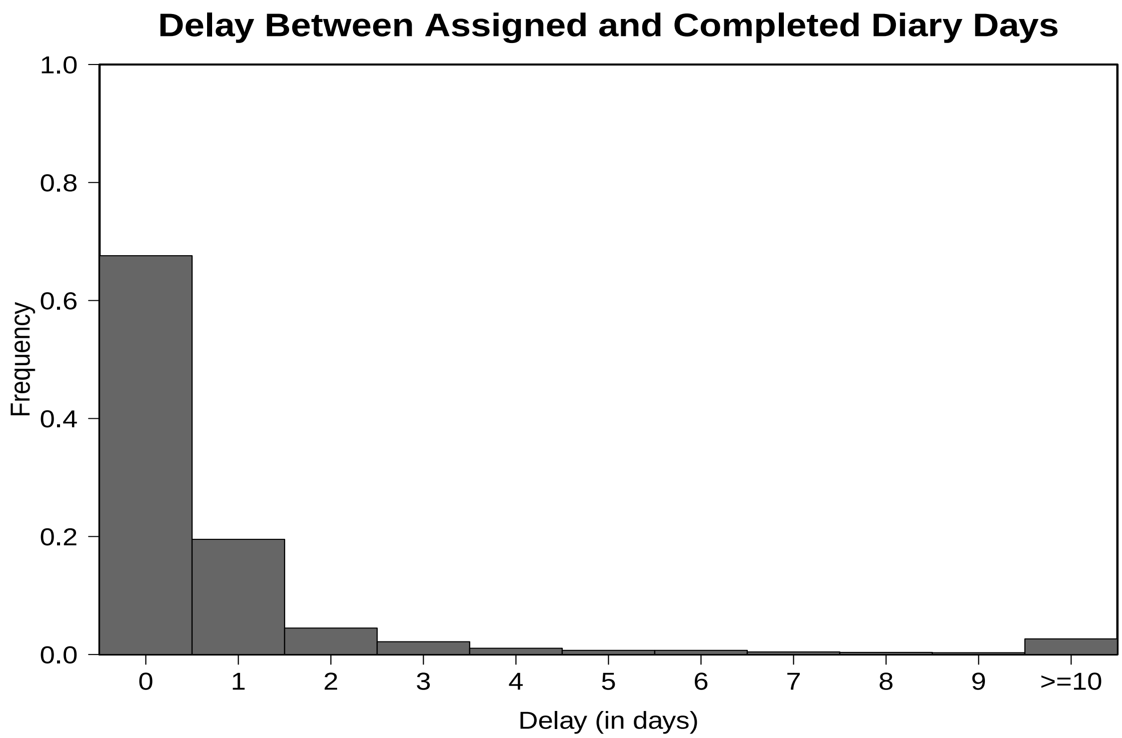

Figure 2: Distribution of the number of days between assigned diary day and day on which respondent logged on to enter data for all completed diary days (up to three per respondent). Source: Authors' calculations.

4.5 Item Nonresponse

For a survey to provide a valid picture of the overall population, it is very important that the item response rates for each question be high. High nonresponse rates not only mean there is less information on which to base estimates but also raise concerns about potential bias in the estimates. If the fact that an observation is missing is independent of the value of the observation, a condition referred to as "missing at random" (Little and Rubin 2002), imputation procedures can be used to generate estimates of sample statistics. However, if there is a confounding variable that relates to both the value of a variable and the likelihood of nonresponse, it is impossible to adjust for the effects on sample statistics. Certain payment details are potentially sensitive topics, and it is possible that the willingness of respondents to provide answers relates to the answers themselves. Naturally, variables with low nonresponse rates are less susceptible to this type of bias.

While we cannot determine whether entire transactions were omitted from online entry, we can look at the fraction of details provided for those transactions that were recorded. Table 4 looks at the nonresponse rates for key attributes of nine different transactions. We see that

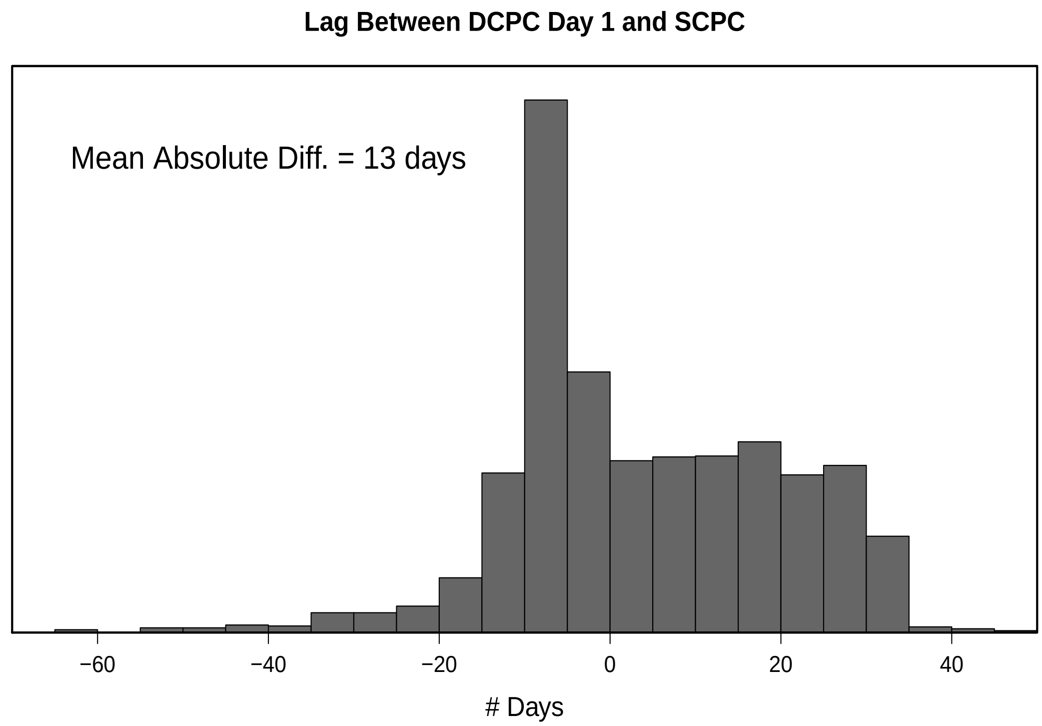

Figure 3: Distribution of the number of days between completion of diary day 1 and completion of the SCPC (Date of diary − Date of SCPC).

Source: Authors' calculations.

overall the item response rates are high, with most respondents providing all information that was requested. The most commonly omitted variable was the time of the transaction, most noticeably the minutes.

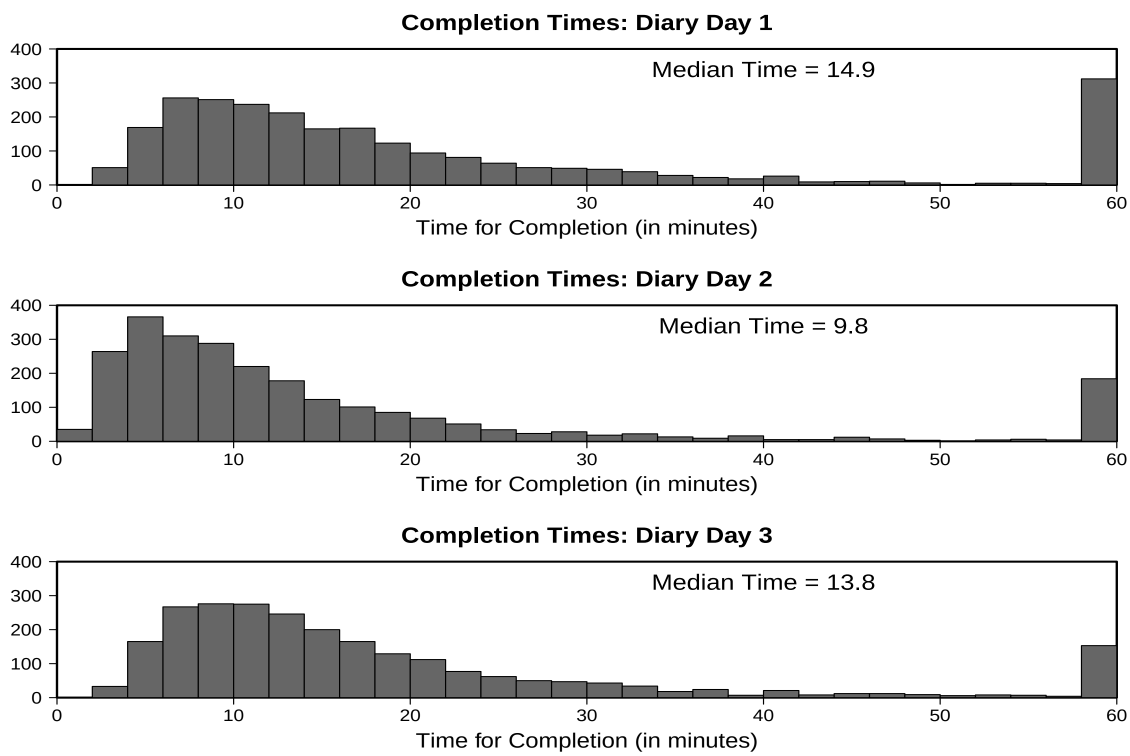

Figure 4: Distribution of the number of minutes to complete each diary day, with truncation to 60 minutes.

Source: Authors' calculations.

Table 4: Item nonresponse rates for nine different transaction types.

Purchases (12,647) observed						
PI	Device	Location	Merchant	Hour	Minute	AM/PM
0.06	2.75	0.05	0.09	0.25	2.99	0.36

Bills (1,356) observed						
PI	Device	Location	Merchant	Hour	Minute	AM/PM
0.07	2.75	0.30	0.30	0.44	4.95	6.80

Automatic Bills (769 observed)	
PI	Merchant
0.13	0.91

Cash Deposits (202 observed)			
Method	Hour	Minute	AM/PM
1.98	4.46	6.44	7.43

Cash Withdrawals (1,237 observed)					
Location	Source	Fee	Hour	Minute	AM/PM
0.81	3.07	4.20	1.94	5.74	5.74

Cash to Coin (556 observed)					
Main Coin	Location	Fee	Hour	Minute	AM/PM
2.34	4.86	7.01	2.34	7.91	7.37

Coin to Cash (64 observed)						
Amount Received	Reimbursed	Location	Fee	Hour	Minute	AM/PM
17.19	3.13	9.38	9.38	0.00	20.31	9.38

Returned Goods (120 observed)			
Method	Hour	Minute	AM/PM
0.00	0.83	0.00	5.00

Prepaid Card Reloads (43 observed)					
PI	Fee	Location	Hour	Minute	AM/PM
0.00	2.33	0.00	2.33	4.65	6.98

Source: Authors' calculations.

4.6 Hurricane Sandy Follow-up Survey

A powerful feature of the ALP is that responses from different surveys can be linked at the individual level. One survey in particular that is relevant to the 2012 DCPC is a follow-up survey conducted by the CPRC relating to the effect of Hurricane Sandy. At the end of October 2012, Hurricane Sandy affected large parts of the East Coast of the United States, mostly between October 26th and November 1st, although with most of the damage done on October 29th. As this time period overlapped with our diary, we took the opportunity to field a follow-up survey to those people who lived in regions that the Federal Emergency Management Agency (FEMA) classified as having been affected by the hurricane. The primary goal of this survey was to determine how the hurricane might have affected payment behavior. The short follow-up survey was fielded in early May 2013.[8]

Table 5 shows the number of individuals who took the diary during Hurricane Sandy, based on the impact level codes provided by FEMA. Areas that were not affected are coded as "no impact," and other areas are classified into one of four impact levels, corresponding to the degree of damage. A map of the impact areas is shown in Figure 5, and more information about these impact levels can be found at the website established by FEMA.[9]

Table 5: Number of respondents with diary days on October 29, by impact of region.

			ZIP Code Impact Level				
Diary Wave			no impact	low	moderate	high	very high
Day 1	Day 2	Day 3		(green)	(yellow)	(red)	(purple)
10/27	10/28	10/29	56	6	2	4	10
10/28	10/29	10/30	53	2	3	4	11
10/29	10/30	11/1	64	6	7	3	5

Source: Authors' calculations.

5 Sampling Weights

5.1 DCPC Sample Demographics

An important goal of the DCPC is to provide estimates of payment statistics for the entire population of U.S. consumers over the age of 18. This can be generally accomplished by using

[8]For more information about Hurricane Sandy data please contact Marcin.Hitczenko@bos.frb.org.
[9]https://www.arcgis.com/home/item.html?id=307dd522499d4a44a33d7296a5da5ea0

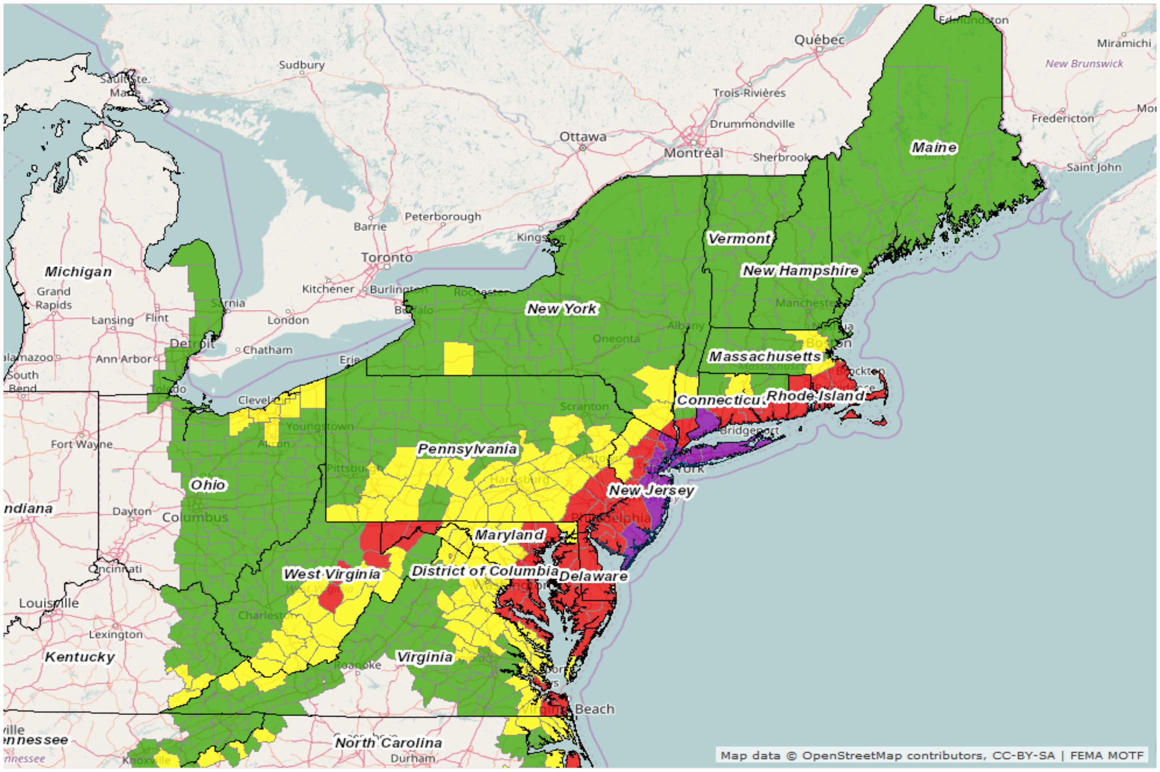

Figure 5: Map of regional impact levels of Hurricane Sandy.
Source: FEMA.

weights to scale sample data appropriately to match the population of interest. As the ALP is a collection of volunteers recruited through various methodologies, the set of DCPC respondents is not a probability sample, making it impossible to use probability-based weighting to generate population-wide inferences. Instead, we rely on poststratification weights that perform the scaling with respect to certain demographic variables. Indeed, recent work by Wang et al. (2009) suggests that nonrepresentative polling can provide relatively accurate estimates with appropriate statistical adjustments.

Although much effort has been expended to make the ALP a better representative sample of the U.S. population, its origin as a sample of older Americans and recruitment rates that vary across demographic groups mean there are non-trivial discrepancies. A notable example is that the proportion of females in the ALP is close to 60 percent, much higher than in the general population. Nevertheless, the ALP has improved with time, and the fact that we are not necessarily sampling the entire ALP means that the SCPC and DCPC subsets can be selected to better match population demographics.

Table 6 shows the unweighted sample proportions for a set of chosen demographic strata for the set of 2012 DCPC respondents and the set of 2012 SCPC respondents used for official estimates (see Angrisani, Foster, and Hitczenko (2014) for details). The right column shows weighted values for the 2012 DCPC, which are intended to match the CPS estimates. We see that the 2012 DCPC unweighted sample is generally more representative of the population than the 2012 SCPC unweighted sample, especially with respect to race and age. This can be explained by the fact that the primary goal of the 2012 DCPC sample selection was to match the population, while the SCPC effort was somewhat hindered by prioritization of maintaining the longitudinal panel. Only gender is more poorly represented in the 2012 DCPC, a result of the stratified recruitment process, which focused on representativeness with respect to variables other than gender (gender is not a characteristic featured in Table 1). The skew in gender is largely due to the fact that the ALP has a much higher share of females, who make up around 60 percent of the panel, than the U.S. population. Overall, the unweighted DCPC sample still tends to be over-educated, more female, and older than the general population. Education level, in particular, shows the worst representation, again, because it is not used to screen in the sample selection process.

5.2 Daily vs. Monthly Weights

To enable better inference of the entire population of U.S. consumers, SCPC respondents are assigned post-stratified survey weights designed to align as much as possible the composition of the SCPC sample with that of a reference population. Specifically, the benchmark distributions against which SCPC surveys are weighted are derived each year from the CPS. This follows common practice in other social science surveys, such as the Consumer Expenditure Survey (CES). The 2012 DCPC generates weights on a monthly basis for all those who participated in the diary, and on a daily basis for all individuals who participated on each day in October.

Developing both sets of weights offers more flexibility in generating population estimates. Daily weights naturally allow calculations for a larger variety of time periods, such as averages corresponding to particular weeks or a particular day of the week. However, for the most commonly used framework of monthly averages, the two sets of weights offer competing estimates. The details of these estimates are described in Section 7, but we discuss their general differences here.

Monthly weights are calculated based on the full sample of around 2,500 people. Daily weights are calculated separately for the set of around 225 respondents who participated on each given day. This difference in the number of people used to post-stratify is important because the lesser number available for daily calculations means that demographic strata

Table 6: Unweighted percentages for various marginal demographics in the 2012 SCPC and DCPC sample, as well as weighted percentages for the 2012 SCPC. The weighted values are based on CPS data.

Demographics		Unweighted 2012 SCPC	Unweighted 2012 DCPC	Weighted 2012 DCPC
Gender	Male	43.6	39.9	48.1
	Female	56.4	60.1	51.9
Age	18−24	3.0	5.6	4.8
	25−34	15.7	21.3	26.1
	35−44	13.1	15.4	15.1
	45−54	22.3	20.3	18.9
	55−64	26.0	22.0	16.7
	65 and older	20.0	15.5	18.5
Race	White	85.5	78.1	75.0
	Black	8.2	11.4	11.7
	Asian	1.8	2.3	2.5
	Other	4.4	8.3	10.8
Ethnicity	Hispanic	7.3	16.0	17.4
Education	No HS diploma	2.7	3.9	6.9
	High School	15.9	15.9	35.0
	Some College	36.8	26.0	28.7
	College	25.2	37.1	16.9
	Post-graduate	19.4	17.0	12.5
Income	$< \$25K$	17.0	22.7	24.1
	$\$25K - \$49K$	24.7	28.0	23.8
	$\$50K - \$74K$	21.6	20.4	20.2
	$\$75K - \$99K$	14.5	11.6	10.5
	$\$100K - \$124K$	9.7	7.7	9.6
	$\$125K - \$199K$	9.0	7.2	9.0
	$\geq \$200K$	3.5	2.3	2.8

Source: Authors' calculations.

must be coarser for daily than for monthly weights. Therefore, the daily weights are less well equipped to adjust for heterogeneity due to demographic differences when individuals are grouped in broader categories. On the other hand, estimates based on monthly weights do not adjust for different numbers of observations on different days of the month, thus giving some daily means more weight than others by virtue of having more respondents on those days. The use of daily weights in estimation accounts for the non-uniform distribution of respondents across days, by assigning equal weight to each daily mean (see Section 7 for details). Based on these principles, daily weights are more appropriate when heterogeneity in behavior comes from temporal changes, such as day effects, while monthly weights are more appropriate when heterogeneity is attributable more to demographic differences.

The result of having monthly and daily weights is that each individual has one monthly weight and up to three daily weights, corresponding to the three days of the diary. Individuals who did not participate on all three assigned days or had some diary days in September or November will have fewer than three daily weights. Figure 6 shows a scatterplot of individuals' monthly weights versus the average of their daily weights. Although there are differences between the two, due to the random fluctuations in daily assignments and a raking procedure based on different demographic strata (see Section 5.3 for details), the general trend shows consistency of the weights within individuals. A more in-depth comparison of daily and monthly weights in the context of population estimates is found in Section 7.

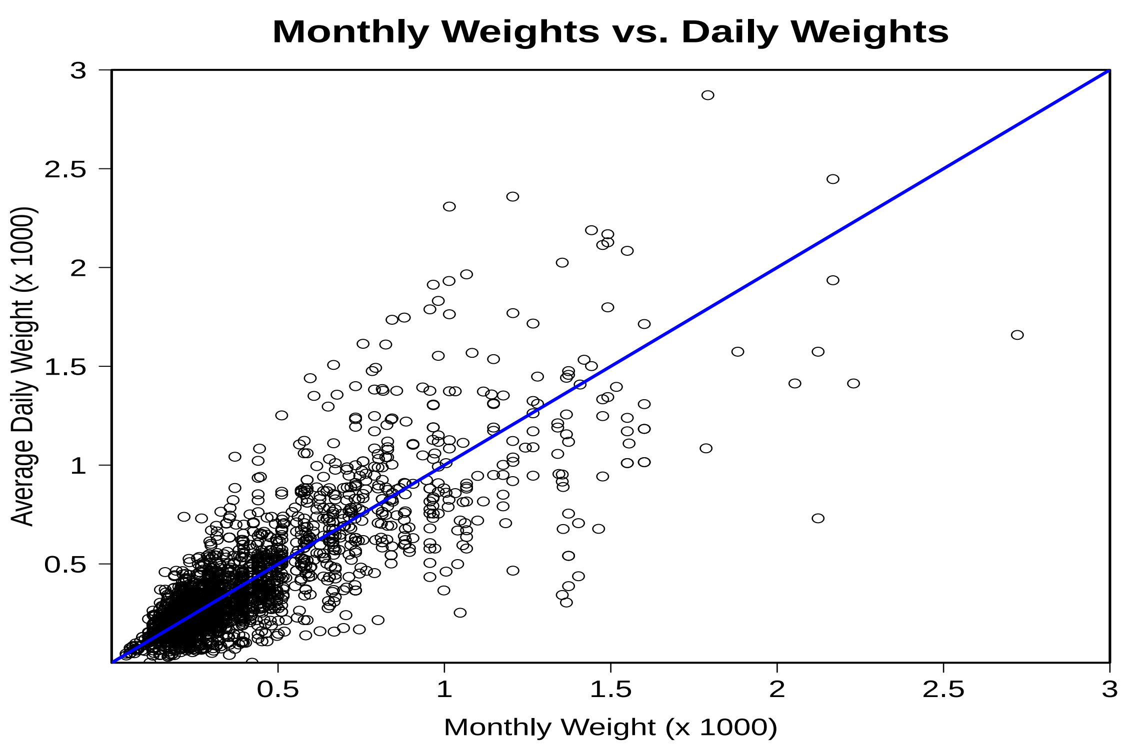

Figure 6: Monthly weights vs. mean of daily weights for each respondent.
Source: Authors' calculations.

5.3 Raking Algorithm

Sampling weights are generated by RAND, using a raking algorithm (Deming and Stephan 1940; Gelman and Lu 2003). This iterative process assigns a weight to each respondent so that the weighted distributions of specific socio-demographic variables in the SCPC sample match their population counterparts (benchmark or target distributions). The weighting procedure consists of two main steps. In the first part, demographic variables from the CPS are chosen and mapped onto those available in the DCPC. Continuous variables such as age and income are recoded as categorical variables by assigning each to one of several disjoint intervals. For example, Table 7 shows five classifications for age and four classifications for income. The number of levels for each variable should be small enough to capture homogeneity within each level, but large enough to prevent strata containing a very small fraction of the sample, which could cause weights to exhibit considerable variability.

Table 7: The set of weighting variables for the monthly weights. "M" stands for male, and "F" stands for female. The highest income brackets for single households were combined to avoid small cell sizes.

Gender × Age				
M, $18-32$	M, $33-43$	M, $44-54$	M, $55-64$	M, 65+
F, $18-32$	F, $33-43$	F, $44-54$	F, $55-64$	F, 65+

Gender × Ethnicity	
M, White	M, Other
F, White	F, Other

Gender × Education		
M, High School or Less	M, Some College	M, Bachelor's Degree or More
F, High School or Less	F, Some College	F, Bachelor's Degree or More

Household Size × Household Income			
Single, $< \$30K$	Single, $\$30K - \$59K$	Single, $\geq 60K$	
Couple, $< \$30K$	Couple, $\$30K - \$59K$	Couple,$\$60K - \$99K$	Couple, $\geq \$100K$
≥ 3 , $< \$30K$	≥ 3 , $\$30K - \$59K$	≥ 3 ,$\$60K - \$99K$	≥ 3 , $\geq \$100K$

Table 7 shows the variables used in weighting as well as the levels within each variable for monthly weights. Comparable information used for daily weights are shown in Table 8. In the second step, the raking algorithm is implemented and sample weights are generated by matching the proportions of predefined demographic groups in the DCPC to those in the CPS. More precisely, the weighting algorithm is performed using the 31 pairs of demographic variables shown in Table 7 for monthly weights and the 20 pairs of demographic variables shown in Table 8 for daily weights.

Table 8: The set of weighting variables for the daily weights. "M" stands for male, and "F" stands for female. The highest income brackets for single households were combined to avoid small cell sizes.

Gender × Age		
M, $18-39$	M, $40-64$	M, 65+
F, $18-39$	F, $40-64$	F, 65+

Gender × Education		
M, High School or Less	M, Some College	M, Bachelor's Degree or More
F, High School or Less	F, Some College	F, Bachelor's Degree or More

Household Size × Household Income		
Single, $< \$40K$	Single, $\geq 40K$	
Couple, $< \$40K$	Couple, $\$40K - \$75K$	Couple, $\geq \$75K$
≥ 3 , $< \$40K$	≥ 3 , $\$40K - \$75K$	≥ 3 , $\geq \$75K$

As the post-stratification weights depend on certain demographic variables, RAND imputes the necessary variables for respondents for whom the information is missing. In the case of many demographic variables, such as age group, gender, or race, missing information can be verified from other surveys taken within the context of the ALP. For household income and household size, both attributes that could easily change within a year, values are imputed by RAND through logistic regression models for the purpose of creating post-stratification weights. The imputations are only used to generate post-stratification weights and are left as missing in the dataset.

The socio-economic variables chosen for the raking procedure result from recent internal research conducted by RAND regarding the sampling properties of weights based on different demographic factors. First, a new imputation algorithm for all possible socio-demographic variables was developed to allow for weights based on a wider range of consumer information. The procedure is sequential, so that variables with the least number of missing values are imputed first and, in turn, used as inputs to impute the variables with the most missing values. Imputations are performed by ordered logistic regression for ordered categorical variables and by multinomial logistic regression for categorical variables. Sample weights produced by different combinations of variables were evaluated on the basis of how well they matched the distributions of demographic variables not used as raking factors (test variables). To assess the robustness and accuracy of different combinations of weighting variables, Monte Carlo samples were drawn and demographic distributions of the test variables were generated based on the weights for that particular sample. Mean deviation from the CPS-defined levels for test variables were estimated by averaging over the samples. The combination of variables in Table 7 consistently matched the target distributions of the CPS for a variety of different sample sizes.

The pairing of gender with other socio-demographic variables allows one to better correct for discrepancies between distributions within each gender, while avoiding the problem of small cell counts. In other words, implementing the raking algorithm on the sets of pairs shown in Table 7 and Table 8 ensures that the distributions of age, ethnicity, and education in the SCPC are matched separately for men and women to their population counterparts in the CPS. Moreover, since bivariate distributions imply marginal distributions for each of the two variables, this approach also guarantees that the distributions of gender, age, ethnicity, and education for the entire SCPC sample are aligned with the corresponding benchmarks in the CPS. The same is true for household size and household income.

The fact that not all strata of interest are represented in the sample makes raking the natural method for assigning weights. However, doing so introduces a few complications related to the statistical framework and analysis of the data. The first relates to the increased difficulty in calculating standard errors of population estimates, which are weighted averages of the sample values. In all tables and publications, the standard errors have been calculated by taking the weights as fixed values, thereby reducing the standard errors. The sampling weights, which are a function of the strata representation in the sample, are random variables, and their variation should be factored into the calculation of standard errors (Gelman and Lu 2003). However, high response rates and targeted sampling (as described in Section 4.2) mean that the variability in the observed sample composition is small, which in turn implies that the variability in the raked weights is small. Therefore, conditional on a chosen weighting scheme, the variance of our estimators is largely attributable to the variation in the observed responses themselves and not to the variance in the sample composition.

The second area of concern regards the effects of the sampling scheme on the weights and on the estimates they produce. In order for the raking algorithm to be appropriate in the sense that the expected weights for each stratum equal those of the population, the sampling procedure must be such that, in expectation, each stratum is proportionally represented in the sample. To be precise, the expected proportion of the sample belonging to a specific stratum is directly proportional to the relative proportion of that stratum within the population. A sampling procedure that does not have this property is likely to consistently produce weights for certain strata that do not reflect the true representation in the entire population. If strata properties correlate with payment behavior, this could lead to biased population-wide estimates. In the case of a sampling procedure in which some strata tend to be over-represented and others under-represented, the raking algorithm, which strives to match marginal proportions rather than those of the cross-sections of all the variables, may generate sample weights that fail to align the sample composition with the reference population. Although the sample from the ALP does not perfectly reflect the U.S. population, the differences between the panel and the broader population are relatively small for the demographics used in weighting. In addition, for many DCPC variables there is little evidence of strong correlations with these variables used in weighting, so any bias is likely to be small.

6 Data Preprocessing

Prior to further statistical analysis, it is important to carefully examine and ensure data quality. In the case of the DCPC, this primarily involves properly identifying all unique and valid transactions and confirming details of those transactions. While this mostly involves reclassifying the attributes of reported transactions, the CPRC also assesses the validity of certain numerical values reported in the data. Below, we detail the individual steps taken to convert the raw data entries to those found in the final dataset.

6.1 Identifying Valid Transactions

The fundamental unit of measure in the 2012 DCPC is the transaction; the official dataset is organized so that each row corresponds to a transaction. At the broadest level, there are two types of transactions: those that indicate participation of a respondent on a particular day and those that represent economic transactions. The latter are defined by some form of exchange or transfer of money, and thus must have a dollar value. So, our first step is to limit the data to only those transactions with non-missing and positive dollar values associated with them. Specifically, we exclude from the final data the small number of cases in which respondents recorded a transaction worth $0 dollars or with the value missing.

The next step of the editing process involves making sure that entered transactions are unique. Duplicate transactions are possible for a variety of reasons. Respondents could enter the same transaction by accident. Perhaps more likely, however, is that the respondent entered the same transaction in separate modules, which are used to prompt recall of payments. The most obvious example involves a separate module on the third diary day that asks about bills over the diary period. Identifying duplicates is not necessarily easy, as it is possible to have two payments with identical attributes within a 15-minute period (the shortest time unit offered). Therefore, the process of identifying duplicates requires making subjective assumptions.

In identifying duplicates, the CPRC currently focuses only on entries entered as one of three transaction types: general payments, bills, and automatic bills. The adopted procedure takes a rather conservative approach, requiring transactions to match on all attributes available for the two relevant transaction types: respondent, diary day, dollar value of payment, payment instrument used, and merchant category selected. Especially for non-standard dollar amounts, the combination of information for respondent, date, and dollar amount represent the most important identifiers of a payment and are sufficiently unique to identify duplicates.

The cleaning algorithm is structured so as to identify duplicates and to categorize the transaction in the narrowest payment group possible. Therefore, we begin by removing duplicates from the purchases module and the bill payments module, deleting any duplicates from the purchases list and inserting them into the bill payments list. After doing this for bills and purchases, we repeat this for bills and automatic bills, and then once more for purchases and automatic bills. In 2012, there were 46 duplicates across payment modules.

The current procedure ignores potential duplicates within a transaction type or between other pairs of transaction types that could feasibly yield duplicates, such as cash deposits and payments. Not many such duplicates exist. For example, among non-bill purchases, around 99 percent of transactions are unique in that the combination of individual, day, amount, payment instrument, and merchant are unmatched by any other transaction.

The final, processed dataset available for public use is based on individuals who participated and completed the diary on all three days. Making this reduction as well as removing duplicates reduces the dataset from 12,891 purchases, 1,402 bills, and 789 automatic bills to 12,647 purchases, 1,356 bills, and 769 automatic bills.

6.2 Editing Transaction Attributes

We also make certain edits to the fields describing transaction attributes, although this is relatively rare. Most often, we use information provided in the text box for the "other (specify)" option in multiple choice questions to recode the response to an existing response option. This most commonly occurs in questions determining the merchant category. In all such cases, the entered text remains as part of the data available to researchers. The original classifications provided by the respondents can be found in the unprocessed dataset.

6.3 Editing Cash Holdings

Every day, respondents are asked to enter the number of bills for each denomination of bill owned (1s,5s, etc.) and confirm the total value of cash on hand, which is automatically calculated for respondents but can also be edited by the respondent. We do not do any cleaning to these data entries as long as they are numerical. In two cases, when non-numeric responses were entered, we reset all related variables to zero.

6.4 Editing Dollar Values

The greatest challenge in data preprocessing for the DCPC comes in the form of dollar values of transactions. Measurement errors in such a context, defined as any incongruity between the data entry and the true value, can be attributed to a variety of sources ranging from recall error to rounding errors to data entry errors. In practice, there is an asymmetry to identifying measurement error, as we can really only hope to identify values that are too high (Chambers and Ren 2004).

In determining the editing philosophy, it is important to distinguish between influential and likely invalid data points. An influential point is one whose inclusion or exclusion in any inferential analysis causes a significant difference in estimates (Bollen and Jackman 1990; Cook and Weisberg 1982), and thus the influence of a point depends on the statistical procedure being performed. These can be identified through a variety of statistical measures, such as Cook's distance (Cook 1977; Cook and Weisberg 1982). An invalid data entry is, technically, any entry that does not represent the truth. An invalid data point need not be influential and an influential point is not necessarily invalid. Our goal is to identify transaction values that are likely impossible.

Identification of data that are technically possible, but very unlikely, is difficult, as it involves comparing a data entry within the context of heterogeneity of behavior within the population, especially for economic variables such as dollar values, which have fat right tails. In other words, it is possible that data entries that by some numerical evaluations are statistical outliers are actually accurate and valid. This issue is not unique to the SCPC. Many consumer surveys, such as the Survey of Consumer Finances (SCF) and the Consumer Expenditure Survey (CES) must also tackle the cleaning of such fat-tailed variables. While the details of the preprocessing of outliers are not provided in either survey, the general strategy is similar to that adopted in the SCPC (Bricker et al. 2012; Bureau of Labor Statistics 2013): use all available intuition and information about value distributions to assign a degree of plausibility. This often involves making subjective decisions.

In the public use dataset, the value provided by the respondent in the diary corresponds to the variable `amnt_orig`. These are unaltered (except in a few cases where respondents were re-surveyed, described below). Our editing approach identifies likely implausible transaction values and sets the variable `amnt` for those transactions to missing. For all other transactions, it is always the case that `amnt` is equal to `amnt_orig`.

It is possible to use model-based methodologies to identify transaction values that are implausible, and the perceived objectivity of such an approach is certainly desirable. However, this is difficult to do in practice. Fitting parametric models is somewhat subjective to begin with, as we do not know how right tails of dollar values should behave. Choosing different distributions, such as a log-Normal or Gamma distribution, will lead to different conclusions, and the distribution parameters are largely determined by the data in the center of the distribution anyway. How one should pool and analyze data is also complicated, as there is a lot of potential information provided about the transaction that should inform the decision. The merchant, payment instrument used, and individual characteristics of the payer, such as household income, naturally relate to the degree of plausibility. Hospital bills are quite plausibly in the thousands of dollars, but this is rare for convenience stores; large-value payments to financial advisers are less likely for those with low income than for those with high income. However, in separating data into finer categories with distinct distributions, we must rely on less data to make inferences about behavior in the right tail and thus lose power to identify extreme outliers.

Instead, the 2012 DCPC dollar-value cleaning relies on follow-up interviews with certain respondents about particular transactions as well as general intuition about validity based on all provided information. We stress that the original data are still found in the final dataset, so researchers are free to edit original transaction values as they like. Table 9 shows a set of 20 transactions for which respondents were asked to confirm values in April, 2013. These were chosen as the transactions that seemed most suspicious based on available information. Of these, nine individuals responded to the inquiry, with one individual confirming the original value and the other eight making corrections. For those nine observations, the confirmed or updated values replace the original amount value, `amnt_orig`, and we accept them as true. For those who did not respond, we considered them along with all other observations in determining whether they are trustworthy or not. For the 11 transactions for which we did not hear back to our inquiry, we decided that 9 are likely incorrect and set `amnt` for those transactions to missing; we kept the other two transactions. These are shown in Figure 9. In addition, another 20 transactions not inquired about were identified as likely errors and set the value of `amnt` to missing. These are shown in Table 10.

Table 9: Large-value transactions for which respondents were asked to confirm values and the edits made. "Inquiry Response" identifies the value of the transaction provided by the respondent upon the follow-up inquiry; those who did not respond are noted by "NR." Note: * respondent dropped from final sample.

Date	Type	Merchant/ Source	PI	Original Amount ($)	Inquiry Response	`amnt_orig`	`amnt`
10/4	Purchase	Other stores	Check	291,060	2,910.60	2,910.60	2,910.60
10/24	Purchase	Repair/maintenance of electronics	BANP	173,071	173.71	173.71	173.71
10/6	Purchase	Veterinarians	Debit	18,313	NR	18,313	NA
10/2	Purchase	Grocery, pharmacy, convenience stores	Debit	17,241	NR	17,241	NA
10/3	Purchase	Mail, delivery, storage	Cash	14,400	144.00	144.00	144.00
10/2	Purchase	Grocery, pharmacy, convenience stores	Credit	9,632	NR	9,632	NA
10/20	Purchase	Financial services	OBBP	8,533.24	NR	8,533.24	8,533.24
10/25	Purchase	Grocery, pharmacy, convenience stores	Prepaid	6,595	NR	6,595	NA
11/1	Purchase	Grocery, pharmacy, convenience stores	Cash	6,045	60.45	60.45	60.45
11/1	Purchase	Electric, natural gas, water, and sewage	Debit	4,500	450.00	450.00	450.00
10/28	Purchase	Department and discount stores	Other	4,500	NR	4,500	NA
10/25	Purchase	Restaurants, bars	Cash	3,800	38.00	38.00	38.00
10/20	Purchase	Employment services, travel agents	Credit	3,400	NR	3,400	3,400
10/18	Purchase	Grocery, pharmacy, convenience stores	Prepaid	3,325	325.00	325.00	325.00
10/5	Purchase	Other stores	Credit	3,220.94	3.22	3.22	3.22
10/1	Bill	Auto and vehicle	Check	24,535	245.35	245.35	245.35
10/5	Bill	Other people	OBBP	9,895.20	NR	9,895.20	NA
10/20	Deposit	Bank teller		14,000	14,000	14,000	14,000
10/16	Deposit	Bank teller		12,600	NR	12,600	NA
10/5	Deposit	Other		5,250	NR	5,250	NA *

Table 10: Edited large-value transactions for which respondents were not asked to confirm values.

Date	Transaction Type	Merchant/ Source	Payment Instrument	Original Amount ($)
10/15	Purchase	Furniture and home goods	Debit	9,923
10/7	Purchase	Grocery, pharmacy, convenience stores	Credit	5,000
10/13	Purchase	Clothing	Credit	4,386
10/27	Purchase	Gas station	Credit	3,826
10/8	Purchase	Financial services	Other	3,651.95
10/3	Purchase	Gas station	Debit	3,521
10/3	Purchase	Personal care	Cash	3,000
10/11	Purchase	Restaurant, bar	Cash	2,947
10/25	Purchase	Grocery, pharmacy, convenience stores	Cash	2,657
10/9	Purchase	Grocery, pharmacy, convenience stores	Cash	2,059
10/15	Purchase	Grocery, pharmacy, convenience stores	Debit	2,018
10/7	Purchase	Mortgage	BANP	1,672
10/18	Purchase	Grocery, pharmacy, convenience stores	Debit	1,620
10/24	Purchase	Grocery, pharmacy, convenience stores	Cash	1,509
11/1	Bill	Online shopping	BANP	3,899.49
11/2	Withdrawal	Savings acct.	NA	4,000
11/1	Withdrawal	Other person	NA	30,000
10/13	Withdrawal	Other source	NA	7,300
10/25	Withdrawal	Checking account	NA	3,790
10/2	Deposit	Other location	NA	3,687

7 Population Parameter Estimation

An important goal of the data collection in the DCPC is to produce estimates of consumer payment behavior for the entire population of U.S. consumers. Doing so can be done using a variety of methodologies that factor in different aspects of the collected data, such as effects associated with days of the week or monthly cycles. In the sections below, we present a simple methodology for generating point estimates and standard errors for two general types of variables: those that define an average consumer and those that define an average transaction. The adopted framework has the benefit of being applicable to wide range of variables and does not require distributional assumptions. We provide formulas based on monthly weights and daily weights and discuss how the various estimates relate to one another.

We begin with some notation. Every transaction is indexed by the triplet (i, d, j), where $i = 1, \ldots, N$ identifies the respondent, d defines the date of the transaction, and j enumerates the transactions made by individual i on day d. Each transaction has various attributes associated with it for which we use the general notation Y_{idj}. These attributes can take the form of a numerical value such as the dollar value of the transaction, a binary variable (for example, 1 if cash was used and 0 otherwise) or a combination of the two, such as the dollar value of the transaction if cash was used and 0 otherwise. We refer to monthly weights as w_i and to daily weights as v_{id}.

7.1 Per-Consumer Estimates

Per-consumer estimates define an average over all consumers. Given that the 2012 DCPC data are collected predominantly throughout the month of October, a natural estimate might correspond to a population average for this period of time. Take $\theta(Y)$ to be some population value corresponding to the attribute Y, such as the average amount spent in October or the average number of cash payments in October.

Our methodology yields estimates that are linear combinations of observed daily totals, given by $Y_{id} = \sum_{(i,d,j)} Y_{idj}$. The daily total is quite natural given that the fundamental sampling unit is the consumer day. It also helps account for the fact that not all individuals participated for three days in October, either because some assigned days were in September or November or because of nonresponse. To factor this in, we let D_i represent the number of individual i's diary days that are in October. Thus, $D_i = 1, 2$, or 3.

To generate monthly estimates for October, we restrict the analysis to data from that month only. For certain economic concepts it may be reasonable to assume that data collected at the end of September or the beginning of November have the same distribution as those from October. In such a case, it is beneficial to incorporate the additional data in the October estimates. However, for variables that have different distributions at the beginning or end of a month than in the middle of the month, doing so would bias results. For example, if rent payments tend to be scheduled for the last day of the month, including data from the end of September as well as October in our estimates would effectively count such payments twice. Therefore, all per-consumer estimates have the general form

$$\hat{\theta}(Y) \;=\; \sum_{i=1}^{N} \sum_{d \in Oct} k_{id} Y_{id}, \qquad (1)$$

where k_{id} represent generic scalar weights. The weights, k_{id}, can be constructed using monthly or daily weights. Naturally, daily weights can be used to generate population estimates for time periods other than October, such as for a particular day or for a portion of the month. However, we do not detail how this would be done, as it is a trivial extension of the monthly estimate discussed below.

7.1.1 Monthly Weights

The general form of an estimate for an average for the month of October based on data from N individuals might be

$$\hat{\theta}^{month}(Y) \;=\; \frac{\sum_{i=1}^{N} w_i \hat{\theta}_i(Y)}{\sum_{i=1}^{N} w_i},$$

where $\hat{\theta}_i(Y)$ is an estimate of the monthly mean for individual i. This estimate will be unbiased if $\mathrm{E}\left[\hat{\theta}_i(Y)\right] = \theta_i(Y)$ and weights are appropriately assigned. Perhaps the simplest estimate of $\theta_i(Y)$ is just the corresponding estimate for the days of participation in October scaled up to a monthly basis, which, in the case of a month with 31 days, takes the form:

$$\hat{\theta}_i(Y) \;=\; \frac{31}{D_i} \sum_{d \in Oct} Y_{id}.$$

If the set of days of participation assigned to individual i span the month uniformly (as if they were chosen at random), then $\mathrm{E}\left[\hat{\theta}_i(Y)\right] = \theta_i(Y)$. Then, the population estimate takes the form

$$\hat{\theta}^{month}(Y) \;=\; \frac{\sum_{i=1}^{N} w_i \left[\frac{31}{D_i} \sum_{d \in Oct} Y_{id}\right]}{\sum_{i=1}^{N} w_i}, \qquad (2)$$

and $k_{id}^{month} = \frac{31 w_i}{D_i \sum_{i=1}^{N} w_i}$. We discuss calculating standard errors for (2) in Section 7.3.

7.1.2 Daily Weights

The principle behind estimates of θ based on daily weights is identical to that for the monthly estimates, except that the methodology is applied to each day and the monthly estimate is a summation over all days in October:

$$\hat{\theta}^{day}(Y) \quad = \quad \sum_{d \in Oct} \frac{\sum_{i=1}^{N} v_{id} Y_{id}}{\sum_{i=1}^{N} v_{id}}. \tag{3}$$

The inner sum in (3) is a population estimate for day d. The sequence of sums in (3) can be rearranged to sum over individuals first and days second:

$$\hat{\theta}^{day}(Y) \quad = \quad \sum_{i=1}^{N} \sum_{d \in Oct} \frac{v_{id}}{\sum_{i'=1}^{N} v_{i'd}} Y_{id}, \tag{4}$$

so that $k_{id}^{day} = \frac{v_{id}}{\sum_{i'=1}^{N} v_{i'd}}$.

It is worth comparing k_{id}^{month} and k_{id}^{day} to see that the relative contributions of Y_{id} to a monthly estimate are similar, especially in expectation. Within k_{id}^{day}, the relative ratio of v_{id} to $\sum_{i=1}^{N} v_{id}$ represents the relative contribution of individual i to the estimate for day d. On average, therefore, $\mathrm{E}\left[k_{id}^{day}\right] \approx \frac{1}{N_d}$, where N_d represents the number of respondents on day d. By design of the survey, three of the 33 diary waves are participating on each day, meaning about one-eleventh of individuals, on average, are participating on day d. Thus, $N_d \approx \frac{N}{11}$ and $\mathrm{E}\left[k_{id}^{day}\right] \approx \frac{11}{N}$. Using similar logic for k_{id}^{month}, we have that $\mathrm{E}\left[k_{id}^{month}\right] \approx \frac{31}{D_i N}$. For virtually all respondents, $D_i = 3$, so that $\mathrm{E}\left[k_{id}^{month}\right] \approx \frac{11}{N}$. Therefore, through this approximation, we can see that, on average, the contribution of each relevant data point in the monthly and daily weights is the same. This is a desirable quality in the estimates. Discrepancies are largely due to the lack of representative sampling across the days of the month and differences in raking algorithms used for daily and monthly weights.

The calculation of standard errors using daily weights is discussed in Section 7.3.

7.2 Per-Transaction Estimates

The use of the 2012 DCPC data to make estimates on a per-transaction basis is perhaps less obvious, because it is less intuitive that using weights associated with consumers can yield values that correspond to an average transaction. An example of such a parameter might be the average value spent at a gas station or the proportion of payments made for medical expenses that are paid with credit cards.

In fact, the per-transaction estimates are simply ratios of per-consumer estimates, as can be seen if one represents:

$$\text{Average value per transaction} = \frac{\text{Average value of transactions per consumer}}{\text{Average number of transactions per consumer}}. \quad (5)$$

Therefore, a per-transaction estimate represents a share. The CPRC has found that such shares are robust and economically meaningful economic quantities. We note that these shares are based on the macroeconomic definitions, meaning that we consider the share of the averages rather than the average of the individual shares. The microeconomic alternative, which calculates a share for each individual and takes the average of those, is not recommended. For one thing, especially with limited data from a three-day diary, this will likely yield undefined estimates $\left(\frac{0}{0}\right)$ for individuals, which are difficult to deal with. In addition, within many frameworks, the macroeconomic estimates correspond to the maximum likelihood estimates of the share. For example, if for $i = 1, \ldots, N$ individuals, the number of observed transaction is T_i, of which $S_i \sim \text{Binomial}(T_i, p)$ are of interest, then the maximum likelihood estimate of the share, p, is $\frac{\sum_{i=1}^N S_i}{\sum_{i=1}^N T_i}$ rather than $N^{-1} \sum_i \frac{S_i}{T_i}$.

Both the numerator and denominator in (5) are per-consumer estimates based on two different attributes, which we generically label $Y^{(1)}$ and $Y^{(2)}$. The denominator can be expressed as $\theta(Y^{(1)})$, using the general notation of (1), where $Y^{(1)}$ is a binary attribute that identifies the transactions of interest. If one is interested in all transactions, then $Y_{idj}^{(1)} = 1$ for all (i, d, j), while in the second example given in the previous paragraph, $Y_{idj} = 1\,[\text{transaction } (i, d, j) \text{ is for medical expenses}]$. The numerator is $\theta(Y^{(2)} \times Y^{(1)})$, where $Y^{(2)}$ can be any type of attribute. We use the notation $Y_{idj}^{(2,1)} = Y_{idj}^{(2)} \times Y_{idj}^{(1)}$. In the example relating to medical expenses, $Y_{idj}^{(2)}$ is a binary variable that is 1 if the the transaction was made with a credit card and 0 otherwise. The per-transaction estimate is thus estimated by

$$\hat{\mu}(Y^{(1)}, Y^{(2)}) = \frac{\hat{\theta}(Y^{(2,1)})}{\hat{\theta}(Y^{(1)})}. \quad (6)$$

The numerator and denominator in (6) can be calculated using monthly or weekly weights according to (2) or (4), respectively.

7.2.1 Monthly Weights

Each per-consumer estimate can be generated using monthly weights:

$$\hat{\mu}^{month}(Y^{(1)}, Y^{(2)}) = \frac{\sum_{i=1}^{N} \frac{w_i}{D_i} \sum_{d \in Oct} Y_{id}^{(2,1)}}{\sum_{i=1}^{N} \frac{w_i}{D_i} \sum_{d \in Oct} Y_{id}^{(1)}}. \tag{7}$$

Standard errors forms are described in Section 7.3.

7.2.2 Daily Weights

Estimates on a per-transaction basis can also be constructed using daily weights:

$$\hat{\mu}^{day}(Y^{(1)}, Y^{(2)}) = \frac{\sum_{d \in Oct} \frac{\sum_{i=1}^{N} v_{id} Y_{id}^{(2,1)}}{\sum_{i=1}^{N} v_{id}}}{\sum_{d \in Oct} \frac{\sum_{i=1}^{N} v_{id} Y_{id}^{(1)}}{\sum_{i=1}^{N} v_{id}}}. \tag{8}$$

Standard errors forms are described in Section 7.3.

7.3 Standard Errors

Standard errors are simply calculated as $\mathrm{SE}\left(\hat{\theta}(Y)\right) = \sqrt{\mathrm{Var}\left(\hat{\theta}(Y)\right)}$ or $\mathrm{SE}\left(\hat{\mu}(Y^{(1)}, Y^{(2,1)}\right) = \sqrt{\mathrm{Var}\left(\hat{\mu}(Y^{(1)}, Y^{(2,1)}\right)}$. We begin by discussing standard errors for per-consumer estimates and then do so for per-transaction estimates.

7.3.1 Per-Consumer Estimates

Assuming independence across individuals, the general form of variances of per-consumer estimates of type (1) is

$$\mathrm{Var}(\hat{\theta}(Y)) = \sum_{i=1}^{N} \mathrm{Var}\left[\sum_{d \in Oct} k_{id} Y_{id}\right]. \tag{9}$$

Treating the weights, k_{id} as fixed, the heart of the calculation involves determining $\mathrm{Cov}(Y_{id}, Y_{id'})$ for all combinations of d, d'. Here, again, various models and assumptions can potentially be used. However, we rely on a relatively simple one. For any individual, let $Y_i = [Y_{id_1} Y_{id_2} Y_{id_3}]$ represent the vector of the measurements for individual i on three consecutive days. Then, the quantity of interest can be determined by modeling the $\mathrm{Cov}(Y_{id}, Y_{id'})$ according to:

$$\mathrm{Var}(\mathrm{Y}_i) = \begin{bmatrix} \sigma^2 & \rho_1 & \rho_2 \\ \rho_1 & \sigma^2 & \rho_1 \\ \rho_2 & \rho_1 & \sigma^2 \end{bmatrix}.$$

Therefore, σ^2 represents the variance of the daily observations, ρ_1 represents the covariance of observations from consecutive days, and ρ_2 represents the covariance of observations from two days apart. This framework allows for the fact that the daily attributes may be correlated. For example, we allow that an individual who makes a large number of purchases on one day may be more likely make fewer purchases on the next day. The key simplification is that we assume stationarity, in that joint distributions depend only on the number of days that separate the observed data and not on when those days occurred.

The weighted sum over observed days in October can be expressed in matrix notation as

$$\sum_{d \in Oct} k_{id} Y_{id} = P_i K_i Y_i^T,$$

where

$$K_i = \begin{bmatrix} k_{id_1} & 0 & 0 \\ 0 & k_{id_2} & 0 \\ 0 & 0 & k_{id_3} \end{bmatrix},$$

and P_i represents the three-dimensional vector with 1 indicating that the data from that day were observed and 0 indicating otherwise. Thus, if individual i participated on all three days, $P_i = [1\ 1\ 1]$, and if individual i participated only on the first assigned day, $P_i = [1\ 0\ 0]$. Then, the variance of the weighted sum is expressed as:

$$\mathrm{Var}\left[\sum_{d \in Oct} k_d Y_{id}\right] = P_i K_i \Sigma\, K_i^T P_i^T. \tag{10}$$

We estimate the variance in (10) by replacing Σ with its estimate, $\hat{\Sigma}$, based on the following calculations:

$$\hat{\sigma}^2 = \frac{1}{\sum_{i=1}^{N} D_i - 1} \sum_{i=1}^{N} \sum_{d \in Oct} \left(Y_{id} - \bar{Y}\right)^2,$$

$$\hat{\rho}_1 = \frac{1}{\sum_{i=1}^{N}(D_i - 1)} \sum_{i=1}^{N} \sum_{d_1,d_2|d_1-d_2=1} (Y_{id_1} - \bar{Y})(Y_{id_2} - \bar{Y}),$$

and

$$\hat{\rho}_2 = \frac{1}{\sum_{i=1}^{N} 1[D_i = 3]} \sum_{i=1}^{N} \sum_{d_1,d_2|d_1-d_2=2} (Y_{id_1} - \bar{Y})(Y_{id_2} - \bar{Y}),$$

where

$$\bar{Y} = \frac{1}{\sum_{i=1}^{N} D_i} \sum_{i=1}^{N} \sum_{d \in Oct} Y_{id}.$$

$\bar{Y}$ is simply the unweighted, sample average, and all other calculations are sample covariances and variances for appropriate pairs of observations. Note that for ρ_1, individuals who participated on all three days have two pairs of observations that contribute to the estimate: (Y_{id_1}, Y_{id_2}) and (Y_{id_2}, Y_{id_3}). Thus, the standard error can be expressed in its most general form as

$$\mathrm{SE}(\hat{\theta}(Y)) = \sum_{i=1}^{N} P_i K_i \hat{\Sigma} K_i^T P_i^T. \tag{11}$$

It is worth considering the specific forms of the standard errors, especially for the case of monthly estimates where there is simplification due to the fact that $k_{id} = k_i$ does not depend on the day d. Table 11 shows the variances for individual sums when using monthly weights. Notice that when cross-day covariances are $\rho_1 = \rho_2 = 0$, then the variance is proportional to D_i, the number of days participated. For certain variables, this may be an adequate simplification.

Table 11: Variance of weighted daily sums based on monthly weights for different participation patterns in October.

Participation Dates	D_i	P_i	Variance
(9/29, 9/30, 10/1)	1	[0 0 1]	$k_i^2 \sigma^2$
(9/30, 10/1, 10/2)	2	[0 1 1]	$k_i^2 (2\sigma^2 - 2\rho_1)$
All 3 days in October	3	[1 1 1]	$k_i^2 (3\sigma^2 - 4\rho_1 - 2\rho_2)$
(10/30, 10/31, 11/1)	2	[1 1 0]	$k_i^2 (2\sigma^2 - 2\rho_1)$
(10/31, 11/1, 11/2)	1	[1 0 0]	$k_i^2 \sigma^2$

7.3.2 Per-Transaction Estimates

In the case of per-transaction estimates, the variances are more complicated, because the form of the estimator is a ratio of two random variables. Using properties of variances, we have

$$\mathrm{Var}(\hat{\mu}(Y^{(1)}, Y^{(2)})) \quad = \quad \mathrm{E}\left(\mathrm{Var}\left[\frac{\hat{\theta}(Y^{(2,1)})}{\hat{\theta}(Y^{(1)})}\middle|\{Y_{id}^{(1)}\}\right]\right) + \mathrm{Var}\left(\mathrm{E}\left[\frac{\hat{\theta}(Y^{(2,1)})}{\hat{\theta}(Y^{(1)})}\middle|\{Y_{id}^{(1)}\}\right]\right) . \tag{12}$$

Because the expected value of $\frac{\hat{\theta}(Y^{(2,1)})}{\hat{\theta}(Y^{(1)})}$ does not depend on the number of observations observed and, thus as some fixed value, has a variance of zero, the second term on the right-hand side of (12) is zero. Conceptually, this corresponds to the idea that the average value of a transaction is the same no matter how many transactions are observed. The variance estimate is thus

$$\mathrm{Var}(\hat{\mu}(Y^{(1)}, Y^{(2)})) \quad = \quad \mathrm{E}\left[\frac{1}{(\hat{\theta}(Y^{(1)}))^2}\right] \mathrm{Var}\left[\hat{\theta}(Y^{(2,1)})\right] .$$

While it is possible to estimate the expected value of the inverse of the square of $\hat{\theta}(Y^{(1)})$, especially using a Normal approximation and the delta method (Casella and Berger 2002), it is common to approximate this term simply by using the inverse square of the observed value itself, $\hat{\theta}(Y^{(1)})$. In practice, when the number of transactions observed is high, this simplification will have a trivial effect on the estimate. Thus, the variance estimate becomes

$$\hat{\mathrm{Var}}\left[\hat{\mu}\left(Y^{(1)}, Y^{(2)}\right)\right] \quad \approx \quad \left[\hat{\theta}(Y^{(1)})\right]^{-2} \mathrm{Var}\left[\hat{\theta}(Y^{(2,1)})\right] .$$

This corresponds to a variance of the average value of a transaction conditional on the number of transactions observed. The variance on the right-hand side above can be calculated using the forms in (11).

References

Ahmed, Naeem, Matthew Brzozowski, and Thomas Crossley. 2006. "Measure Errors in Recall Food Consumption Data." IFS Working Paper W06/21. Institute for Fiscal Studies.

Angrisani, Marco, Kevin Foster, and Marcin Hitczenko. 2013. "The 2010 Survey of Consumer Payment Choice: Technical Appendix." Federal Reserve Bank of Boston Research Data Report 13-3.

Angrisani, Marco, Kevin Foster, and Marcin Hitczenko. 2014. "The 2011 – 2012 Survey of Consumer Payment Choice: Technical Appendix." Federal Reserve Bank of Boston Research Data Report 14-2.

Angrisani, Marco, Kevin Foster, and Marcin Hitczenko. 2015. "The 2013 Survey of Consumer Payment Choice: Technical Appendix." Federal Reserve Bank of Boston Research Data Report 15-5.

Baltagi, Badi H. 2008. *Econometric Analysis of Panel Data*. Hoboken, New Jersey: John Wiley and Sons.

Bollen, Kenneth A., and Robert W. Jackman. 1990. "Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases." In *Modern Methods of Data Analysis*, eds. John Fox and J. Scott Long, 257–291. Newbury Park, CA: Sage.

Bricker, Jesse, Arthur B. Kennickell, Kevin B. Moore, and John Sabelhaus. 2012. "Changes in U.S. Family Finances from 2007 to 2010: Evidence from the Survey of Consumer Finances." *Federal Reserve Bulletin* 98(2).

Bureau of Labor Statistics. 2013. "Consumer Expenditures and Income." In *BLS Handbook of Methods*. BLS Publishing.

Casella, George, and Roger L. Berger. 2002. *Statistical Inference*. California: Thomson Learning.

CES. Various Years. "Consumer Expenditure Survey." http://www.bls.gov/cex/home.htm.

Chambers, Raymond L., and Ruilin Ren. 2004. "Outlier Robust Imputation of Survey Data." *The Proceedings of the American Statistical Association*.

Cook, R. Dennis. 1977. "Detection of Influential Observations in Linear Regression." *Technometrics* 19(1): 15–18.

Cook, R. Dennis, and Sanford Weisberg. 1982. *Residuals and Influence in Regression*. New York, New York: Chapman and Hall.

CPS. 2012. "Current Population Survey." http://www.census.gov/cps/.

Daniel, Wayne W. 1990. *Applied Nonparametric Statistics*. Boston, MA: PBS-Kent Publishing.

De Leeuw, Edith D. 2005. "To Mix or Not to Mix Data Collection Modes in Surveys." *Journal of Official Statistics* 21(5): 233–255.

Deming, W. Edwards, and Frederick F. Stephan. 1940. "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Tables are Known." *The Annuals of Mathematical Statistics* 11: 427–444.

Duncan, Greg J., and Graham Kalton. 1987. "Issues of Design and Analysis of Surveys Across Time." *International Statistical Review* 55: 97–117.

Frees, Edward W. 2004. *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*. Cambridge, UK: Cambridge University Press.

Gelman, Andrew, and Hao Lu. 2003. "Sampling Variances for Surveys with Weighting, Post-stratification, and Raking." *Journal of Official Statistics* 19(2): 133–151.

Greene, Claire, Shaun O'Brien, and Scott Schuh. 2017. "U.S. Consumer Cash Use, 2012 – 2015: An Introduction to the Diary of Consumer Payment Choice." Federal Reserve Bank of Boston Research Data Report 17-x.

Hitczenko, Marcin. 2015. "Identifying and Evaluating Selection Bias in Consumer Payment Surveys." Federal Reserve Bank of Boston Research Data Report 15-7.

Jonker, Nicole, and Anneke Kosse. 2009. "The Impact of Survey Design on Research Outcomes: A Case Study of Seven Pilots Measuring Cash Usage in the Netherlands." DNB Working Paper 221. De Nederlandsche Bank.

Little, Roderick J. A., and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data*. New York, New York: Wiley.

Lynn, Peter. 2009. *Methodology of Longitudinal Surveys*. Hoboken, New Jersey: John Wiley and Sons.

Samphantharak, Krislert, Scott Schuh, and Robert M. Townsend. 2017. "Integrated Household Financial Surveys: An Assessment of U.S. Methods and an Innovation." Federal Reserve Bank of Boston Research Data Report 17-7.

SCF. Various Years. "Survey of Consumer Finances." http://www.federalreserve.gov/econresdata/scf/scfindex.htm.

Schmidt, Tobias. 2011. "Fatigue in Payment Diaries - Empirical Evidence from Germany." Discussion Paper Series 1: Economic Studies No 11/2011. Deutsche Bundesbank.

Schuh, Scott, and Joanna Stavins. 2017. "The 2012 Diary of Consumer Payment Choice." Federal Reserve Bank of Boston Research Data Report 17-x.

Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman. 2009. "Forecasting Elections with Non-Representative Polls." *Public Opinion Quarterly* 73(5): 895–916.