Measuring Household Spending and Payment Habits: The Role of "Typical" and "Specific" Time Frames in Survey Questions

Marco Angrisani* and Arie Kapteyn

RAND Corporation

Scott Schuh

Federal Reserve Bank of Boston

April 2012

Abstract

We have designed and fielded an experimental module in the American Life Panel (ALP) where we ask individuals to report the number of their purchases and the amount spent by debit cards, cash, credit cards, and personal checks. The experimental design features several stages of randomization. First, three different groups of sample participants are randomly assigned to an entry month (July, August, or September, 2011) and will be interviewed four times during a year, once every quarter. Second, for each method of payment a sequence of questions elicits spending behavior during a day, week, month, and year. At the time of the first interview, this sequence is randomly assigned to refer to "specific" time spans or to "typical" time spans. In all subsequent interviews, a "specific" sequence becomes a "typical" sequence and vice versa. In this paper, we analyze the data from the first wave of the survey. We show that the type – specific or typical – and length of recall periods greatly influence household reporting behavior.

1 Introduction

The rapid transformation of the U.S. payment system and the increasing availability of new payment instruments have greatly changed household spending habits and use of payment methods. Understanding these trends has important policy implications. First, an assessment of consumers' preferences and financial literacy may help enact regulations, laws, and educational programs to protect and support consumer payment choices. Second, identifying which

*Corresponding author, e-mail: mangrisa@rand.org. We are especially grateful to Angus Deaton, Martha Starr, and Tom Crossley for helpful comments and discussions. We also thank Barry Bosworth, Jennifer Edgar, and participants in the 2011 CRIW/NBER Conference on Improving the Measurement of Consumer Expenditures. Usual disclaimer applies. The views expressed in this paper are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Boston or the Federal Reserve System.

individual characteristics and personal traits drive such preferences and determine spending attitudes is critical to targeting interventions aimed at reducing households' exposure to consumer debt and boosting lifetime savings.

The Survey of Consumer Payment Choice (SCPC), developed by the Federal Reserve Bank of Boston and administered in the RAND American Life Panel (ALP), offers a unique opportunity to study these questions. While it is seldom done in practice, there seem to be clear potential advantages in allowing the respondent to choose the frequency in reporting behavior in surveys. The fundamental reason is that this gives the respondent the flexibility to select a time frame of recall which is best suited to their way of thinking and their habits. The hope is that this will provide more accurate individual results and, thus, more reliable global results. The intuition that certain payments naturally correspond to certain frequencies seems to be verified by the results of the 2010 SCPC. For example, when asked to provide information about cash expenditures in retail, 52.7% of respondents chose the weekly frequency, with only 10.8% answering on a per annum basis. An even stronger example relates to check usage for bill payments, where 67% of respondents answered using the monthly frequency, which might be expected as many bills are due on a monthly basis. However, when adopting such a novel survey approach, it is important to understand the nature of the collected data and how the specifics of the question might influence the response. In the SCPC, those who answered on a weekly basis on average reported 173.3 yearly cash transactions in retail, while those who reported on a monthly basis reported an average of 51.9 and those who reported on an annual basis averaged 11.2. Of course, it might be expected that the choice of reporting frequency is not independent of usage frequency, with those that use a payment type more often finding it easier to think on a weekly basis. However, the differences observed are quite large and it might be that at least part of this is due to bias imposed by the frequency choice.

Measuring the frequency with which people perform regular actions, such as purchasing consumer goods, is not a simple task. The cognitive process used by subjects to answer a frequency question, in fact, may differ substantially depending on the question content and format (Chang and Krosnick, 2003). The SCPC asks respondents about their spending and payment behavior during a "usual" or "typical" period (week, month, or year). This type of question may conceivably trigger a rate-based estimation, in which individuals construct an occurrence rule and apply it to the reference time frame. An alternative approach is to elicit behavior frequency within "specific" time periods, such as past day, week, month, or year. In this case, respondents may be more likely to use episode enumeration, in which they recall and count episodes from a well-specified time frame. The reason for the SCPC to choose "typical" is that its aim is to develop aggregate U.S. estimates of payment use that accurately reflect the trend of payment use. A concern with the use of a specific period is that it has at least two components in it – trend and non-trend, where the latter may include seasonal and other deterministic effects, cyclical effects, and idiosyncratic consumer effects. Using "typical" may help respondents focus on the trends and strip away the other sources of volatility.

Individuals tend to balance effort and accuracy in selecting formulation processes and the trade-off is often determined by the accessibility of the information in memory. The answer to a question about a specific recent period entails shorter-term recall than does one about a typical period and may therefore be subject to smaller recall error. On the other hand, it may represent a less accurate description of average behavioral frequencies, especially when sample sizes are not too large. The issue of determining the optimal recall period has a long history of study in several disciplines (for instance, Mahalanobis and Sen, 1954; Deaton and Kozel, 2005). In the measurement of expenditures, recall periods may vary from one day to a year. Often different periods are chosen for different types of expenditures: long periods for major purchases of durables, for instance, and short periods for small, frequently purchased items. There are various cognitive processes determining the accuracy of retrospective reports including telescoping (events that took place in the past, are reported as more recent than they really were) and straightforward forgetting. The latter is particularly relevant for the measurement of small expenditures. Deaton and Grosh (2000) and Deaton (2001) provide an extensive discussion of the effects of varying recall periods on measured consumption (and its distribution). Assessing the quality and validity of individual reports referring to specific and typical periods of different lengths is an interesting methodological question with important implications for the design of consumer spending surveys and their use for policy analysis.

With this objective in mind, we have designed and fielded an experimental module in the ALP where we ask individuals to report the number of their purchases and the amount spent by debit card, cash, credit card, and check. The experimental design features several stages of randomization. First, three different groups of sample participants are invited every month to answer the survey. Each respondent is randomly assigned to an entry month (July, August, or September, 2011) and is interviewed four times during a year, once every quarter (e.g. the respondents entering in July are re-interviewed in October, respondents entering in August are re-interviewed in November, etc.). Second, for each method of payment a sequence of questions elicits spending behavior during a day, week, month, and year. At the time of the first interview, this sequence is randomly assigned to refer to "specific" time spans or to "typical" time spans. In all subsequent interviews, a "specific" sequence becomes a "typical" sequence and vice versa. Finally, the order of the time frames (day, week, month, year) within a sequence is randomly determined so as to reduce anchoring or order effects.

This design generates both between- and within-subjects variation for our research purposes. In each quarter, we will have one group of respondents answering about specific periods and another group answering about typical periods. Within these two sub-samples, we will compare answers to different reference periods and evaluate the effect of shorter vs. longer recall spans. Also, the randomization of the period sequence (day, week, month, year) will allow us to gauge the degree of dependency among answers referring to different time spans. For instance, is the number of payments in a typical week consistent with the number of payments in a typical day or month? At the same time, we will be able to compare, for a given reference period, reported frequencies within a specific time frame and a typical time frame.

Over two subsequent quarters, we will have individual changes from a specific to a typical period and individual changes from a typical to a specific period. By studying the direction of these changes, we will get insights on whether any of the two formats leads to systematic over- or under-reporting and on whether the"intensity" of the bias differs depending on the length of the reference period (day, week, month, or year).

Over the four planned waves, we will have changes over time for each "specific" and "typical" period. Hence, we can analyze how stable answers are for different question formats. A priori, one would expect reported payment frequencies and spending amounts within typical periods to be less volatile than those within specific periods. Moreover, one would expect such differences to decrease with the length of the reference time frame. Consistency of answers could be treated as an indicator of reliability of the measurements.

An interesting output of this analysis is an assessment of how alternative measures obtained from different question formats correlate with individual characteristics such as education, cognitive ability, and wealth. We will also test the validity of such measures by evaluating their association with criterion variables (i.e. variables with which we expect spending and payment habits to correlate relatively strongly and in a specific way). Possible criterion variables among those already collected by the SCPC are household income, respondents' financial responsibility within the household, individual financial literacy and cognitive capability, and consumers' opinion about the characteristics - security, convenience, acceptance for payment, and cost - of a particular payment instrument.

The first wave of this experimental module has now been completed. In this paper, we describe the experimental design and the characteristics of the sample (Section 2) and provide some preliminary evidence of the role played by time frames when eliciting spending and payment habits in household surveys (Section 3).

Our main findings are two. First, when referring to short reference periods, such as a day or a week, respondents tend to report higher number of payments and amounts spent. Differences between answers to "monthly" and "yearly" questions are relatively small. Second, the probability of reporting non-zero payments by debit cards, cash, and credit cards, is significantly higher when reporting for typical than for specific periods, while there is no differential effect for checks. At the same time, reported amounts spent are systematically lower for typical than for specific reference periods across the four payment instruments.

2 Data and Experimental Design

2.1 The sample

The study is carried out on a sample of individuals participating in the American Life Panel (ALP), an internet-based survey administered by the RAND Corporation. Respondents in the ALP either use their own computer to log on to the Internet or they are provided with a small laptop or a Web TV to access the Internet. About twice a month, sample participants receive an email with a request to visit the ALP URL and fill out specific questionnaires. Typically an interview takes no more than 30 minutes and respondents are paid a monetary incentive proportional to the length of the interview (about 70 cents per minute, or $20 per 30 minutes). Most respondents respond within one week and the vast majority within three weeks. To further increase response rates reminders are sent each week. For the current study, 97% of the sampled individuals completed the survey within one week, 2.5% between two to three weeks, and only 0.5% took four weeks.

There are currently 5,000 members in the ALP mainly recruited from survey programs that collect representative samples of U.S. consumers.[1] For this study we rely on a sample of 3,285 individuals, whose characteristics are summarized in Table 1 below.

Table 1: Sample Characteristics

Gender/Age			Gender/Education			Gender/Income		
	Freq.	Perc.		Freq.	Perc.		Freq.	Perc.
M, Age 18-34	248	7.55	M, High School or less	268	8.16	M, Inc<35k	375	11.45
M, Age 35-54	507	15.43	M, Some College	476	14.49	M, Inc 35-59k	352	10.75
M, Age 55+	578	17.60	M, College+	589	17.93	M, Inc 60k+	601	18.35
F, Age 18-34	475	14.46	F, High School	426	12.97	F, Inc<35k	746	22.78
F, Age 35-54	774	23.56	F, Some College	823	25.05	F, Inc 35-59k	510	15.57
F, Age 55+	703	21.40	F, College +	703	21.40	F, Inc 60k+	691	21.10
Total	3,285	100.00	Total	3,285	100.00	Total	3,275	100.00

2.2 The experiment

About one third of the selected sample is invited every month to answer the experimental module. Each participant is interviewed four times during a year, once every quarter. The first wave of the survey was fielded during the summer of 2011. Specifically, respondents were randomly assigned to three different entry dates – July 15^{th}, August 15^{th}, and September 15^{th} – and are scheduled to be re-interviewed every three months since then. For instance, those who started on July 15^{th} 2011 are asked to take the second wave of the survey on October 15^{th} 2011, the third wave on January 15^{th} 2012, and the fourth wave on March 15^{th} 2012.

Table 2: Randomization 1 – Entry Date

	Freq.	Perc.
July 15^{th}	1,067	32.48
August 15^{th}	1,079	32.85
September 15^{th}	1,139	34.67
Total	3,285	100.00

[1] Until August 2008, most participants were recruited from the pool of individuals age 18 and older who were respondents to the Monthly Survey (MS) of the University of Michigan's Survey Research Center (SRC). The MS is the leading consumer sentiments survey that incorporates the long-standing Survey of Consumer Attitudes (SCA) and produces, among others, the widely used Index of Consumer Expectations. After August 2008, the ALP did not receive new members from the University of Michigan's MS. A subset of participants (approximately 550) have been recruited through a "snowball" sample. That is, respondents were given the opportunity to suggest friends or acquaintances who might also want to participate in the panel. These were then contacted and asked if they wanted to join the ALP. In the fall of 2009, a new group of respondents (approximately 600) was recruited from the National Survey Project (NSP), an NSF-funded panel of Stanford University and Abt SRBI. More recently, the ALP has begun recruiting from a random mail and telephone sample using the Dillman method as well as from vulnerable populations so as to increase the representation of minorities and less affluent individuals.

The survey features questions about the four most common methods of payment adopted by U.S. consumers in recent years, as documented by Foster et al. (2008) and (2009). These are, in order of importance, debit cards, cash, credit cards, and personal checks. For each method of payment, sample participants are asked to report first the number of transactions made and then the amount spent in four recall periods, a day, a week, a month, and a year. At the time of the first interview, each respondent is randomly assigned to answer about "specific past" recall periods or "typical" recall periods. In all subsequent waves, those who answered about "specific past" recall periods in the previous interview are asked to answer about "typical" recall periods and vice versa. Thus, each sample participant faces two possible initial options – "specific past" and "typical" recall periods – and two possible paths over the entire survey originating from them as shown in Table 3.

Table 3: Randomization 2 – "Specific Past" and "Typical" Recall Periods

1st Interview		2nd Interview		3rd Interview		4th Interview
"Specific Past"	$\longrightarrow$	"Typical"	$\longrightarrow$	"Specific Past"	$\longrightarrow$	"Typical"
"Typical"	$\longrightarrow$	"Specific Past"	$\longrightarrow$	"Typical"	$\longrightarrow$	"Specific Past"

After the type of recall period (specific or typical) has been assigned, a further stage of randomization determines, at each interview and for each respondent, the order in which the four payment instruments appear in the questionnaire. Moreover, the order of the recall period sequence (day/week/month) is randomly allocated to each method of payment so as to reduce mechanical answers and systematic anchoring or order effects. Questions referring to the year are always asked after the respondent has reported about all other recall periods.[2] Table 4 illustrates the random assignments.

Our experiment design does not allow the respondent to choose a particular frequency (as in the SCPC), but each survey participant answers about four possible recall periods. This choice prevents us from studying how the rate of payment use (e.g. very frequent use of cash for daily purchases) induces selection into particular time frames (e.g. choosing day as a reference period when answering about cash payments). On the other hand, it enables us to analyze whether reporting behavior exhibits systematic differences for each method of payments across recall periods of different length. It should be noted that blocking questions by payment method and not by recall periods has the advantage of attenuating possible "seam" effects (Rips et al., 2003; Ham et al., 2007; Moore et al., 2009). That is, the tendency of providing relatively similar answers for each recall period within one wave and relatively different answers across waves. This issue may conceivably arise if respondents adopt "constant responding" strategies so as to simplify the reporting task. For instance, when asked about the number of payments in a week,

[2] In a pilot test we randomized the whole period sequence (day/week/month/year). Respondents' feedback revealed strong reluctance to answer the "year" question at the beginning of the recall period sequence. We therefore decided to permute only day, week, and month, while keeping the "year" question at the end of the sequence for each method of payment. We acknowledge that this may cause some anchoring effects. On the other hand, however, it makes it easier for survey participants to approximate the number of payments and the amount spent over a long time span such as one year.

survey participants may be inclined to provide the same answer for all payment instruments in order to minimize the mental effort. Our design should discourage such behaviors and therefore reduce the importance of "seam" effects in our survey.

Table 4: Randomization 3 – Recall Period Sequence and Payment Methods

	Specific Past Period					Typical Period				
	Debit	Cash	Credit	Check	Total	Debit	Cash	Credit	Check	Total
Day/Week/Month	263	257	271	273	1,064	305	267	263	268	1,103
Day/Month/Week	272	261	243	277	1,053	284	287	274	287	1,132
Week/Day/Month	230	272	274	275	1,051	265	282	285	278	1,110
Week/Month/Day	309	277	252	261	1,099	274	276	279	268	1,097
Month/Day/Week	278	274	287	238	1,077	278	255	295	272	1,100
Month/Week/Day	277	288	302	305	1,172	250	289	260	283	1,082
Total	1,629	1,629	1,629	1,629		1,656	1,656	1,656	1,656	

2.2.1 Defining "specific past" recall periods

In this section, we briefly discuss how "specific past" recall periods are defined in our study. A "specific past" day is determined by randomly drawing a number from 1 to 7 which pins down the specific recent day the respondent has to refer to. For example, if the respondent answers the survey on a Tuesday and the random number is 5, he/she will have to refer to the previous Thursday when answering questions about "specific past" day.

An alternative design would be to ask individuals about payments executed during the day prior to the interview. While this choice would reduce the time of recollection and perhaps increase response accuracy, it has a substantial drawback. Since sample participants are more likely to answer the questionnaire during the first three days after receiving the ALP URL, referring to the day prior to the interview would cluster the reference day on specific days of the week and, hence, reduce its representativeness.[3] For this reason, a design that randomly selects a specific day during the week prior to the interview is to be preferred.

The "specific past" week is defined as follows. For each interview date, an algorithm goes back 7 days and pins down the reference week. Thus, if the respondent answers the interview on July 27^{th}, the "specific past" week is defined as the time since July 20^{th}. Similarly, the "specific past" month and "specific past" year are anchored to the interview date. Thus, if the respondent answers the questionnaire on July 27^{th} 2011, the "specific past" month is defined as the time since June 27^{th} 2011, whereas the "specific past" year is defined as the time since July 2010.

[3] Among those who entered the survey on July 15^{th} 2011, 41% answered the survey during the first three days after receiving the ALP URL and 55% during the first five days. Among those who entered the survey on August 15^{th} 2011, 57% answered the survey during the first three days after receiving the ALP URL and 65% during the first five days. Among those who entered the survey on September 15^{th} 2011, 55% answered the survey during the first three days after receiving the ALP URL and 65% during the first five days.

This procedure avoids variation across individuals in the difficulty of their recall task. For instance, if we were to define the "specific past" month as the month prior to the one when the interview took place, we would have two persons, one answering on July 2^{nd} 2011 and one on July 27^{th} 2011, referring both to June 2011 while facing substantially different recollection times.

3 Results

3.1 Descriptive statistics

Summary statistics reported in Tables 5 and 6 reveal interesting results and, when comparison is possible, confirm the findings by Foster et al., (2008) and (2009). Across all instruments, both the median and the average number of reported payments are mostly higher in typical recall periods than in specific ones. Credit cards are somewhat of an exception in that the mean number of credit card payments per year and per month are higher for specific than for typical periods. This reflects a more skewed distribution of the number of payments in specific years and months than in typical ones.

The difference in skewness between specific and typical distributions is most pronounced when we consider the amounts spent. For all four payment instruments and for day, week, and month, average amounts are larger when we ask for specific periods than when we ask for typical ones, while median amounts are smaller. The differences between specific and typical periods decrease as the length of the recall period increases. In fact, when the reference period is a year, differences are rather modest. These patterns point to higher variances in the reported specific amounts than in the typical amounts. This is consistent with the notion that specific amounts are noisier, since these include intertemporal variation that gets smoothed out when asking for typical periods.

Across all possible payment instruments we compute that the median (average) consumer conducts 22 (36) transactions in the previous month, spending \$1,320 (\$1,839). When considering a typical month, we find a median number (average) of payments equal to and 29 (40) and median (average) spending of \$1,300 (\$1,599). Respondents rely more heavily on debit cards and cash to make their transactions, while credit cards and personal checks are the third and fourth most common methods of payment, respectively. As for the amount spent, survey participants indicate using mainly personal checks and credit cards for large purchases and debit cards and cash to pay for relatively smaller amounts. Such rankings appear to be robust to variations in the type and length of the recall period.[4]

[4]The 2007 Survey of Consumer Finance (SCF) is perhaps the closest source of comparable information for the data collected in this study. The SCF, however, only contains information about the adoption of some non-cash payment instruments and the amount spent by credit card. In the 2007 SCF, the percentage of consumers who had adopted debit cards was 67, the percentage of those who had adopted credit cards was 73, and the percentage of those who had adopted checks 89.7. Using answers to typical-year questions, the percentages of ALP respondents reporting a non-zero number of transaction by debit card, debit card and check are 67, 63, and 77, respectively. In the 2007 SCF the average U.S. household made \$850 worth of credit card charges per month. Table 6 shows that the average monthly amount spent by ALP respondents in 2011 using credit cards is roughly \$500 (in current dollars). Although the information collected in the two surveys is not fully comparable (SCF has household as the unit of measurement, while our analysis is based on individuals), these statistics seem reasonably in line, especially after taking into account that households have significantly decreased the use of

Given the randomization of the sequence (day/week/month), our experimental design allows us to assess the degree of dependency among answers referring to different recall periods. For instance, is the number of payments in a specific or typical week consistent with the number of payments in a specific or typical month? Also, is the answer to a particular reference period systematically anchored by the one given in the preceding question? We investigate these issues in Table 7, where, to help the comparison, we express reported values for day, week, and month in yearly equivalents.

Table 5: Number of Payments

		Specific Past Period				Typical Period			
		Day	Week	Month	Year	Day	Week	Month	Year
Debit	*1st quartile*	0	0	0	0	0	0	0	0
	2nd quartile	0	1	3	20	0	2	4	39
	3rd quartile	1	5	12	140	2	5	20	204
	Mean	1	4	13	171	1	5	15	291
	N of obs.	1,460	1,463	1,464	1,445	1,524	1,527	1,525	1,524
Cash	*1st quartile*	0	0	0	0	0	0	0	5
	2nd quartile	0	1	4	24	0	2	5	50
	3rd quartile	1	4	10	100	1	5	15	200
	Mean	1	5	15	152	1	4	15	260
	N of obs.	1,467	1,469	1,464	1,441	1,529	1,529	1,525	1,521
Credit	*1st quartile*	0	0	0	0	0	0	0	0
	2nd quartile	0	0	2	10	0	0	2	12
	3rd quartile	0	3	10	85	1	3	8	108
	Mean	1	3	12	161	1	3	8	135
	N of obs.	1,464	1,464	1,467	1,448	1,529	1,529	1,530	1,530
Check	*1st quartile*	0	0	0	1	0	0	0	4
	2nd quartile	0	0	2	20	0	0	2	24
	3rd quartile	0	2	6	63	0	1	6	60
	Mean	0	2	6	78	0	1	5	105
	N of obs.	1,468	1,470	1,470	1,454	1,528	1,519	1,534	1,527

Statistics are computed excluding the top 1% of the variables' distribution.

Overall, answers to month and year questions are reasonably consistent, while relatively large discrepancies can be observed between spending reports referring to short (day and week) and long (month and year) recall periods. There is also evidence that answers are anchored to

credit cards during the recent economic turmoil.

those given in the preceding question. Particularly for checks, the total number of payments for both specific and typical reference periods is highest for the sequence D/W/M/Y, followed by W/M/D/Y. For debit cards a somewhat similar pattern seems to emerge, but it is less uniform. Looking across reporting periods, we observe that when day is the first reference period, annualized frequencies of payments tend to be higher when based on daily reports.

Table 6: Amount Spent (in current dollars)

		Specific Past Period				Typical Period			
		Day	Week	Month	Year	Day	Week	Month	Year
Debit	*1st quartile*	0	0	0	0	0	0	0	0
	2nd quartile	0	10	150	800	0	35	200	1,200
	3rd quartile	25	200	586	5,000	25	140	600	6,000
	Mean	39	141	430	4,332	17	90	409	4,864
	N of obs.	1,475	1,475	1,466	1,466	1,542	1,542	1,543	1,543
Cash	*1st quartile*	0	0	0	0	0	0	0	30
	2nd quartile	0	20	75	500	0	20	100	1,000
	3rd quartile	15	95	300	2,080	10	70	300	3,000
	Mean	21	81	230	1,981	10	52	200	2,295
	N of obs.	1,472	1,475	1,475	1,475	1,543	1,543	1,543	1,543
Credit	*1st quartile*	0	0	0	0	0	0	0	0
	2nd quartile	0	0	82	750	0	0	100	882
	3rd quartile	0	160	650	6,000	20	100	500	6,000
	Mean	29	162	605	5,677	15	88	477	5,560
	N of obs.	1,475	1,473	1,475	1,475	1,539	1,522	1,540	1,542
Check	*1st quartile*	0	0	0	0	0	0	0	100
	2nd quartile	0	0	240	2,134	0	0	260	2,400
	3rd quartile	0	215	900	9,600	0	100	875	9,000
	Mean	47	252	727	7,282	11	86	634	6,663
	N of obs.	1,475	1,475	1,474	1,475	1,543	1,543	1,543	1,538

Statistics are computed excluding the top 1% of the variables' distribution.

The order of the recall period sequence also influences reported values. An interesting contrast emerges when comparing number of payments for checks and the total value of check payments. The annualized values across the different sequences are perfectly negatively correlated with the annualized frequencies. That is, the higher the reported number, the lower the annualized value. For cash, the amount spent tends to be higher for the "increasing" sequence

Table 7: Mean Values in Yearly Equivalents for Different Recall Period Sequences

Number of Payments

		Specific Past Period				Typical Period			
		Day	Week	Month	Year	Day	Week	Month	Year
Debit	$D/W/M/Y$	612	223	247	175	430	301	225	242
	$W/M/D/Y$	376	381	198	134	394	275	250	255
	$M/W/D/Y$	243	118	164	189	272	139	92	145
Cash	$D/W/M/Y$	226	77	51	53	95	49	111	208
	$W/M/D/Y$	238	171	130	156	354	235	341	421
	$M/W/D/Y$	188	202	144	136	391	181	233	238
Credit	$D/W/M/Y$	197	143	221	136	180	124	88	125
	$W/M/D/Y$	98	92	239	69	88	52	61	56
	$M/W/D/Y$	220	172	136	162	240	163	126	156
Check	$D/W/M/Y$	222	158	112	141	300	242	149	183
	$W/M/D/Y$	98	123	92	110	153	117	97	106
	$M/W/D/Y$	80	75	54	64	76	57	52	56

Amount Spent

		Specific Past Period				Typical Period			
		Day	Week	Month	Year	Day	Week	Month	Year
Debit	$D/W/M/Y$	20,765	7,869	5,139	3,935	7,471	4,328	4,264	3,880
	$W/M/D/Y$	11,065	4,648	2,237	1,776	3,760	2,547	2,263	2,208
	$M/W/D/Y$	16,837	8,302	6,484	5,317	5,836	4,516	5,350	5,515
Cash	$D/W/M/Y$	27,917	12,683	8,710	8,645	4,560	3,341	7,126	6,584
	$W/M/D/Y$	10,649	7,609	5,179	4,153	5,527	5,001	5,848	5,844
	$M/W/D/Y$	6,136	4,022	2,272	1,805	3,427	2,704	2,469	1,862
Credit	$D/W/M/Y$	7,872	11,576	7,887	6,151	5,103	4,836	5,428	5,812
	$W/M/D/Y$	8,652	14,825	10,164	7,520	3,490	4,276	7,110	7,619
	$M/W/D/Y$	7,700	5,902	5,529	5,040	5,040	3,695	4,724	3,827
Check	$D/W/M/Y$	4,998	3,372	2,948	1,949	2,360	2,380	2,376	2,449
	$W/M/D/Y$	5,087	5,437	6,694	4,382	4,875	5,346	6,367	5,592
	$M/W/D/Y$	7,858	12,834	8,547	7,442	3,755	5,715	7,911	6,456

Statistics are computed excluding the top 1% of the variables' distribution. Reported number of payments and amount spent for day, week and month are expressed in yearly equivalents.

day/week/month than for the “decreasing” sequence month/week/day.[5]

3.2 Regression analysis

We now turn to the analysis of the experimental data in a regression framework so as to quantify the effect that different type – specific or typical – and length of recall periods have on household spending habits as elicited by our module. Throughout this section, we will focus on two outcomes: the reported number of payments and the amount spent using one of the four payment methods in a particular time frame. As a preliminary step, we express these two variables in yearly equivalents, whenever the recall period is a day, a week or a month. This transformation will ease the interpretation and help the comparison of estimated coefficients across recall periods of different length.

Given the experimental design described above, we have four individual reports for each method of payment, one per day, one per week, one per month, and one per year. Our strategy is to express these individual reports in yearly equivalents and regress them on question format indicators. We use relatively flexible specifications allowing the length of the reference period to interact with the type of recall frame - specific or typical - and with an indicator for the starting period in the reference period sequence. We control for a set of individual characteristics including gender, age, education, and family income, as well as for survey specific factors such as the time it took the respondent to complete the questionnaire. In order to account for correlation between observations within each individual unit, we cluster standard errors at the respondent level.

In Tables 8 and 9 we focus on the number of payments. Specifically, we first present OLS estimates and then test hypotheses across various question formats.[6] The regression results confirm the patterns of the descriptive analysis in the previous section. Respondents report a substantially higher number of payments when referring to short time spans, such as a day or a week, than when referring to longer spans, such as a month or a year. For instance, the marginal effects (shown in Table 10 below) implied by the regressions in Table 8 reveal that individuals report 51 more debit card payments when referring to a week than to a month, 30 more cash payments, 33 more credit card payments, and 12 more check payments. These differences more than double if we compare reports referring to a day with those referring to a month. On the other hand, there are relatively small discrepancies between frequencies elicited using month and year as reference periods. Comparing typical and specific reference periods, we see that asking for the number of payments with debit cards or cash yields frequencies that

[5] For all the other recall period sequences not reported in Table 7, there are no appreciable differences with respect to the patterns commented above.

[6] Zero payments could reflect either non-adoption of the payment instrument by the respondent or spending inactivity by the respondent; the latter could occur even if the respondent adopted the instrument. Count data models for the number of payments give very similar results to the OLS estimates presented here. Specifically, allowing for unobserved heterogeneity, which would imply overdispersion in the number of reported transactions, we estimate a Negative Binomial model with quadratic variance. Moreover, in order to deal with the large number of reported zeros for short recall periods and/or for less common payment instruments (e.g. personal checks), we consider a Zero-Inflated Negative Binomial Model (Cameron and Trivedi, 1998), for which the process generating zero observations differs from the one producing positive values. The results of these regressions are available upon request.

Table 8: OLS Regressions for Number of Payments

Recall Period	*Version*	*Sequence Starting Period*	Debit	Cash	Credit	Check
Day	**Specific**	*D*	262.2***	216.5***	131.7***	170.7***
			(41.5)	(43.4)	(28.8)	(26.2)
		W	105.7***	77.6*	112.3***	47.7**
			(30.6)	(42.1)	(28.5)	(19.4)
		M	133.5***	82.7**	53.4**	22.0
			(32.0)	(40.8)	(21.3)	(14.9)
Day	**Typical**	*D*	306.4***	186.8***	121.5***	39.8***
			(33.1)	(35.9)	(22.0)	(13.9)
		W	203.9***	160.0***	80.6***	19.6
			(30.1)	(38.2)	(19.1)	(12.9)
		M	160.9***	132.8***	79.2***	17.5*
			(22.6)	(31.3)	(14.2)	(9.8)
Week	**Specific**	*D*	50.8*	2.9	19.3	14.1
			(26.5)	(34.8)	(18.5)	(11.2)
		W	46.2*	8.5	58.3***	55.0***
			(24.3)	(33.9)	(20.6)	(13.8)
		M	9.8	39.9	34.1	14.7
			(24.4)	(40.4)	(21.3)	(9.3)
Week	**Typical**	*D*	132.2***	66.8*	28.1*	-10.6
			(28.2)	(34.8)	(17.0)	(7.9)
		W	75.4***	35.1	15.1	-9.4
			(24.2)	(34.9)	(15.2)	(7.8)
		M	46.9**	71.6**	15.2	-6.0
			(20.8)	(33.1)	(9.8)	(6.3)
Month	**Specific**	*D*	-8.1	35.7	-22.1	-14.8**
			(26.7)	(42.1)	(15.8)	(7.4)
		W	-13.4	0.3	12.3	10.5
			(23.8)	(41.4)	(18.6)	(13.1)
		M	-8.9	-49.0	22.4	1.2
			(26.1)	(34.3)	(22.2)	(10.5)
Month	**Typical**	*D*	73.5***	37.6	-7.0	-8.6
			(25.7)	(34.7)	(14.9)	(7.7)
		W	48.9*	28.3	-16.9	1.2
			(27.5)	(37.2)	(13.7)	(8.6)
		M	-32.6**	-4.4	-19.2**	-6.2
			(13.2)	(24.8)	(8.6)	(6.3)
Year	**Specific**	*D*	34.3	-34.5	15.2	-5.0
			(30.1)	(35.6)	(25.2)	(10.1)
		W	-12.3	-36.2	81.7**	2.9
			(23.7)	(34.1)	(32.9)	(10.7)
		M	28.4	-50.2	4.3	2.4
			(30.5)	(33.2)	(21.6)	(10.4)
Year	**Typical**	*D*	108.5***	74.1*	2.1	-2.6
			(31.4)	(41.2)	(15.9)	(8.8)
		W	73.6**	66.8*	-4.8	0.2
			(30.0)	(40.6)	(15.3)	(9.5)
Number of Observations			11,905	11,918	11,932	11,941

Dependent variable: number of payments in yearly equivalents. Regressions include controls for gender, age, education, family income and survey time. Standard errors are clustered at the individual level. The omitted category is $Year \times Typical \times M$. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 9: OLS Regressions for Number of Payments: Testing Differences across Time Frames

Panel A		Debit	Cash	Credit	Check
Specific	H_0: Day = Week	***	***	***	***
Specific	H_0: Day = Month	***	***	***	***
Specific	H_0: Day = Year	***	***	***	***
Specific	H_0: Week = Month	***	**	***	***
Specific	H_0: Week = Year	**	***	$\circ$	***
Specific	H_0: Month = Year	$\circ$	$\circ$	**	$\circ$
Typical	H_0: Day = Week	***	***	***	***
Typical	H_0: Day = Month	***	***	***	***
Typical	H_0: Day = Year	***	***	***	***
Typical	H_0: Week = Month	***	*	***	$\circ$
Typical	H_0: Week = Year	$\circ$	$\circ$	**	$\circ$
Typical	H_0: Month = Year	**	$\circ$	**	$\circ$
Panel B		Debit	Cash	Credit	Check
Day	H_0: Specific = Typical	**	$\circ$	$\circ$	***
Week	H_0: Specific = Typical	**	$\circ$	$\circ$	***
Month	H_0: Specific = Typical	***	$\circ$	*	$\circ$
Year	H_0: Specific = Typical	***	***	*	$\circ$
Panel C		Debit	Cash	Credit	Check
Day-Specific	H_0: Starting D = Starting W	***	***	$\circ$	***
	H_0: Starting D = Starting M	***	***	**	***
	H_0: Starting W = Starting M	$\circ$	$\circ$	*	$\circ$
Day-Typical	H_0: Starting D = Starting W	***	$\circ$	*	$\circ$
	H_0: Starting D = Starting M	***	$\circ$	*	$\circ$
	H_0: Starting W = Starting M	$\circ$	$\circ$	$\circ$	$\circ$
Week-Specific	H_0: Starting D = Starting W	$\circ$	$\circ$	*	***
	H_0: Starting D = Starting M	$\circ$	$\circ$	$\circ$	$\circ$
	H_0: Starting W = Starting M	$\circ$	$\circ$	$\circ$	***
Week-Typical	H_0: Starting D = Starting W	*	$\circ$	$\circ$	$\circ$
	H_0: Starting D = Starting M	***	$\circ$	$\circ$	$\circ$
	H_0: Starting W = Starting M	$\circ$	$\circ$	$\circ$	$\circ$
Month-Specific	H_0: Starting D = Starting W	$\circ$	$\circ$	*	**
	H_0: Starting D = Starting M	$\circ$	**	**	*
	H_0: Starting W = Starting M	$\circ$	$\circ$	$\circ$	$\circ$
Month-Typical	H_0: Starting D = Starting W	$\circ$	$\circ$	$\circ$	$\circ$
	H_0: Starting D = Starting M	***	$\circ$	$\circ$	$\circ$
	H_0: Starting W = Starting M	***	$\circ$	$\circ$	$\circ$
Year-Specific	H_0: Starting D = Starting W	$\circ$	$\circ$	*	$\circ$
	H_0: Starting D = Starting M	$\circ$	$\circ$	$\circ$	$\circ$
	H_0: Starting W = Starting M	$\circ$	$\circ$	**	$\circ$
Year-Typical	H_0: Starting D = Starting W	$\circ$	$\circ$	$\circ$	$\circ$
	H_0: Starting D = Starting M	***	*	$\circ$	$\circ$
	H_0: Starting W = Starting M	**	*	$\circ$	$\circ$

Tests use estimates from OLS regressions in Table 8. The reference distribution in Panels A and B is χ^2_3; the reference distribution in Panel C is $N(0,1)$. ***, ** and * indicate that the null H_0 is rejected at the 1%, 5% and 10% level, respectively. $\circ$ indicates that the null H_0 is not rejected.

Table 10: OLS Regressions for Number of Payments: Marginal Effects

	Debit	Cash	Credit	Check
Week	-135.8***	-104.1***	-67.3***	-43.4***
	(11.6)	(14.0)	(7.1)	(6.4)
Month	-186.1***	-134.4***	-100.3***	-55.5***
	(12.1)	(15.2)	(8.3)	(6.4)
Year	-157.1***	-137.8***	-79.9***	-53.0***
	(13.2)	(15.0)	(9.9)	(6.7)
Typical	47.8***	47.5***	-18.8*	-23.9***
	(13.6)	(15.5)	(10.0)	(5.2)
Starting W	-54.0***	5.9	-27.3**	2.2
	(16.8)	(18.8)	(11.9)	(6.2)
Starting M	-78.5***	-5.1	-5.8	2.7
	(16.7)	(19.0)	(12.8)	(6.2)

Marginal effects after the OLS regressions in Table 8. Omitted categories are: "Day" for the length of the reference period; "Specific" for the type of reference period; "Starting D" that is reference period sequence starting with day. Standard errors are clustered at the individual level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

are about 48 higher when referring to typical periods than when we use specific periods; on the other hand for credit cards and checks, typical periods yield respectively 18 and 24 fewer reports per annum than when asking for frequencies in specific periods. The hypothesis tests in Panel A of Table 9 show that these differences are highly significant.

Given the mixture of observations with zero and positive values for spending amounts and its different balance across the various methods of payment, we estimate a Hurdle model for the reported amount spent. Compared to OLS, this approach allows to relax the assumption that zero payments and positive amounts spent are produced by the same data generating process.[7] Specifically, indicating with y_1 the number of payments and with y_2 the amount spent, we model the conditional probability of a non-zero payment as a Probit:

$$\Pr\left[y_1 > 0|\mathbf{x}\right] = \Phi\left(\mathbf{x}'\beta\right), \tag{1}$$

and the expected value of a positive reported amount as a linear function

$$E_{y_2>0}\left[y_2|y_1 > 0, \mathbf{x}\right] = \mathbf{x}'\gamma. \tag{2}$$

The unconditional mean for the amount spent is therefore:

$$E\left[y_2|\mathbf{x}\right] = \Phi\left(\mathbf{x}'\beta\right) \times \mathbf{x}'\gamma. \tag{3}$$

We separately estimate equations (1) and (2) and compute the "combined" marginal effects for a discrete explanatory variable x_j using

$$E\left[y_2|\mathbf{x}\right]_{x_j=1} - E\left[y_2|\mathbf{x}\right]_{x_j=0} = \left[\Phi\left(\mathbf{x}'\beta\right) \times \mathbf{x}'\gamma\right]_{x_j=1} - \left[\Phi\left(\mathbf{x}'\beta\right) \times \mathbf{x}'\gamma\right]_{x_j=0}. \tag{4}$$

[7] Model specifications addressing these issues are discussed, among others, by Deaton and Irish (1984), Blundell and Meghir (1987), Chesher and Irish (1987), and Robin (1993).

In Tables 11 and 13 we report average partial effects defined as:

$$\frac{1}{n}\sum_{i=1}^{n}\left\{\left[\Phi\left(\mathbf{x}_i'\hat{\beta}\right)\times\mathbf{x}_i'\hat{\gamma}\right]_{x_ij=1}-\left[\Phi\left(\mathbf{x}_i'\hat{\beta}\right)\times\mathbf{x}_i'\hat{\gamma}\right]_{x_ij=0}\right\}. \tag{5}$$

with i indicating the i^{th} observation from a sample of size n.[8]

The estimated coefficients of the Hurdle model provide some insights on the mechanisms driving reporting behaviors. First, as one would expect, the probability of reporting a positive number of payments increases with the length of the reference period. However, the extent to which this happens varies substantially across payment methods. The likelihood of reporting positive purchases by debit card when referring to a week, month, and year is, respectively, 16, 23, and 27 percentage points higher than when referring to a day. For transactions using checks, differences are on the order of 30, 55, and 60 percentage points. Within a "typical" framework the probability of reporting positive purchases increases by 9 percentage points for debit cards and cash and by 4 percentage points for credit card. On the other hand, there is no differential effect for personal checks.

Second, conditional on non-zero payments, answering about short recall periods significantly increases the reported amount in yearly equivalents. After computing the marginal effects implied by the estimates in Table 11, we find that, when they refer to a week, respondents report about \$2,500 more spent by debit card and cash, \$3,000 more spent by credit card, and \$6,000 spent by check than when they refer to a month. These differences are much more pronounced when answers to questions about the day are compared to those about the month. On the other hand, less marked discrepancies are observed between answers to a month and to a year, ranging from \$1,000 for debit cards to \$2,000 for checks.

Third, with the exception of checks, a typical framework increases the probability of reporting non-zero payments by 8-9 percentage points. At the same time, it lowers the reported amount spent, conditional on it being positive. Specifically, individuals who conduct a non-zero number of transactions report \$9,000 less spent by check, \$6,500 less spent by debit and credit card, and \$3,000 less spent in cash when they are asked to refer to a typical rather than to a specific past period (comparison of average partial effects for specific and typical periods computed taking all interactions into account).

The combination of these mechanisms produces the results in Table 12. Panel A shows that the length of the reference period greatly affects household reporting behavior. Answers to shorter time spans are systematically different from those to longer ones. Within either a specific or a typical framework, this is true across all four payment instruments. Discrepancies between answers to monthly and yearly questions tend to be economically less sizeable and not statistically significant when respondents are asked to refer to typical periods.

Panel B in Table 12 reveals that the question frame matters as long as the length of the reference period is short enough. That is, answers referring to a specific day or week are systematically different from those referring to a typical day or week. On the other hand,

[8] Estimated coefficients for the Probit model in equation (1) and the OLS regression in equation (2) are available upon request.

Table 11: Hurdle Model – Average Partial Effects

Recall Period	*Version*	*Sequence Starting Period*	Debit	Cash	Credit	Check
Day	**Specific**	*D*	5.16***	1.68***	1.11	0.92
			(1.26)	(0.58)	(0.99)	(1.31)
		W	0.40	-0.02	-0.30	-3.16***
			(0.92)	(0.46)	(0.88)	(1.03)
		M	0.80	-0.02	-1.43*	-3.27***
			(0.93)	(0.45)	(0.80)	(1.02)
Day	**Typical**	*D*	0.08	-0.55**	-1.75***	-6.29***
			(0.56)	(0.27)	(0.54)	(0.41)
		W	-1.32**	-0.80***	-2.30***	-6.56***
			(0.52)	(0.26)	(0.50)	(0.41)
		M	-1.01***	-1.07***	-2.07***	-6.52***
			(0.39)	(0.20)	(0.38)	(0.36)
Week	**Specific**	*D*	1.63**	0.41	0.62	-0.57
			(0.81)	(0.40)	(0.79)	(0.90)
		W	2.59***	1.34***	2.76***	3.58***
			(0.82)	(0.44)	(0.87)	(1.18)
		M	0.08	0.01	-0.28	0.10
			(0.68)	(0.33)	(0.66)	(1.02)
Week	**Typical**	*D*	-0.48	-0.65***	-1.36***	-5.03***
			(0.48)	(0.24)	(0.50)	(0.43)
		W	-0.63	-0.17	-1.66***	-4.66***
			(0.49)	(0.28)	(0.48)	(0.46)
		M	-0.99***	-0.40**	-0.84***	-4.27***
			(0.29)	(0.16)	(0.32)	(0.36)
Month	**Specific**	*D*	-0.27	-0.08	0.33	-0.67
			(0.58)	(0.31)	(0.66)	(0.68)
		W	0.25	-0.44	1.47**	0.92
			(0.60)	(0.29)	(0.71)	(0.88)
		M	1.10*	0.15	1.89**	0.94
			(0.66)	(0.34)	(0.74)	(0.88)
Month	**Typical**	*D*	-0.12	-0.41*	0.05	-0.89
			(0.50)	(0.23)	(0.55)	(0.63)
		W	0.14	-0.34	-0.64	-0.13
			(0.52)	(0.25)	(0.51)	(0.69)
		M	-0.11	-0.41***	-0.05	0.03
			(0.24)	(0.13)	(0.24)	(0.35)
Year	**Specific**	*D*	-0.48	-0.73***	0.29	0.30
			(0.59)	(0.26)	(0.64)	(0.78)
		W	-0.17	-0.70***	0.84	0.65
			(0.60)	(0.25)	(0.63)	(0.86)
		M	0.47	-0.43	0.24	-0.48
			(0.62)	(0.29)	(0.59)	(0.68)
Year	**Typical**	*D*	0.05	-0.16	0.30	-0.45
			(0.51)	(0.25)	(0.58)	(0.65)
		W	0.37	-0.49**	-0.43	0.17
			(0.54)	(0.24)	(0.52)	(0.69)
Number of Observations			12,021	12,048	12,043	12,046

"Combined" average partial effects from Probit and OLS regressions are reported. The dependent variable for Probit is an indicator for non-zero number of payments. The dependent variable for OLS is the amount spent in yearly equivalents expressed in 1,000 dollars. Regressions include controls for gender, age, education, family income and survey time. The omitted category is $Year \times Typical \times M$. Bootstrap standard errors (500 replications) are clustered at the individual level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 12: Testing Differences across Time Frames (Hurdle Model)

Panel A		Debit	Cash	Credit	Check
Specific	H_0: Day = Week	***	***	***	***
Specific	H_0: Day = Month	***	***	***	***
Specific	H_0: Day = Year	***	***	**	***
Specific	H_0: Week = Month	***	***	***	**
Specific	H_0: Week = Year	***	***	**	**
Specific	H_0: Month = Year	○	***	***	**
Typical	H_0: Day = Week	**	***	***	***
Typical	H_0: Day = Month	***	***	***	***
Typical	H_0: Day = Year	***	***	***	***
Typical	H_0: Week = Month	***	○	***	***
Typical	H_0: Week = Year	***	***	***	***
Typical	H_0: Month = Year	○	***	○	○

Panel B		Debit	Cash	Credit	Check
Day	H_0: Specific = Typical	***	***	***	***
Week	H_0: Specific = Typical	***	***	***	***
Month	H_0: Specific = Typical	○	○	***	○
Year	H_0: Specific = Typical	○	*	○	○

Panel C		Debit	Cash	Credit	Check
Day-Specific	H_0: Starting D = Starting W	***	***	○	***
	H_0: Starting D = Starting M	***	***	**	***
	H_0: Starting W = Starting M	○	○	○	○
Day-Typical	H_0: Starting D = Starting W	**	○	○	○
	H_0: Starting D = Starting M	*	*	○	○
	H_0: Starting W = Starting M	○	○	○	○
Week-Specific	H_0: Starting D = Starting W	○	*	**	***
	H_0: Starting D = Starting M	*	○	○	○
	H_0: Starting W = Starting M	***	***	***	***
Week-Typical	H_0: Starting D = Starting W	○	*	○	○
	H_0: Starting D = Starting M	○	○	○	*
	H_0: Starting W = Starting M	○	○	○	○
Month-Specific	H_0: Starting D = Starting W	○	○	○	*
	H_0: Starting D = Starting M	**	○	**	*
	H_0: Starting W = Starting M	○	○	○	○
Month-Typical	H_0: Starting D = Starting W	○	○	○	○
	H_0: Starting D = Starting M	○	○	○	○
	H_0: Starting W = Starting M	○	○	○	○
Year-Specific	H_0: Starting D = Starting W	○	○	○	○
	H_0: Starting D = Starting M	○	○	○	○
	H_0: Starting W = Starting M	○	○	○	○
Year-Typical	H_0: Starting D = Starting W	○	○	○	○
	H_0: Starting D = Starting M	○	○	○	○
	H_0: Starting W = Starting M	○	**	○	○

Tests use estimates from the hurdle model in Table 11. The reference distribution in Panels A and B is χ^2_3; the reference distribution in Panel C is $N(0,1)$. ***, ** and * indicate that the null H_0 is rejected at the 1%, 5% and 10% level, respectively. ○ indicates that the null H_0 is not rejected.

answers about month and year are fairly similar independently of the question frame. The tests in Panel C confirm that the order of the reference period sequence has very little effect on individual answers. We only find evidence that respondents report higher frequencies and amounts when they are asked about daily payments and the day features as first in the sequence of reference periods. Respondents exhibit a similar behavior when they are asked to recall payments during a specific past week and the sequence of reference periods starts with week instead of month.

Since different question frames affect the propensity with which positive payments are reported, treatment variables in equation (2) could potentially be correlated with unobserved characteristics driving reporting behavior. In other words, if there is selection on unobservables, the estimated coefficients on treatment variables in equation (2) may be biased. A Heckman's selection model would allow for selection on unobservables. The absence of plausible exclusion restrictions, however, makes the estimation of such a model entirely dependent on functional form assumptions. Rather than relying on arbitrary exclusion restrictions, we prefer a different approach. As is well-known, if the errors in the Probit equation and the amount equation are correlated this leads to the addition of a Mills ratio to (2), where its coefficient is the product of the correlation between the error terms and the standard deviation of the error term in the amount equation. We calculate the Mills ratio from the Probit equation and add it to (2). Next we vary the size of the correlation coefficient from 0 to 1. We find that although the estimated marginal effects do vary as the size of the correlation coefficient increases, these changes are not dramatic and in no case is the sign of a statistically significant coefficient reverted.[9]

In Table 13 we report the estimated coefficients for the control variables used in the Hurdle model regressions.[10] The coefficients on income and education have the expected sign. Compared to those whose income is less than $35,000 and accounting for the probability of reporting non-zero payments, individuals with more than $60,000 spend $2,000 more by debit card and about $5,500 more by credit card and check. At the same time they rely substantially less on cash payments spending, on average, $700 less. Having a college degree appears to have a combined positive effect for credit card and check payments, but it seems to have no impact on the use of debit cards and cash.

The estimated coefficients on age dummies reveal an interesting pattern too. Relatively older respondents are found to use debit cards and cash less frequently, while relying more on personal checks.[11] Specifically, being in the group of those age 55 and over decreases the amount spent by debit card by $3,500, but increases the amount spent using checks by $5,300.

A further interesting result is the effect of survey time on reported payment frequencies and spending habits. As mentioned above, we include in our regression a control for the time taken by the respondent to complete the questionnaire.[12] We observe a strong, positive relationship

[9]For correlation values up to 0.4, estimated marginal effects change very little. For larger values of the correlation parameter, some of the magnitudes change substantially more, but that is only true for a small minority of (typically not statistically significant) coefficients. The results of this exercise are available upon request.

[10]The same set of controls was used for the OLS regressions commented above, but the corresponding estimated coefficients were omitted for brevity.

[11]This is consistent with the trends in the use of paper checks documented by Schuh and Stavins (2010).

[12]We computed that the questionnaire could be completed in 5 to 10 minutes, depending on the number

Table 13: Hurdle Model – Individual Characteristics

	Debit	Cash	Credit	Check
Female	0.82**	-0.44**	-0.82**	-0.58
	(0.33)	(0.19)	(0.39)	(0.49)
Age 35-54	-0.50	-0.48*	0.11	2.30***
	(0.42)	(0.26)	(0.51)	(0.66)
Age 55+	-3.54***	-1.30***	0.31	5.33***
	(0.37)	(0.26)	(0.49)	(0.74)
Inc 35-59k	1.66***	-0.96***	0.73	3.51***
	(0.40)	(0.22)	(0.47)	(0.60)
Inc 60k+	2.07***	-0.67***	5.53***	5.92***
	(0.42)	(0.22)	(0.51)	(0.63)
Some College	0.87*	-0.52*	0.07	0.14
	(0.46)	(0.27)	(0.49)	(0.64)
College+	-0.35	0.01	4.35***	1.79***
	(0.46)	(0.27)	(0.52)	(0.67)
ST q2	1.38***	0.68***	2.39***	1.75***
	(0.49)	(0.26)	(0.56)	(0.67)
ST q3	1.00**	1.11***	3.43***	3.12***
	(0.45)	(0.29)	(0.57)	(0.76)
ST q4	1.92***	1.08***	4.37***	6.26***
	(0.51)	(0.32)	(0.67)	(0.94)

Average partial effects for the control variables used in the hurdle model regression (Table 11). $ST\ q(k)$ is an indicator for the k^{th} quartile of the survey time distribution. The omitted categories are $Income < 35k$, $Education \leq High\ School$, $18 \leq Age < 35$, the indicator for $Survey\ Time \leq q1$. Bootstrap standard errors (500 replications) are clustered at the individual level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

between such a variable and both the likelihood of reporting non-zero payments and the amount spent conditional on it being positive. These two effects produce sizeable and statistically significant coefficients for the survey time indicators in Table 13. For instance, passing from the first quartile ($ST\ q1$ corresponding to 5 minutes) of the survey time distribution to the fourth ($ST\ q4$ corresponding to 14 minutes) increases the reported amount of debit card charges by \$2,000 and the one of credit card charges by \$4,400. Needless to say, these effects are not necessarily causal. Someone who reports more transactions may need more time to think about the correct number of transactions and the correct total amount than someone whose total number of transactions is lower.

4 Conclusion

In this paper we investigate the role of different time frames (specific or typical recall periods of different length) in survey questions measuring household payment and spending habits. For this purpose, we have designed and fielded an experimental module in the American Life Panel

of payment instruments adopted by the respondent. This is confirmed by the data. The median respondent answered in 8 minutes, while respondents at the first and third quartile of the survey time distribution answered in 5 and 14 minutes, respectively. In our analysis we exclude all those who completed the questionnaire in less than 2 minutes – 48 – and those who did so over multiple days – 187 (in the ALP respondents can pause the survey and resume it later as long as the survey is still "open").

(ALP) where we ask individuals to report the number of their purchases and the amount spent using four common payment instruments, debit cards, cash, credit cards, and personal checks. Three different groups of sample participants are randomly assigned to an entry month (July, August, or September, 2011) and interviewed four times during a year, once every quarter. For each method of payment, a sequence of questions elicits spending behavior during a day, week, month, and year. At the time of the first interview, this sequence is randomly assigned to refer to "specific" time spans or to "typical" time spans. In all subsequent interviews, a "specific" sequence becomes a "typical" sequence and vice versa.

Accounting for all possible payment instruments we compute that the median (average) consumer makes 22 (36) transactions in the previous month, spending $1,320 ($1,839). In comparison, when asked to refer to a typical month, respondents report 29 (40) transactions, spending $1,300 ($1,599). Respondents rely more heavily on debit cards and cash to make their transactions, while credit cards and personal checks are used less frequently to pay for relatively large expenses.

Regression analysis shows that, when referring to short reference periods, such as a day or a week, respondents tend to report higher numbers of payments and amounts spent. Differences between answers to "monthly" and "yearly" questions are relatively small. Within a "typical" framework the probability of reporting non-zero payments increases significantly for debit cards, cash, and credit cards, while there is no differential effect for checks. At the same time, reported amounts spent are systematically lower for "typical" than for "specific" reference periods across the four payment instruments.

The present analysis is very preliminary as it only uses the data from the first completed wave of our survey. Further evidence will be provided as data from subsequent waves will become available. Notably, given our experimental design, we will exploit in the future both cross-section and within-subject variations to assess the effect of different time frames on individual reporting behavior.

References

[1] Blundell, R. and Meghir, C. (1987), "Bivariate Alternatives to the Tobit Model," *Journal of Econometrics*, Vol. 34, pp. 179-200.

[2] Cameron, A. C. and Trivedi, P. K. (1998), "Regression Analysis of Count Data." Cambridge University Press, Cambridge.

[3] Chang, L. and Krosnick, J. A. (2003), "Measuring the Frequency of Regular Behaviors: Comparing the Typical Week to the Past Week," *Sociological Methodology*, Vol. 33, pp. 55-80.

[4] Chesher, A. and Irish, M. (1987), "Residual Analysis in the Grouped and Censored Normal Linear Model," *Journal of Econometrics*, Vol. 34, pp. 33-61.

[5] Deaton, A. (2001), "Survey Design and Poverty Monitoring in India," *Research Program in Development Studies*, Princeton University.

[6] Deaton, A. and Grosh, M. (2000), "Consumption," in *Designing Household Survey Questionnaires for Developing Countries: Lessons from ten Years of LSMS Experience*, edited by Grosh, M. and Glewwe, P., Chapter 5, pp.91-133.

[7] Deaton, A. and Irish, M. (1984), "Statistical Models for Zero Expenditures in Household Budgets," *Journal of Public Economics*, Vol. 23, pp. 59-80.

[8] Deaton, A. and Kozel, V. (2005), "Data and Dogma: The Great Indian Poverty Debate," *The World Bank Research Observer*, Vol. 20(2), pp. 177-199.

[9] Foster, K., Meijer, E., Schuh, S., and Zabeck, M. (2009), "The 2009 Survey of Consumer Payment Choice," Public Policy Discussion Paper, Federal Reserve Bank of Boston.

[10] Foster, K., Meijer, E., Schuh, S., and Zabeck, M. (2008), "The 2008 Survey of Consumer Payment Choice," Public Policy Discussion Paper, Federal Reserve Bank of Boston.

[11] Ham, J., Li, X., and Sheppard, L. (2007), "Correcting for Seam Bias when Estimating Discrete Variable Models, with an Application to Analyzing the Employment Dynamics of Disadvantaged Women in the SIPP," *mimeo*, USC.

[12] Moore, J., Bates, N., Pascale, J., and Okon, A. (2009), "Tackling seam bias through questionnaire design," in *Methodology of Longitudinal Surveys*, , edited by Lynn, P., John Wiley & Sons, Ltd.

[13] Rips, L., Conrad, F., and Fricker, S. (2003), "Straightening the Seam Effect in Panel Surveys," *The Public Opinion Quarterly*, Vol. 67(4), pp. 522-554.

[14] Robin, J. M. (1993), "Econometric Analysis of the Short-run Fluctuations of Households Purchases," *Review of Economic Studies*, Vol. 60, pp. 923-934.

[15] Schuh, S and Stavins, J. (2010), "Why are (some) consumers (finally) writing fewer checks? The role of payment characteristics," *Journal of Banking & Finance*, Vol. 34, pp. 1745-1758.