

DNB Working Paper

No. 546 / February 2017

Survey shortcuts? Evidence from a payment diary survey

Lola Hernandez, Robbert-Jan 't Hoen and Juanita Raat

DeNederlandscheBank

EUROSYSTEEM

Survey shortcuts? Evidence from a payment diary survey

Lola Hernandez, Robbert-Jan 't Hoen and Juanita Raat *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 546

February 2017

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

Survey shortcuts? Evidence from a payment diary survey^{*}

Lola Hernandez^a, Robbert-Jan 't Hoen^b and Juanita Raat^b

^a *De Nederlandsche Bank, The Netherlands*

^b *Erasmus School of Economics, Erasmus University Rotterdam*

16 February 2017

Abstract

In this paper we examine the presence of panel conditioning in a payment diary survey studying the payment behavior of Dutch consumers. We analyze whether the reporting behavior of frequent participants in the payment survey systematically differs from that of less-frequent survey participants. We introduce refreshment groups that allow us to compare the reporting behavior of ‘trained’ and ‘fresh’ respondents, where the differences between reported values are used as a proxy for the panel conditioning effect. We find no consistent significant differences for the number and value of cash and debit card payments between trained and fresh respondents, for any demographic group, sector or transaction size. Likewise we find no consistent significant differences for the value of cash withdrawals between trained and fresh respondents. However we do find some signs of panel conditioning in the reported number of cash withdrawals.

Keywords: payment behavior, diary studies, survey design, panel conditioning.

JEL classifications: C81, C93, D12.

^{*} Lola Hernandez is the corresponding author, e-mail: L.Hernandez@dnb.nl. Views expressed in this paper do not necessarily coincide with those of De Nederlandsche Bank. We thank participants in the DNB research seminar and Hans Brits for their comments on a previous version of this paper.

1 Introduction

Understanding consumers' payment behavior and payment choices is important for central banks and relevant stakeholder of the payment system. It allows them to come up with improvements to the continuously changing payment landscape and ultimately increase the efficiency of the system as a whole. One of the big challenges when analyzing payment choices is to have accurate information about cash usage.

Cash has been traditionally the most widely used payment instrument and yet one of the most challenging in terms of measurement. Unlike electronic payments, cash payments are not centrally registered. Estimating cash usage is an imprecise process due to its unique properties of being an anonymous, untraceable as well as a storable payment instrument (Boeschoten, 1992). We therefore rely on survey techniques to collect payment data from consumers. A common approach is to use payment diaries asking consumers to register their payments over one or more days. In order to estimate cash usage by taking into consideration the heterogeneity of consumers, as well as the fact that payment behavior varies over the year, the appropriate approach would be to contact individuals more than once a year, e.g. every three or four months. This paper addresses the question of whether this approach could incorporate additional biases that affect the survey results.

Earlier literature on survey methods points to the risk of panel conditioning which may lead to respondents changing their payment or reporting behavior. The objective of this paper is to examine panel conditional effects in the survey on consumers' payments at the point-of-sale (SCP) commissioned by *De Nederlandsche Bank* (DNB) and the *Dutch Payment Association* (DPA). The survey was held in September and November 2013, and March 2014 among a total of 21,954 respondents. The contribution of this paper is threefold. First, it assesses whether the payment behavior of respondents, who participated in the same survey two or four months before, differs from the registered behavior of those participating for the very first time. Second, it makes an important contribution to the existing literature as it is the first study on consumers' behavior that analyzes panel conditioning effects using a significantly large sample size that allows meaningful conclusions to be drawn. Third, it enables DNB and the DPA to take a well-considered decision on how to select and sample the respondents for their annual study to measure cash usage in the Netherlands.

Findings show that there is no significant evidence of a panel conditioning effect in the number and the value of payments that respondents report. Neither is there evidence of panel conditioning in the value of withdrawals nor in the proportion of respondents' who do not report any payments at all. Results show that, overall, participants' reporting behavior does not change when they are interviewed multiple times in a year. However, we do find some evidence of panel conditioning in the proportion of respondents who do not report any cash withdrawals at all. We find that trained respondents are significantly more likely to report zero withdrawals than first-time participants in both, the second and third wave of the study (November and March, respectively).

The paper proceeds as follows: section 2 presents the literature review, section 3 presents the paper's objective, section 4 describes the data and methodology, section 5 presents descriptive statistics on the reported number and value of payments, cash withdrawals and the number of respondents who do not report any payments at all. Section 6 presents the results and section 7 concludes.

2 Literature Review

Diary surveys have gained popularity in the last few decades because of high reporting levels and good external validity, which is the extent to which the results of a study can be generally applied to other situations and to other people (Verbrugge, 1980; Cheak-Zamora, Wyrwich, & McBride, 2009).

Much research has been conducted on cross-section surveys on consumers' behavior, whereas panel surveys are relatively new to the field. The first panel surveys stem from the middle of the last century. Back then, panel surveys were mostly used in electoral surveys, like the *Colombia University Study* conducted by Lazarsfeld, Berelson, and Gaudet (1944), who monitored the changes of people's voting preferences during a presidential campaign. Panel surveys are also used for transportation planning as discussed by Kitamura (1990) and medical expenditures (Cohen et al., 1996).

The change from cross-section to panel data has both advantages and disadvantages. The main advantage is the increase in information, which allows for more efficient econometric estimates (due to the increase in degrees of freedom) (Hsiao, 2007). Furthermore, a panel dataset also has an increased capacity to capture the complexity of human behavior and to find time-varying relations. Although these advantages look convincing, possible drawbacks should also be carefully considered, such as panel conditioning and other time dependent biases (Neter & Waksberg, 1964; Golob & Meurs, 1986; Sharot, 1991).

Panel conditioning is the phenomenon that arises if respondents are influenced by having participated in a previous wave of a survey, whereby their answers differ systematically from those who participate for the first time in the same survey (Das, Toepoel, & van Soest, 2011). One reason for panel conditioning is that respondents gain a better understanding of a survey's questions after their first participation in the same survey. As a result, reported answers of experienced participants are a better reflection of the true answers which consequently reduces the measurement error. However, panel conditioning does not always lead to smaller measurement errors. For example, respondents who have been interviewed multiple times learned that they can shorten the length of the interview by avoiding follow up questions (Meurs, Van Wissen, & Visser, 1989). As a result of this, respondents change their reporting behavior and the obtained answers become less representative. Another frequently cited reason is known as the cognitive stimulus hypothesis. Being interviewed multiple times stimulates respondents to change their real behavior. This is based on the idea that first time respondents lack knowledge in the topic of the survey and eventually develop affinity with it. The time between the first and second participation stimulates respondents to further deliberate on the subject of the survey. The result is that respondents change their actual attitude toward the subject of the survey (Sturgis, Allum, & Brunton-Smith, 2009).

The research results of studies exploring panel conditioning are mixed. Alessie, Gradus, and Melenberg (1990) show that the small cash expenditures are the most difficult to keep track of and are therefore more prone to underreporting. Yan and Copeland (2010) did a study on underreporting of total expenditure and expenditure types in the *Consumer Expenditure Survey*. They used the difference between aggregated values in different waves as a proxy for the panel conditioning effect, but found no significant evidence of underreporting due to panel conditioning. In this study the non-respondents were excluded from the analysis. Silberstein and Jacobs (1989) did a similar study and compared the mean expenditure from the *Consumer Expenditure Survey*, a five wave panel survey. They found no evidence of change in mean expenditures due to panel conditioning, when comparing all types of purchases at once. This result is in line with the findings of Yan and Copeland. However, when dividing the purchase types in subgroups, they found significant changes in the reported expenditures. Yet, the changes in reporting behavior were inconsistent, and not substantively large.

Another type of panel conditioning arises due to non-responding, referred as "non-buying" in case of payment surveys, which is largely discussed by De Leeuw (1997). Non-respondents are people who are included in the sample of a survey, but

who do not respond at all or do not respond a specific question. This non-responding of a participant can occur at the outset or start somewhere during the process of the research. Non-responding in surveys poses a serious threat as it affects the data in two ways. First, it introduces a bias in estimated parameters, and second, it increases the total variance of estimates because the actual sample size is lower than expected. The *American Association of Public Opinion Research* warns that the group of non-respondents in survey samples is growing as of the 1990's (Bradburn, 1992). Shields and To (2005) found partial evidence that respondents who reported a purchase in a previous interview were more prone to non-responding.

3 Research objective

The objective of this paper is to explore the existence of panel conditioning effects in the survey of consumers' payments (SCP) once survey respondents participate more than once a year in the same survey. The main purpose of this study is to find support for or to reject the hypothesis of panel conditioning in the diary panel obtained from the experimental three wave diary survey on cash expenditures. If there is evidence of panel conditioning we will look at the effect regarding: a) the number and value of cash & debit card transactions, with a main focus on cash, b) number and value of cash withdrawals and c) non-buyers (respondents not reporting any transactions at all from the beginning or from half way through the experiment). Furthermore, we are interested in whether the bias, due to panel conditioning, tends to concentrate in certain transaction values, sectors, or demographic characteristics (age, district, ethnicity, education, gender, and income).

3.1 Hypotheses

Base on the available literature on panel conditioning and expenditure behavior we hypothesize the following:

HYPOTHESIS 1: The proportion of non-respondents increases with the number of waves participated in. In line with previous literature we expect to observe a higher proportion of non-buyers for second and third-time participants compared to first-time participants (Shields & To, 2005). This is the result of participants that change their respondents behavior when being interviewed multiple times.

HYPOTHESIS 2: There is positive correlation between underreporting and the number of waves participated in. We expect a lower number of reported payments (withdrawals) for second and third-time participants compared to first-time participants. In addition, we expect third-time participants to report even fewer payments or withdrawals than second-time participants.

HYPOTHESIS 3: Respondents underreport the number of point-of-sale payments in large cash-intensive sectors. We expect that respondents underreport the number of payments in sectors which up to the year of our study (2013) were characterized for being high volume cash-intensive sectors. These are: supermarkets, restaurants, vending machines, non-food small payments and food and tobacco shops. In particular, we expect the number of cash payments to be underreported.

HYPOTHESIS 4: The average value of payments is higher for respondents who participated in the survey before. We expect that respondents are more likely to underreport the number of small payments, resulting in a higher average value per payment. This hypothesis is based on the literature that small expenditures, in particular cash expenditures, are most frequently omitted when reporting in an

expenditure diary (Alessie et al., 1990).

HYPOTHESIS 5: The bias due to panel conditioning is concentrated in the following demographic subgroups: youths (aged 12-24), elderly people (aged 65+), lower educated people and respondents with lower income. These subgroups are chosen, because they are known as the most cash intensive groups. Also, these demographic subgroups make a relatively high number of small payments and are therefore more prone to forget payments (Alessie et al., 1990).

HYPOTHESIS 6: Responding behavior is similar in all months. We expect to find a consistent pattern for underreporting between November and March: the differences between second time and first-time participants in November and March are comparable (same sign and both differences significant or both differences insignificant).

4 Data and methodology

The dataset that we use in this paper corresponds to an experimental phase in the ongoing survey on consumers' payments commissioned by DNB/DPA. The survey gathers transaction level data as well as information about cash withdrawals. It was carried out in three waves in order to allow for comparisons between respondents' single or multiple participation in the study. A total of 21,954 respondents participated, spread during the months of September and November 2013, and March 2014. In each of these months the participants were evenly distributed over the days of the week, including weekends and holidays. The survey consisted of two parts: a payment diary and a questionnaire. The first gathered information about the characteristics of each payment or cash withdrawal carried out during the diary day. For each payment, respondents had to record the instrument (cash, debit card or other), the place of purchase (18 branches, e.g. a supermarket, street vendor, non-food retail shop, charity), the value of the purchase, and whether they were able to use their preferred payment method. For each cash withdrawal they had to report the source (ATM, bank counter, other) and the value of the withdrawal had to be recorded. The second part of the survey consisted of a questionnaire gathering background information on respondents' demographic characteristics as well as questions on issues that may influence their payment behavior, for example, payment preferences, reasons to use cash or cards, and other.

Not all respondents that took the survey in the first wave participated in the second and third wave (see Figure 1), therefore a refreshment group was introduced in both the second and the third wave. In the first wave, in September 2013, 8,707 consumers participated in the survey, divided over two groups: Group A consisting of 8,537 first-time participants and Group D, which consists of 170 consumers that participated in September and March, but not in November. We do not know what effect the gap of six months has on the reporting behavior of the respondents, and we therefore decided to exclude this group from the research. In November 2013 there were a total of 14,939 respondents, divided over 2 groups: Group A1 consisting of the consumers that participated in both the first and the second wave ($n=6,930$, response rate of 81.2%), and Group B, which is a refreshment group (first time participating) of 8,009 consumers. Then in the last wave, in March 2014, there was a total of 16,652 participants. These can be broken down in three groups: Group A2 consisting of consumers that participated in all three waves ($n=5,719$, response rate of 82.5%), group B1 consisting of the consumers that participated in wave 2 and 3 ($n=5,355$, response rate of 66.9%), and finally, group C, which is a refreshment group of 5,408 consumers.

4.1 Methodology

In this paper we compare average number and value of payments and withdrawals, and the proportion of non-buyers for different groups of participants in the survey in the same

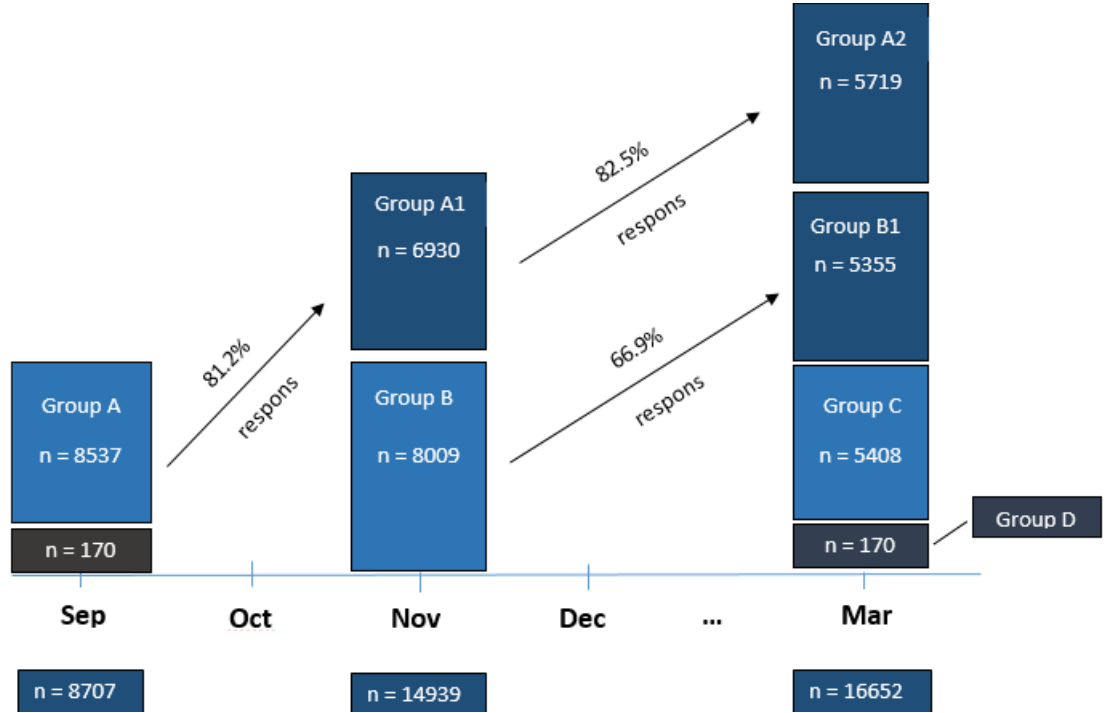


Figure 1: Overview of all participants and when they participated. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively. Group D is omitted from the research because of a six month gap between responses.

month³. For each of those variables there are four comparisons that are of interest:

1. First-time participants versus second-time participants in November (B and A1)
2. First-time participants versus second-time participants in March (C and B1)
3. First-time participants versus third-time participants in March (C and A2)
4. Second-time participants versus third-time participants in March (B1 and A2)

To estimate the panel conditioning effect we use a method proposed by Yan and Copeland (2010) in their paper on panel conditioning in the American expenditures diary book survey. In that paper Yan and Copeland argue that the first difference of the mean of the self reported answers in group j and k (where group j and k are in the same month) is equal to the measurement error and can be used as a proxy for the panel conditioning effect. This can be modeled as follows:

$$\bar{y}_j - \bar{y}_k = \gamma_{jk}, \quad (1)$$

where \bar{y}_i is the average value of a variable of interest in group i , and γ_{jk} the panel conditioning effect between groups j and k . Since we only compare groups that participate during the same month there is no time effect that could possibly influence the differences in reported answers. A two sample t -test on the difference of means is then used to test if the obtained panel conditioning effect is significantly different from zero. Finally, robust standard errors are used to account for heterogeneity between the groups.

³In this paper we focus on low-value payments (up to EUR 100) corresponding to the 95th percentile of the whole sample.

5 Payments, Withdrawals and Non-buyers

In this section, we look at payments, withdrawals and the proportion of non-buyers to see if there are any indications of panel conditioning with respect to these categories or when specific demographic groups are targeted. All variables are studied per group and demography, however, only the most relevant observations are presented in the paper⁴.

5.1 Non-buyers

The total number of non-buyers per day, month or group is defined as the sum of participants that made no actual purchases and participants that did not record any of their actual purchases that day. The percentage of non-buyers is known to be fairly constant for a country and was determined at 30%⁵ for the (same) Dutch population in 2013. In 2014 (Figure 2) we find that, using the new survey design implemented as of 2014, the percentage of non-buyers is slightly higher than 31%.

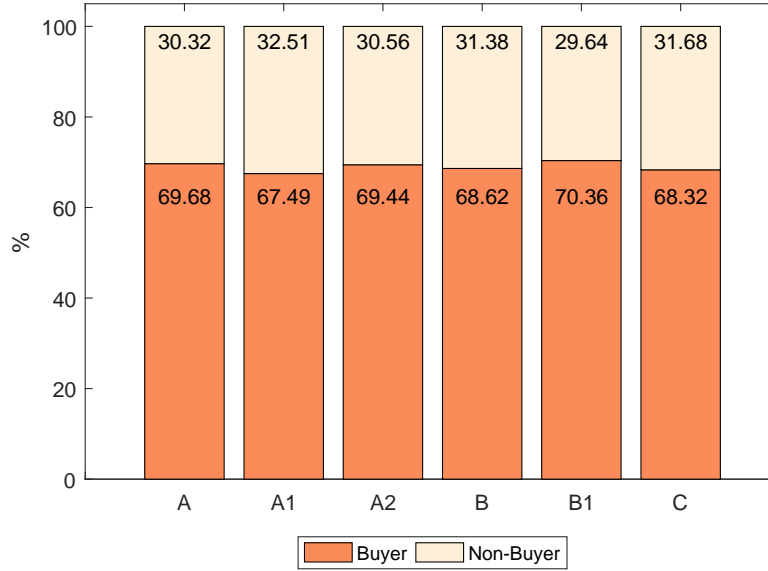


Figure 2: Overview of the percentage of buyers and non-buyers per group. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

The percentage of respondents who did not report any purchases (non-buyers) varies between 29.6% and 32.5% (Figure 2). This small gap ascertains that the actual percentage of non-buyers should be close to 31% and that there is no strong evidence of a panel conditioning effect for non-buyers. Furthermore, the percentage of non-buyers increases with time. For first-time participants this increase is from 30.3% in September to 31.4% in November and 31.7% in March). Note, that second-time participants in March (Group B1) have the lowest percentage of non-buyers (29.6%). The percentage of non-buyers is almost 1% lower than in all other groups and could therefore negatively affect the results on the average number of payments in Group B1, since a higher (lower) number of non-buyers artificially decreases (increases) the average number of payments in a group.

A priori, we expected that the distribution of non-buyers over age categories was skewed to the right and that the panel conditioning effect could be present for non-buyers

⁴The tables and figures regarding payments, withdrawals, and non-buyers per demographic are available from the authors upon request

⁵This number was obtained by the the SCP survey results from 2013, a cross section survey, where all participants only participated once

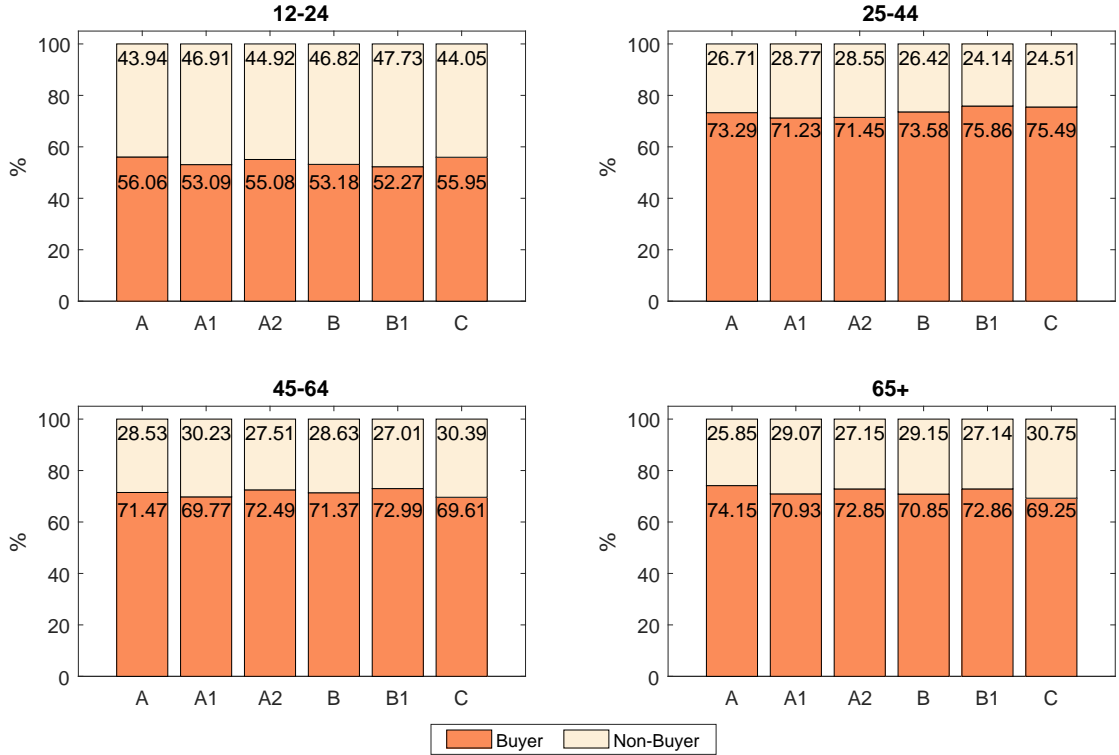


Figure 3: Overview of the percentage of buyers and non-buyers per group per age category. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

in some age categories. In this experiment, we indeed find that the percentage of non-buyers for ages below 25 is equal to 45.5%, and lower for the other age categories (Figure 3). We also observe differences in percentages between groups in the same month within specific age categories. For example, if we compare first, second and third-time participants in March (Group C, B1 and A2, respectively) for age category ‘45-64’ we find a 3% gap between groups that participated more than once (B1 and A2) and the first-time participants (C). So, overall there are no large differences in percentages of non-buyers, which means that the panel conditioning effect due to non-buyers in general is almost non-existent, but, if we look more closely, in specific demographic groups for example, we may find significant differences. The proportion of non-buyers over all demographic groups is further discussed in the results section.

5.2 Payments

In this section we look for potential differences in reported number and value of payments between the different groups within a month. We focus on the payments carried out at the point-of-sale (POS)⁶, which are defined as all payments below EUR 100 excluding person-to-person transactions (P2P), donations, and money that is stored at a bank or at home. We omit the higher values because they act as noise in the comparison of average value of payments. This is justified, because the number of payments with a value higher than EUR 100 is small (EUR 100 is the 98th percentile).

⁶Point-of-sale payments refers to purchases done in shops, restaurants, petrol stations, vending machines, market stalls, service-providers, among others.

5.2.1 Number and value of payments

Table 1 presents an overview of the number of payments per participating group. Groups which are compared to each other are shown together. In total, approximately 57,000 payments are made amongst 40,000 respondents, which leads to an average number of payments of 1.43 per person. The average number of payments varies within the groups, with the lowest average number of payments observed for third-time participants in March (Group A2) and the highest average number of payments for first-time participants in November (Group B).

Table 1: Comparison of number of point-of-sale payments below EUR 100 per group

Payments	Total			Cash		Debit Card	
	Respondents	Number	Avg. Number	Number	Avg. Number	Number	Avg. Number
A1	6,930	9,873	1.4247	4,708	0.6794	4,297	0.6201
B	8,009	11,927	1.4892	5,787	0.7226	5,142	0.6420
B1	5,355	7,475	1.3959	3,534	0.6599	3,313	0.6187
C	5,408	7,445	1.3767	3,654	0.6757	3,253	0.6015
A2	5,719	7,838	1.3705	3,813	0.6667	3,436	0.6008
C	5,408	7,445	1.3767	3,654	0.6757	3,253	0.6015
A2	5,719	7,838	1.3705	3,813	0.6667	3,436	0.6008
B1	5,355	7,475	1.3959	3,534	0.6599	3,313	0.6187
Total	39,958	57,292	1.4338	28,180	0.7052	24,470	0.6124

Note: This table reports the total number and average number of point-of-sale payments below EUR 100 for all payment instruments, cash, and debit card in the different groups. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

When we compare the two groups that participated in the second wave (November) we notice that second-time participants (Group A1) report a lower average number of payments than first-time participants (Group B). A lower number of reported payments does not necessarily mean that there is underreporting for each payment instrument. Since our paper mainly focuses on cash payments, we broke down payments by cash and debit card to more closely examine reporting behavior when using these two instruments. When comparing the different groups we notice that purchases using both, cash and the debit card, are underreported in November for second-time participants.

When comparing the groups participating in the third wave (March), we assess whether the findings are in line with those coming from the second wave, and notice that on average the number of reported payments for second-time participants (Group B1) is slightly higher than for first-time participants (Group C).

The average number of cash payments is lower for second-time participants, just as we saw in November. However, the average number of debit card payments is actually higher for second-time participants. This may seem strange, since this is contrary to the findings in November when we found that the average number of reported payments was lower for second-time participants, for both cash and debit card payments. This rather strange finding could be explained by the percentage of non-buyers in each group – the percentage of non-buyers is substantially lower for Group B1 as for Group C, with respectively 29.6% and 31.7% of non-buyers per group. So group C has a relatively high number of non-buyers (compared to Group B1), which brings down the number of average payments, since there is a relatively large group that makes zero payments. Therefore, it is not surprising that the average number of debit card payments is lower for Group C than for Group B1. For third-time participants (Group A2) the average number of payments is lower than for respondents who participated for the first time. However, note that the differences are small. Taking a more detailed look, we notice that both cash and debit card payments are lower for third-time participants.

To examine if there is any difference in reporting behavior between third and second-time participants, we also compared Group A2 with Group B1. As shown in

Table 1, respondents who participated for the third time reported fewer payments per person than second-time participants. There seems to be a relation between the number of payments per person and the number of times a respondent participated in previous surveys.

Overall, the average number of payments is lower for respondents who participated more than once. This suggests that there might be underreporting for the number of reported payments. However, we have to test the differences on statistical significance.

Table 2: Comparison of value of point-of-sale payments below €100 per group

Payments	Respondents	Total		Cash		Debit Card	
		Value	Avg. Value	Value	Avg. Value	Value	Avg. Value
A1	6,930	158,952	16.0997	49,637	10.5431	100,431	23.3722
B	8,009	192,365	16.1285	63,351	10.9472	117,712	22.8922
B1	5,355	117,598	15.7322	35,245	9.9730	76,088	22.9665
C	5,408	122,756	16.4883	38,964	10.6635	77,709	23.8883
A2	5,719	125,291	15.9850	39,585	10.3816	79,826	23.2323
C	5,408	122,756	16.4883	38,964	10.6635	77,709	23.8883
A2	5,719	125,291	15.9850	39,585	10.3816	79,826	23.2323
B1	5,355	117,598	15.7322	35,245	9.9730	76,088	22.9665
Total	39,958	920,033	16.0587	299,422	10.6253	571,138	23.3403

Note: This table reports the total value and average value of point-of-sale payments below EUR 100 for all payment instruments, cash, and debit card in the different groups. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

We also look at the value of the reported payments and the results are presented in Table 2. On average, respondents reported a value of EUR 16.06 per POS payment. The preference to use cash for small valued payments and a debit card, when a major purchase is made, is reflected in the higher average value for debit card payments, as well as in the distributions of the transaction sizes for cash and debit card payments below EUR 100 (Figure 4a and 4b).

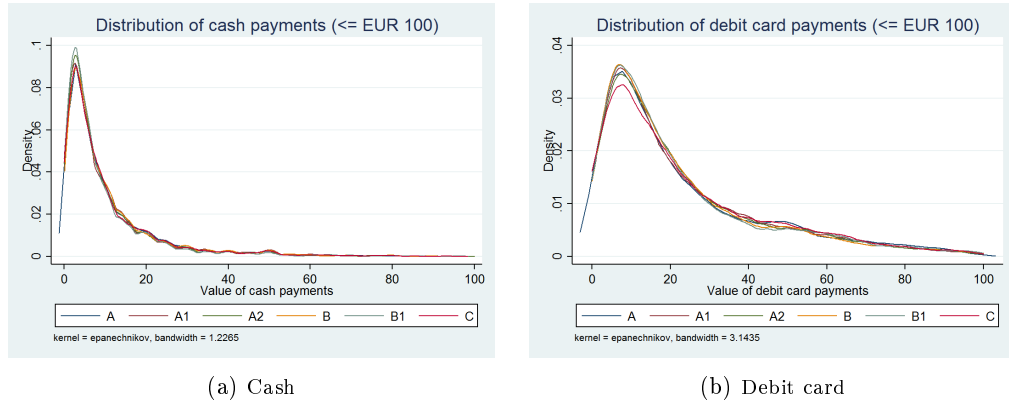


Figure 4: These figures display the distribution of point-of-sale (a) cash and (b) debit card payments below €100. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

When comparing the different groups, we find that respondents who participated more than once tend to underreport the value of their payments. This is due to the increase in proportion of small payments reported in the survey for second and third-time participants. Particularly, the difference between first and second-time participants in March seems large (Group B1 and C), where a decline of almost 5% is observed. What

is interesting is the difference between third and second-time participants (Group A2 and B1): second-time participants seem to be more prone to underreport the value of their payments than third-time participants.

5.2.2 Sectors

We extended our research and divided the payments by sector. The proportion of payments made in each sector and the frequency of each payment instrument per sector are provided in Figures 5a and 5b. Most payments are made in supermarkets and restaurants and similar, respectively 33.2 % and 13.6 %. We focus our analysis on the sectors that up to the year of the study (2013) were characterized for being high volume cash-intensive sectors, namely: supermarket, restaurant and similar, non-food retail (low value), CTN (confectionery, tobacconist, newsagent) and vending machine. This is in line with the 2013 findings on payments at the point-of-sale (Hernandez, Jonker, & Zwaan, 2014).

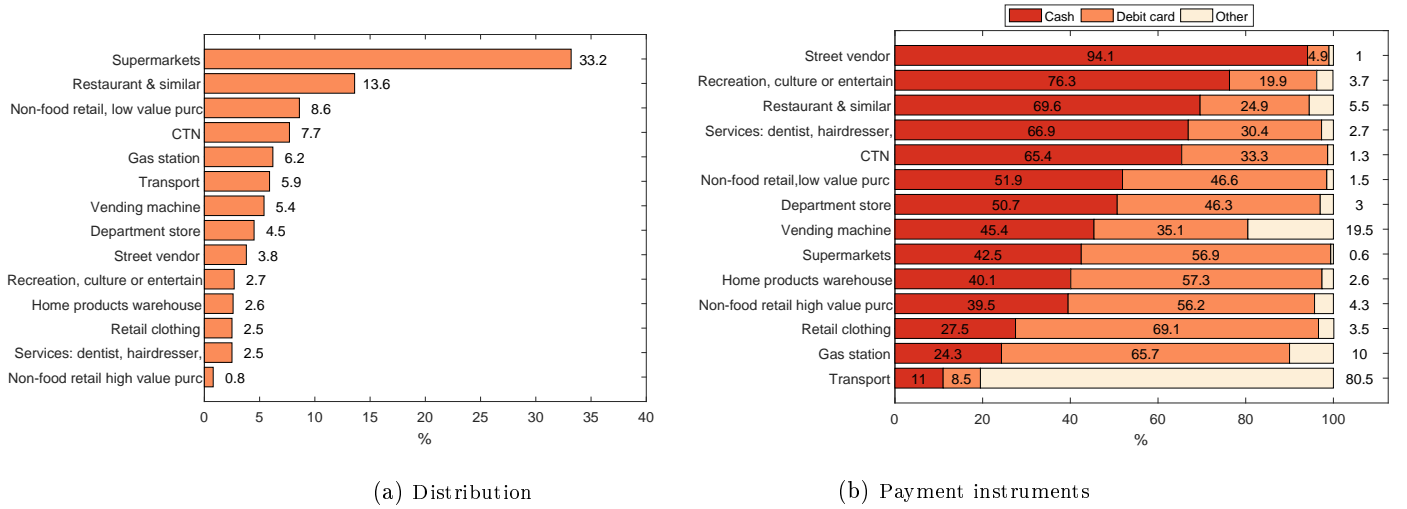


Figure 5: These figures display the distribution of low-value payments (up to EUR 100) (a) the number of point-of-sale payments over the different sectors and (b) the distribution by instrument of payment (cash, debit card, and other) in each sector.

5.2.3 Payments per demographic characteristics

Our fourth hypothesis focuses on the payments by participants in specific demographic groups. The hypothesis states that bias, due to panel conditioning, is concentrated in specific demographic groups. Therefore, we take a closer look at the payments per demographic group.

The average number and value of payments is calculated for each age category and each group individually. Furthermore, we divided the total payments in cash and debit card payments to get a more detailed look on the reported payments (Figure 6 and Figure 7)⁷. We notice that the mean number as well as the mean value is substantially lower for the youngest age category (age 12-24). What are noteworthy are the different patterns in number of payments for both payments instruments. The number of cash payments increases linearly with age, whereas debit card payments peak in age category ‘25-44’ and decrease thereafter.

When looking in more detail at debit card payments we note that for both months the mean number of reported payments by respondents aged 25 to 44 years old vary widely over the different groups (Figure 7). This suggests that there might be underreporting for the number of debit card payments for young adults (age 25-44) that have participated in the survey previously. The differences in reported payments for the other age categories

⁷The mean value for each age category is presented in the in Figure 8 and Figure 9

are in general small (Figure 6). While the average number of debit card payments seems roughly similar over all age categories, the average number of cash payments seems to differ over the different groups of participants. Large differences are observed for: respondents aged between 45 and 64 in November, youths in March and the oldest age category in March. The significance of the differences will be assessed in the results section.

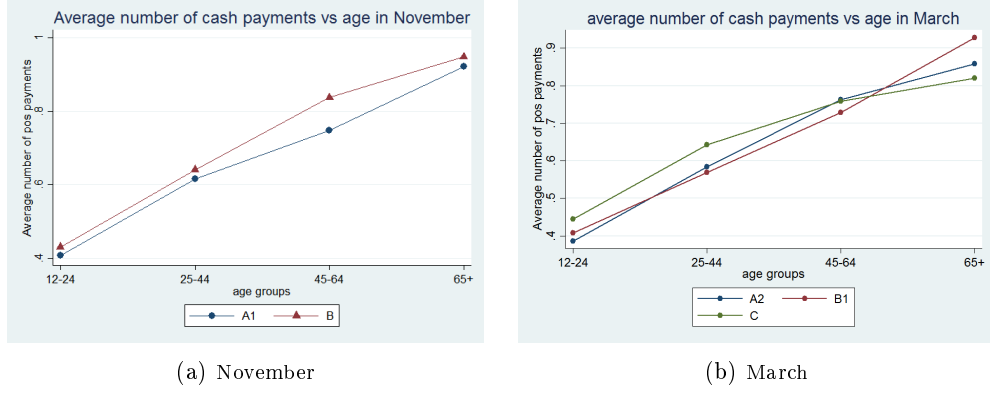


Figure 6: Average number of cash point-of-sale payments in November and March per age category. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

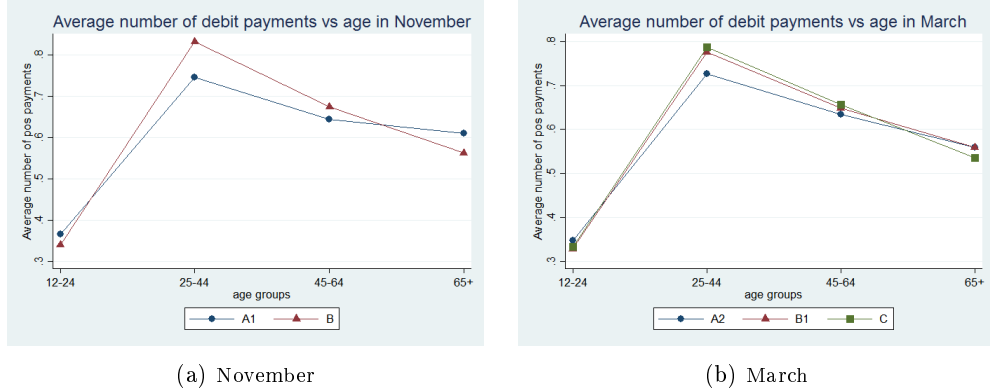


Figure 7: Average number of debit card point-of-sale payments in November and March per age category. Group A are first-time participants in September, Group B and A1 are first- and second-time participants in November, respectively, and Group C, B1, and A2 are first-, second-, and third-time participants in March, respectively.

The findings for the average value per payments over the age categories are remarkable. We expect to find a consistent pattern of reporting behavior over the months and a higher average value per payment for second and third-time participants, however, we observe almost the opposite.

For the average value of debit card payments, it is clear that the reporting behavior is inconsistent over the months. The average value of a debit card payment in November is in general higher for second-time participants compared to first-time participants, while the average value of debit card payment in March is lower for trained participants compared to first-time participants. Contrarily, the reported behavior towards the average value of a cash payment in both months show similarities. The average value of cash payment is lower for trained participants in both November and March. However, the size of the differences between fresh and trained participants varies between both months. In November we observe small differences in reporting behavior; in contrast, in March we observe large differences that are not consistent in sign. Therefore, a

consistent pattern for reporting behavior over the months is not observed. Furthermore, the hypothesis that the average transaction size is higher for second and third-time participants compared to first-time participants does not seem to hold.

We also look at the other demographic subgroups, in particular the other high cash intensive groups, such as lower educated people and respondents with a lower income, to see if there are signs of panel conditioning.

First, we look at the number of payments. Education seems unrelated to the average number of debit card payments. However, for the average number of cash payments we do find large differences between trained and first-time participants in educational subgroups. Yet, the differences are inconsistent over the months. For income, the bias due to panel conditioning seems to be concentrated in the second and third income category (EUR 14,300-38,300 and EUR 38,300-65,000). For district, the average number of cash payments is lower for second and third-time participants compared to first-time participants. In contrast, the pattern in the average number of reported debit card payments per district looks odd and illogical (Figure 10). For the demographics gender and ethnicity we notice that, even though the differences are small, trained respondents reported in general a lower average number of payments.

Second, we look at the value of payments. The analysis at the district level shows an inconsistent pattern of respondents' behavior towards reported value of payments for both debit and cash. Also for ethnicity, we do not see a clear relation between the reported value of payments and the ethnicity of the respondents. For gender however, we do notice consistent differences between trained and first-time participants for the average value of payments. We find that on average women tend to report a lower value of payments when being interviewed multiple times, in particular when reporting cash payments. Males however, are more likely to report a higher value of cash payment when having participated in the survey before. When examining the transaction size per educational level we see differences in responding behavior in March, but not in November. Finally, respondents' behavior regarding their reported values of transactions when paying by cash or debit card lacks consistency for the individual income and age groups.

5.3 Withdrawals

Another possible source of panel conditioning are cash withdrawals – where cash withdrawals are defined as withdrawals from an ATM, the bank counter (up to EUR 200)⁸, cash backs, cash that is found, cash gifts from family, friends, and others. Since the average person makes far fewer withdrawals than payments per day (in our dataset the average number of withdrawals is below 0.2), there are logically fewer observations on withdrawals than on payments. However, due to the size of the survey there are still more than 6,500 observations to examine.

⁸We omitted the higher values because they acted as noise in the comparison of average value of withdrawals. The threshold is set at the 95th percentile of withdrawal value, which is approximately €200.

Table 3: Comparison of number and value of withdrawals per group

Withdrawals	Number	Avg.	Value	Avg.
A1	1,154	0.167	51,220	44.38
B	1,400	0.175	62,070	44.34
B1	781	0.146	31,776	40.69
C	916	0.169	38,404	41.93
A2	837	0.146	36,050	43.07
C	916	0.169	38,404	41.93
A2	837	0.146	36,050	43.07
B1	781	0.146	31,776	40.69
Total	6,666	0.167	291,211	43.69

Note: This table reports the total and average number and value of withdrawals in the different groups. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

There is a clear decline in average number of withdrawals from groups that participated more than once compared to those who participated only once. For example, the second-time participants in November have an average number of 0.167 compared to an average number of 0.175 for the group of first-time participants, which is a 5% decrease of withdrawals per person per day. Furthermore a difference in percentage of zero withdrawals is observed between fresh and trained participants in November and March — the trained participants were 1.5% more likely to report zero withdrawals than the participants in the refreshment groups. The above indicates that panel conditioning affects the reporting of withdrawals in the experiment. Contrary to the findings on the average number of withdrawals, the average values of withdrawals do not differ in a systematic way. In November we observe no difference between first- and second-time participants (B and A1) and in March we observe, for second and third-time participants (B1 and A2), respectively higher and lower average values compared to the average value of first-time participants.

To investigate whether the above findings on average number and value of withdrawals are concentrated in specific demographic groups we also examined the withdrawals per demographic group. In November, we observe for all age categories, districts, education levels, ethnicity groups, genders and income groups, except for district ‘East’ and education level ‘HBO’(college education), that the average number of withdrawals recorded by second-time participants is lower than the recorded average number by first-time participants. In March, the second and third-time participants even reported a lower average number of withdrawals over all demographic subgroups. Looking at the average value of withdrawals, we find almost no effect of age, education level and income class on the differences between average value of withdrawals between groups with different number of participations in both November and March.

6 Results

In this section we provide and discuss the results of non-buyers, payments, transaction size and withdrawals. We have tested all demographic groups and present the most important observations.

6.1 Non-buyers

We hypothesized that respondents that participated more than once are more likely to report not to have made any payments than respondents who participated less or only once. Not rejecting this hypothesis would mean that the survey suffers from panel

conditioning and that the sample no longer suffices as a representative sample of the target population. To find statistical evidence in favor of or against a relative increase of non-buyers in the diary book survey we conducted a two sample proportion z -test on the difference of the proportion non-buyers between groups. The results of the tests are presented in Table 4.

Table 4: Significance test on the difference in the proportion of non-buyers between groups

Non-buyers	
A1 vs B	1.4825 (0.1382)
B1 vs C	-2.2944** (0.0218)
A2 vs C	-1.2648 (0.2060)
A2 vs B1	1.0648 (0.2869)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of proportion of non-buyers between the groups. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

The difference in proportion of non-buyers is not significantly different from zero for second and first-time participants in November. Thus, respondents that had a two-month gap between their first and second participation are not more likely to report zero payments than respondents in a refreshment group. This is in contrast to respondents that had a four-month gap between their first and second participation, who are significantly less likely to report not having made any payments than their first time counterparts. Third-time participants are not significantly more likely to report zero payments in March compared to first-time participants. Lastly, there is no significant difference in the proportion non-buyers between third and second-time participants in March. These findings are not in line with the first hypothesis posited earlier in the paper.

6.2 Payments

6.2.1 Number and value of payments

We examined whether there is a difference within reported number and value of payments for first-, second- or third-time participants, the results are shown in Table 5. The results show that only first- and second-time participants in November differ significantly in average number of reported payments (group A1 vs B). The sign of the t -statistic is negative, which reveals that second-time participants tend to report a lower number of payments, on average. Specifically, the average number of purchases made in cash seems to be lower for first-time participants. Unlike the number of payments, there is no significant difference in the average value of payments between first- and second-time participants in November.

Contrary to the findings from the second wave (November), we find a significant difference in average value of payments but no significant difference in number of payments for first- and second-time participants in the third wave (March). The negative t -statistic indicates that on average ‘fresh’ respondents report a higher value per payment than respondents that already participated in the survey. Also third-time participants

Table 5: Significance test on the difference of means for number and value of total, cash, and debit card point-of-sale payments between the groups

Payments	Total		Cash		Debit Card	
	Number	Value	Number	Value	Number	Value
A1 vs B	-2.2715** (0.0231)	-0.1141 (0.9092)	-2.2592** (0.0239)	-1.5580 (0.1193)	-1.2354 (0.2167)	1.1021 (0.2705)
B1 vs C	0.6295 (0.5290)	-2.4575** (0.0140)	-0.7612 (0.4466)	-2.2397** (0.0251)	0.8599 (0.3899)	-1.7421* (0.0815)
A2 vs C	-0.2002 (0.8413)	-1.6635* (0.0962)	-0.4306 (0.6667)	-0.9237 (0.3557)	-0.036 (0.9713)	-1.2622 (0.2069)
A2 vs B1	-0.8422 (0.3997)	0.8430 (0.3993)	0.3315 (0.7403)	1.3441 (0.1790)	-0.9192 (0.3580)	0.5175 (0.6048)

Note: This table reports t -statistics obtained by a two sample t -test on the difference of means for number and value of total, cash, and debit card point-of-sale payments between the groups. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first- and second-time participants in November, respectively, and Group C, B1, and A2 are first-, second-, and third-time participants in March, respectively.

tend to report a lower value per payments, although it should be noticed that this difference in reporting behavior is only significant at the 10 % significance level.

Furthermore, we find no significant differences between second-time and third-time participants (group A2 vs B1) for number as well as for value. This suggests that third-time participants are not more likely to underreport the number and value of their payments than respondents who only participated once before.

6.2.2 Transaction size

Our fourth hypothesis suggested that small payments and in particular small purchases paid in cash are most prone to underreporting, since those small payments are in general most frequently omitted when reporting all payments. However, we find results that are not in line with this hypothesis (Table 6). In fact, the opposite is true — higher frequency of small valued payments ($< \text{EUR } 5$) is observed for second-time participants compared to first-time participants in November as well as for March. This also holds for third-time participants. However, it should be noted that the differences are not significant at the traditional 5% significance level.

Of interest are the number of significant differences between the comparable groups. Only a small number of differences are significant, suggesting that there is no correlation between the transaction size and the number of waves participated in. This means that respondents who have participated in the survey before are not more likely to report a certain transaction size than fresh respondents.

Table 6: Significance test on the difference of means for the percentage of payments for a range of transaction sizes.

Transaction size	< EUR 5	EUR 5–15	EUR 15–25	EUR 25–50	EUR 50–100
A1 vs B	1.7946* (0.0727)	-1.2617 (0.2070)	-1.8747* (0.0608)	0.7881 (0.4307)	0.4574 (0.6474)
B1 vs C	0.6849 (0.4934)	2.3275** (0.0199)	0.0155 (0.9876)	-3.0288** (0.0025)	-1.4634 (0.1434)
A2 vs C	0.3348 (0.7378)	1.5099 (0.1311)	-0.2415 (0.8092)	-0.5562 (0.5781)	-2.1922** (0.0284)
A2 vs B1	-0.3585 (0.7200)	-0.8459 (0.3976)	-0.2575 (0.7968)	2.5100** (0.0121)	-0.7117 (0.4767)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of means for the percentage of payments for a range of transaction sizes between the groups. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

6.2.3 Sectors

Table 7: Significance test on the difference of the proportion of the number of total, cash and debit card payments in cash intensive sectors between groups

Sectors	Total	Cash	Debit Card
A1 vs B	-2.7283** (0.0064)	-0.4596 (0.6458)	-2.0522** (0.0402)
B1 vs C	-0.6848 (0.4934)	-1.3784 (0.1681)	-0.6848 (0.4934)
A2 vs C	0.0663 (0.9471)	0.3074 (0.7585)	0.6515 (0.5148)
A2 vs B1	0.7601 (0.4472)	1.6979* (0.0895)	-0.9287 (0.3530)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of the proportion of the number of total, cash and debit card payments in cash intensive sectors between groups. The five large and cash intensive sectors are: supermarket, non-food small payments, food and entertainment, vending machine, and restaurant & similar. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

To further explore if cash (and other) payments are significantly underreported in this survey we focus on five large sectors in which up to 2013 the year of this study, the use of cash was mostly dominant: supermarket, restaurants and similar, non-food retail (low value), food and CTN and vending machines (Hernandez et al., 2014). These five sectors together cover about 68% of all point-of-sale payments in the survey. To examine if respondents did underreport the number of payments in these cash intensive sectors we compare the proportion of the number of payments in one of the five cash intensive groups versus all other point-of-sale payments. The results of the comparison of total, cash and debit card payments are given in Table 7. We find only two significant results at the 5% significance level: the total number of payments is underreported by second-time participants in November due to underreporting of their debit card payments. Thus we

found significant differences for debit card payments, but again, these effects are not consistent over the months. Moreover we do not find any significant differences for the number of cash payments, which indicates that there is no significant panel conditioning effect for cash payments regarding sectors.

We also analyzed the sectors at an individual level. Although we find some signs of change in the reporting of cash payments in November and March in supermarkets and for debit card payments in November in vending machines, these are again not consistent over the months. The significant difference for cash payments is positive in November, while it is negative in March and the difference for number of debit card payments is significant between trained and fresh respondents in November, whereas it is not significant in March.

6.2.4 Payments per demographic

Furthermore, we looked for the presence of panel conditioning in specific demographic groups (age, district, education, ethnicity, gender, and income level). Contrary to our expectations, we did not find evidence of a significant panel conditioning effect in any of the studied demographic groups. The results of the test outcomes on the number and value of cash and debit card payments for individual demographic groups are provided in Table 8.⁹ The results reveal that the reporting behavior of second and third-time participants is in general roughly similar to reporting behavior of first-time participants.

However, there are a few individual significant differences in subcategories of the demographic groups. For example, adults (age 45-65) that participate for the second time in November underreport the number of cash payments compared to first-time participants. Whereas, younger adults (age 25-44) that participate for the second time in March underreport the number of cash payments compared to first-time participants. These results could be an indication of a panel conditioning effect, however, they are not consistent over the other months and therefore we cannot interpret it as such. There is one example of a nearly consistent panel conditioning effect in a subcategory. Second-time participants in the EUR 38,300 - 65,000 income level group underreport the number of cash payments in both November and March compared to first-time participants, however, in November this effect is only significant at a significance level of 10%.

The results for the value of cash payments are in line with those found for the number of cash payments, some significant differences are found that are not consistent over time. There is one exception — females that participate for the second time in November and March underreport the average value of cash payments at a 5% significance level, indicating that females might be influenced by the panel conditioning effect. However, third-time female participants in March do not significantly underreport compared to second-time participants.

The results for number and value of debit card payments reveal that there is no consistent panel conditioning effect in any of the subcategories.

6.3 Withdrawals

Results of the tests on the number and value of withdrawals are displayed in Table 9. The negative test statistics indicate that trained respondents have the tendency to underreport the number of withdrawals, although it must be noted that, the panel conditioning effect is only significant in March. Once more we also find that groups that participated three times are not significantly more likely to underreport the number of withdrawals in March than groups that participated only twice. Besides that, we also tested on the difference of the proportion of zero withdrawals and at least one withdrawal per person per day, for first and second-time participants in November and March. In both these months the obtained test statistics are positive and significant,

⁹Test result of the unpaired two sample t -test of total payments and debit card payments are shown in the Appendix Table 13 and Table

Table 8: Results of the unpaired two sample t-test on the difference of means for number and value of point-of-sale cash payments between groups, for age, district, education level, ethnicity, gender, and income class.

Number	A1 vs B	B1 vs C	A2 vs C	A2 vs B1	Value	A1 vs B	B1 vs C	A2 vs C	A2 vs B1
Age									
12-24	-0.7174 (0.4732)	-1.0259 (0.3051)	-1.6020 (0.1093)	-0.5943 (0.5524)		-0.0545 (0.9566)	1.1640 (0.2448)	2.0269** (0.0431)	0.9007 (0.3681)
25-44	-0.7501 (0.4532)	-2.0983** (0.0360)	-1.6166 (0.1061)	0.4356 (0.6632)		-1.0225 (0.3066)	-0.0958 (0.9237)	-1.2903 (0.1971)	-1.1671 (0.2433)
45-65	-2.6022** (0.0093)	-0.7948 (0.4268)	0.0961 (0.9235)	0.9167 (0.3594)		-0.2436 (0.8075)	-3.1543** (0.0016)	-1.1510 (0.2498)	2.0280** (0.0427)
65+	-0.4876 (0.6259)	1.8815* (0.0601)	0.6956 (0.4868)	-1.2143 (0.2248)		-1.8687* (0.0618)	-2.1143** (0.0346)	-1.529 (0.1264)	0.6165 (0.5376)
District									
Big cities + aggl.	-0.9709 (0.3317)	-0.5598 (0.5757)	0.1800 (0.8571)	0.7577 (0.4488)		-1.5416 (0.1234)	-0.6968 (0.4861)	-0.4656 (0.6416)	0.2215 (0.8247)
West	-0.6165 (0.5376)	-0.4473 (0.6547)	0.3149 (0.7529)	0.7706 (0.4410)		-1.8400* (0.0659)	-1.5005 (0.1336)	-0.6907 (0.4898)	0.8671 (0.3860)
North	-1.4487 (0.1476)	-0.9156 (0.3601)	-0.9877 (0.3235)	-0.0733 (0.9415)		-0.3390 (0.7347)	-1.383 (0.1672)	0.0571 (0.9545)	1.382 (0.1674)
East	-0.4387 (0.6609)	-0.3198 (0.7491)	-0.2415 (0.8092)	0.0766 (0.9389)		2.2718** (0.0232)	-1.0237 (0.3061)	0.1841 (0.8539)	1.2289 (0.2193)
South	-1.6258 (0.1041)	0.2433 (0.8078)	-0.2904 (0.7715)	-0.5233 (0.6008)		-2.0451** (0.0409)	-0.5371 (0.5913)	-0.7897 (0.4298)	-0.2387 (0.8114)
Education									
Primary education	1.7613* (0.0785)	0.5144 (0.6071)	-0.5954 (0.5517)	-1.0083 (0.3137)		-0.1474 (0.8829)	0.6687 (0.5043)	0.7814 (0.4353)	0.1103 (0.9123)
Lower voc.	0.0746 (0.9406)	0.6029 (0.5466)	1.1173 (0.2640)	0.4934 (0.6218)		-1.5561 (0.1198)	-1.4794 (0.1392)	-0.5463 (0.5849)	1.0306 (0.3029)
Senior sec. voc.	-3.8931** (0.0001)	-0.3303 (0.7412)	-0.606 (0.5446)	-0.2773 (0.7815)		0.1881 (0.8508)	-0.755 (0.4503)	-1.0271 (0.3045)	-0.2562 (0.7978)
Higher professional	-0.3176 (0.7508)	-2.5022** (0.0124)	-1.7932* (0.0731)	0.6732 (0.5009)		-0.9347 (0.3500)	-1.7953* (0.0728)	-0.4921 (0.6227)	1.2903 (0.1971)
University	-1.6257 (0.1043)	0.3406 (0.7335)	0.1853 (0.8531)	-0.1440 (0.8856)		-0.5545 (0.5795)	-1.5341 (0.1256)	-0.4652 (0.6420)	0.9179 (0.3591)
Ethnicity									
Native	-1.5064 (0.1320)	-1.0538 (0.2920)	-0.4049 (0.6856)	0.6672 (0.5047)		-0.7806 (0.4351)	-2.2419** (0.0250)	-1.1086 (0.2676)	1.1932 (0.2328)
Foreign	-1.9100* (0.0563)	-0.1071 (0.9148)	-0.0428 (0.9659)	0.055 (0.9562)		-1.1678 (0.2431)	-0.6600 (0.5094)	0.2672 (0.7894)	0.9685 (0.3330)
Gender									
Male	-1.9936** (0.0462)	-0.8334 (0.4047)	0.1065 (0.9152)	0.9577 (0.3382)		0.2135 (0.8309)	0.3974 (0.6911)	0.2530 (0.8003)	-0.1551 (0.8768)
Female	-1.1950 (0.2321)	-0.2965 (0.7668)	-0.6303 (0.5285)	-0.3355 (0.7373)		-2.3831** (0.0172)	-3.5683** (0.0004)	-1.6086 (0.1078)	2.0102** (0.0445)
Income									
< EUR 14,300	-0.0710 (0.9434)	-0.4288 (0.6682)	-1.4985 (0.1344)	-0.9520 (0.3414)		-2.0841** (0.0375)	-0.4138 (0.6792)	-0.6314 (0.5281)	-0.1787 (0.8582)
EUR 14,300–38,300	-1.9480* (0.0515)	0.7850 (0.4325)	0.4666 (0.6408)	-0.3377 (0.7356)		-1.1358 (0.2561)	-0.7490 (0.4539)	-1.2375 (0.2160)	-0.4494 (0.6532)
EUR 38,300–65,000	-1.8596* (0.0630)	-2.0772** (0.0379)	-0.5600 (0.5755)	1.5720 (0.1161)		0.9705 (0.3319)	-0.1255 (0.9002)	0.3712 (0.7106)	0.4951 (0.6206)
> EUR 65,000	-1.6133 (0.1069)	-0.6336 (0.5265)	-0.8897 (0.3738)	-0.3031 (0.7618)		-0.4459 (0.6558)	-1.3742 (0.1698)	0.3606 (0.7185)	1.5935 (0.1115)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of means for number and value of point-of-sale cash payments between groups, for age, district, education level, ethnicity, gender, and income class. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

Table 9: Significance test on the difference of means for number and value of withdrawals between the groups

Withdrawals	Number	Value
A1 vs B	-1.2621 (0.2069)	0.0271 (0.9784)
B1 vs C	-3.1878** (0.0014)	-0.5899 (0.5553)
A2 vs C	-3.1234** (0.0018)	0.5530 (0.5804)
A2 vs B1	0.0737 (0.9413)	1.1257 (0.2605)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of means for number and value of withdrawals between the groups. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

2.4599 (0.0139) in November and 1.9929 (0.0463) in March, where the p -values are given in parenthesis. This confirms the observation from the previous section on withdrawals, that the percentage of zero reported withdrawals in a group is higher if the group has participated before. For the reported value of withdrawals there are no significant differences between the groups, which indicates that there is not one specific withdrawal size that is underreported.

To examine if the panel conditioning effect for the number of withdrawals is related to specific demographic groups, we looked at all demographic groups in more detail. Figures containing three demographic groups (age, gender, and income) are displayed in appendix A (Figure 11, 12, and 13). These figures reveal, as expected, that the trained respondents are, in general, underreporting their number of withdrawals per day.

We find that the average number of reported withdrawals is lower for trained respondents for all age groups, gender, income categories, ethnicity, districts and education levels but this panel conditioning effect is not similar in the second wave (November) and in the third wave (March). For example, in November only trained respondents, between 25 and 44 year of age, report a lower average number of withdrawals, whereas in March all trained respondents report a lower number of withdrawals, except for third-time participants older than 65. Also in November, trained males reported a lower number of withdrawals than first time participating males, while trained woman did not report so much less, contrary in March, the trained females reported relatively less withdrawals than the trained males. The only effect that carries over from one month to the other is in category ‘District’. For this we find a significantly lower average number of reported withdrawals, in both November and March, for trained respondents living in the West or East of the Netherlands, excluding the three biggest cities. The effect is significant at the traditional 5% significance level for trained respondents living in the East, whereas it is only significant at the 10% significance level for trained respondents living in the West (highest p -value is 0.0755, see Table 12).

7 Conclusion

In this paper we analyse data from the survey on consumers’ payments in order to detect panel conditioning and to see whether any panel conditioning effect is concentrated in specific demographic groups.

Overall, we conclude that there is no significant evidence of a panel conditioning

effect in the payment survey. This is a very positive finding as it suggests that reporting behavior does not change when participants are interviewed multiple times. Since the survey does not suffer from panel conditioning, we conclude that the survey is well constructed and that the survey results can be used to make inferences about the Dutch population, without taking into account the number of times a respondent participated in the survey. We now more closely examine each hypothesis individually.

Our first hypothesis considered the proportion of respondents who do not report any payments (non-buyers). We found no significant evidence of a panel conditioning effect due to non-buyers in any of the groups. Moreover, the trained respondents were less likely to report zero payments than first-time participants, albeit not significant at the traditional 5% significance level. Hence, we rejected the first hypothesis.

We suspected that overall, the average number of reported payments and withdrawals would decrease with the number of waves participated in. This was our second hypothesis. Our results reveal that underreporting of the number of payments and withdrawals for trained respondents appears in both months. Yet, only a few differences were significant. For payments only the difference between first and second-time participants in the second wave of the study (November) were statistically significant. For withdrawals we found the opposite: statistically significant differences were only observed in March between second and first-time participants and between third and first-time participants. Since the significant differences lack consistency over the months, we reject the second hypothesis. Furthermore, we found no significant differences between second time and third-time participants, suggesting that second time and third-time participants have similar reporting behavior.

In our third hypothesis, we expected the number cash payments, carried out in known cash-intensive sectors, to be underreported¹⁰. Although we did find some signs of change in the reporting of cash payments in supermarkets for the second wave (November) and third wave (March) of the study, the findings are not consistent over time and among the different groups of respondents. For this reason, we reject the third hypothesis.

In our fourth hypothesis, we suggested that the average value per payment would be higher for trained respondents compared to first-time participants. We found no significant evidence for this hypothesis when examining the responding behavior of withdrawals. When looking at payments, we found no supporting evidence for this hypothesis. In fact, second time and third-time participants tend to report a higher number of small payments. This is a sign that trained respondents better report their small payments. For this reason, it can be said that the new methodology leads to improved results on the use of cash, especially for low-value payments.

We extended the research and divided the respondents in subcategories based on their demographic characteristics and we hypothesized that the bias due to panel conditioning would be concentrated in specific demographic subgroups. In particular, we expected concentrated bias in the subgroups that are known for being cash-intensive users: children, elderly people, respondents with a lower income and lower educated people. The analysis per demographic subgroup did not show any consistent patterns and only few differences were statistically significant. As a result, we found no evidence to support our fifth hypothesis. Furthermore, we found again no significant evidence that second and third-time participants differ in reporting behavior.

The last hypothesis looked for the consistency of the panel conditioning effect in all months. We found that trained respondents were significantly more likely to report zero withdrawals than first-time participants in both November and March. For the other variables of interest, number and value of payments, value of withdrawals and proportion of non-buyers, we did not find significant panel conditioning effects in all months. Therefore we had to reject the hypothesis for all variables except for the proportion of non zero withdrawals for trained respondents.

¹⁰We focus our analysis on the sectors that, up to the year of the study, characterized as being cash-intensive sectors. These are: supermarkets, restaurants and similar, non-food retail shops (low value), TCN (confectionery, tobacconist, newsagent) and vending machines. These five sectors together cover about 68% of all point-of-sale payments in the survey (Hernandez et al., 2014).

Finally, we conclude that quantifying a panel conditioning effect in panel surveys on payments and withdrawals is still very difficult and that further research is required.

Possible future advancements can be made by, for example, conducting a longitudinal survey which is able to capture specific month effects — we suspect that there are different reporting patterns over the different months (panel conditioning behaves differently in one month compared to the other), this could be tested by conducting a survey that spans more than one year.

Another advancement can be made by revising the questionnaire structure. We initially expected to find possible changes on reporting behavior regarding consumers' payments rather than on reporting of their cash withdrawals. This is because the frequency of daily payments is higher than that of withdrawals, which could make the recording process more tedious. Contrary to our expectations, we find a clear panel conditioning effect for the number of cash withdrawals over all demographic groups, while we did not find consistent changes of behavior in the reporting of payments made by respondents. We think that this could be due to the structure of the questionnaire. In the questionnaire respondents first had to report their payments and after that their withdrawals. For this reason we propose a revision of the design of the questionnaire in combination with an improvement of the online tool which most respondents use to complete the survey.

References

- Alessie, R., Gradus, R., & Melenberg, B. (1990). The problem of not observing small expenditures in a consumer expenditure survey.
- Boeschoten, W. (1992). Currency use and payment patterns. *Kluwer Academic Publishers. Dordrecht*.
- Bradburn, N. (1992). A response to the nonresponse problem. *Public Opinion Quarterly*, 391–397.
- Cheak-Zamora, N. C., Wyrwich, K. W., & McBride, T. D. (2009). Reliability and validity of the sf-12v2 in the medical expenditure panel survey. *Quality of Life Research*, 18(6), pp. 727–735.
- Cohen, J. W., Monheit, A. C., Beauregard, K. M., Cohen, S. B., Lefkowitz, D. C., Potter, D., . . . III, R. H. A. (1996). The medical expenditure panel survey: A national health information resource. *Inquiry*, 33(4), pp. 373–389.
- Das, M., Toepoel, V., & van Soest, A. (2011). Nonparametric tests of panel conditioning and attrition bias in panel surveys. *Sociological Methods & Research*, 40(1), 32–56. doi: 10.1177/0049124110390765
- De Leeuw, E. D. (1997). Nonrespons: Wat weten we en wat kunnen we eraan doen. *Recente Ontwikkelingen in het Marktonderzoek*, 39–59.
- Golob, T. T., & Meurs, H. (1986). Biases in response over time in a seven-day travel diary. *Transportation*, 13(2), 163–181. doi: 10.1007/BF00165546
- Hernandez, L., Jonker, N., & Zwaan, P. (2014). Point-of-sale payments in 2013. http://www.dnb.nl/en/binaries/fs%20uk_tcm47-316904.pdf. (Accessed: 2017-01-30)
- Hsiao, C. (2007). Panel data analysis: advantages and challenges. *TEST*, 16(1), 1–22. doi: 10.1007/s11749-007-0046-x
- Kitamura, R. (1990). Panel analysis in transportation planning: An overview. *Transportation Research Part A: General*, 24(6), 401–415.
- Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). The people’s choice: How the voter makes up his mind in a presidential campaign. *The ANNALS of the American Academy of Political and Social Science*, 261(1), 194. doi: 10.1177/000271624926100137
- Meurs, H., Van Wissen, L., & Visser, J. (1989). Measurement biases in panel data. *Transportation*, 16(2), 175–194.
- Neter, J., & Waksberg, J. (1964). Conditioning effects from repeated household interviews. *The Journal of Marketing*, 51–56.
- Sharot, T. (1991). Attrition and rotation in panel surveys. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 40(3), pp. 325–331.
- Shields, J., & To, N. (2005). Learning to say no: conditioned underreporting in an expenditure survey.
- Silberstein, A. R., & Jacobs, C. A. (1989). Symptoms of repeated interview effects in the consumer expenditure interview survey. *Panel Surveys*, 289–303.
- Sturgis, P., Allum, N., & Brunton-Smith, I. (2009). Attitudes over time: The psychology of panel conditioning. *Methodology of longitudinal surveys*, 113–126.
- Verbrugge, L. M. (1980). Health diaries. *Medical care*, 18(1), 73–95.
- Yan, T., & Copeland, K. (2010). Panel conditioning in the consumer expenditure quarterly interview survey. *JSM Proceedings. Survey Research Methods Section. Alexandria, VA: American Statistical Association*, 3618–3628.

A Figures and Tables

Table 10: Significance test on the difference of means for the percentage of debit payments for a range of transaction sizes

Transaction size	< EUR 5	EUR 5–15	EUR 15–25	EUR 25–50	EUR 50–100
A1 vs B	0.1156 (0.9080)	-0.3069 (0.7589)	-2.2785** (0.0227)	1.6677* (0.0954)	1.0849 (0.2780)
B1 vs C	-1.9695** (0.0489)	3.6501** (0.0003)	0.8226 (0.4107)	-2.3203** (0.0203)	-1.1191 (0.2631)
A2 vs C	-1.5607 (0.1186)	2.1723** (0.0298)	0.8491 (0.3958)	-0.3597 (0.7191)	-1.7438* (0.0812)
A2 vs B1	0.4293 (0.6677)	-1.519 (0.1288)	0.0193 (0.9846)	1.9899** (0.0466)	-0.6171 (0.5372)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of means for the percentage of debit payments for a range of transaction sizes. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

Table 11: Significance test on the difference of means for the percentage of cash payments for a range of transaction sizes

Transaction size	< EUR 5	EUR 5–15	EUR 15–25	EUR 25–50	EUR 50–100
A1 vs B	2.1525** (0.0314)	-1.2214 (0.2219)	-0.2843 (0.7762)	-1.245 (0.2131)	-0.3546 (0.7229)
B1 vs C	1.9564* (0.0504)	0.0400 (0.9681)	-0.7196 (0.4718)	-2.3259** (0.0200)	-0.929 (0.3529)
A2 vs C	0.9063 (0.3648)	0.6275 (0.5303)	-1.2772 (0.2015)	-0.9608 (0.3367)	-0.7345 (0.4627)
A2 vs B1	-1.0783 (0.2809)	0.5817 (0.5608)	-0.5390 (0.5899)	1.4003 (0.1614)	0.2115 (0.8325)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of means for the percentage of cash payments for a range of transaction sizes. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

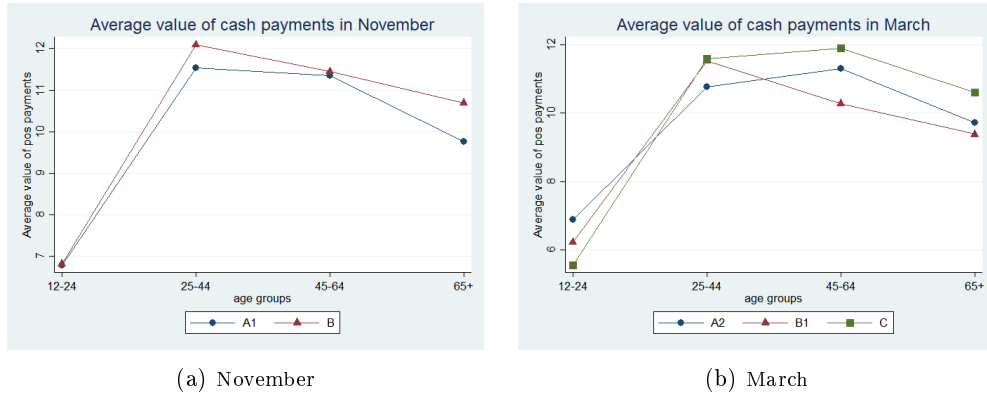


Figure 8: Average value of cash point-of-sale payments in November and March per age category. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

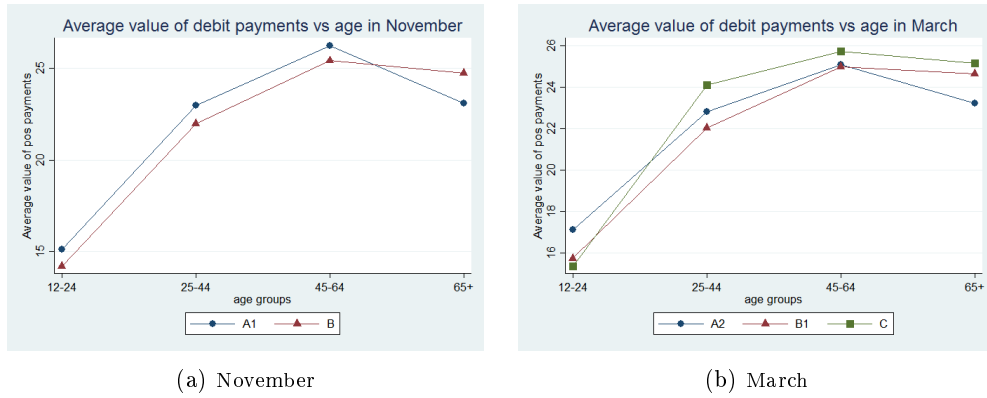


Figure 9: Average value of debit card point-of-sale payments in November and March per age category. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

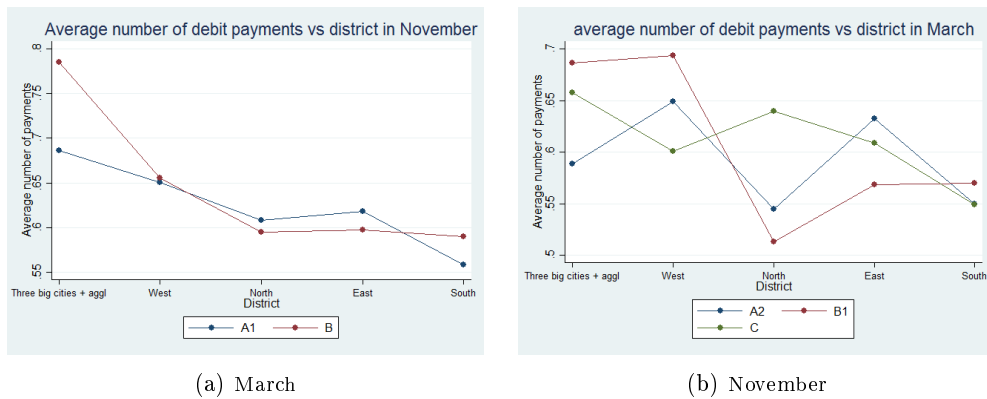


Figure 10: Average number of debit card point-of-sale payments in November and March per district. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

Table 12: Significance test on the difference of means for number of withdrawals between groups for age, district, education level, ethnicity, gender, and income class

Number	A1 vs B	B1 vs C	A2 vs C	A2 vs B1		A1 vs B	B1 vs C	A2 vs C	A2 vs B1
Age					Income				
12-24	0.0518 (0.9587)	-2.4517** (0.0143)	-2.7749** (0.0056)	-0.3502 (0.7262)	< EUR 14,300	-0.9038 (0.3663)	-1.3973 (0.1627)	-1.0304 (0.3031)	0.3961 (0.6922)
25-44	-1.8127* (0.0699)	-1.1311 (0.2581)	-1.8166* (0.0694)	-0.7632 (0.4454)	EUR 14,300-38,300	-1.5263 (0.127)	-1.998** (0.0458)	-1.5054 (0.1323)	0.5956 (0.5515)
45-65	-0.1977 (0.8433)	-1.3446 (0.1788)	-1.2442 (0.2135)	0.1252 (0.9004)	EUR 38,300-65,000	-0.1871 (0.8516)	-1.5274 (0.1268)	-1.6791* (0.0933)	-0.2005 (0.8411)
65+	-0.3295 (0.7418)	-1.5301 (0.1261)	0.8988 (0.1272)	1.6952* (0.0902)	> EUR 65,000	0.5308 (0.5308)	-1.0498 (0.2940)	-0.5548 (0.5791)	0.4960 (0.6200)
Ethnicity					Gender				
Native	-1.1936 (0.2327)	-3.0169** (0.0026)	-2.8512** (0.0044)	0.2493 (0.8032)	Male	-1.3868 (0.1656)	-1.6409 (0.1009)	-1.6052 (0.1085)	0.0541 (0.9568)
Foreign	-0.2100 (0.8337)	-1.3983 (0.1622)	-1.1979 (0.2311)	0.0654 (0.9479)	Female	-0.4138 (0.6790)	-2.8365** (0.0046)	-2.7578** (0.0058)	0.0659 (0.9475)
Education					District				
Primary education	-0.6129 (0.5400)	-1.7344* (0.0833)	-0.6462 (0.5183)	1.0533 (0.2926)	Big cities + aggl.	-0.6353 (0.5253)	0.1977 (0.8433)	-0.9157 (0.3600)	-1.1779 (0.2390)
Lower voc.	-1.275 (0.2024)	-1.6676* (0.0955)	-1.3889 (0.1650)	0.3154 (0.7525)	West	-1.7781* (0.0755)	-2.5693** (0.0102)	-1.8192* (0.0690)	0.7884 (0.4305)
Senior sec. voc.	-0.8337 (0.4045)	-0.3859 (0.6996)	-1.4035 (0.1606)	-1.0390 (0.2988)	North	-0.4626 (0.6437)	-0.4319 (0.6659)	-1.4757 (0.1403)	-1.0349 (0.3009)
Higher professional	1.022 (0.3069)	-2.0345** (0.0420)	-2.3639** (0.0182)	-0.3691 (0.7121)	East	2.1543** (0.0313)	-2.5400** (0.0112)	-2.2025** (0.0277)	0.3646 (0.7155)
University	-1.9745** (0.0486)	-1.5103 (0.1314)	-0.4911 (0.6235)	0.9796 (0.3276)	South	-1.8463* (0.0649)	-1.1437 (0.2529)	-0.6649 (0.5062)	0.4982 (0.6184)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of means for number of withdrawals between groups for age, district, education level, ethnicity, gender, and income class. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

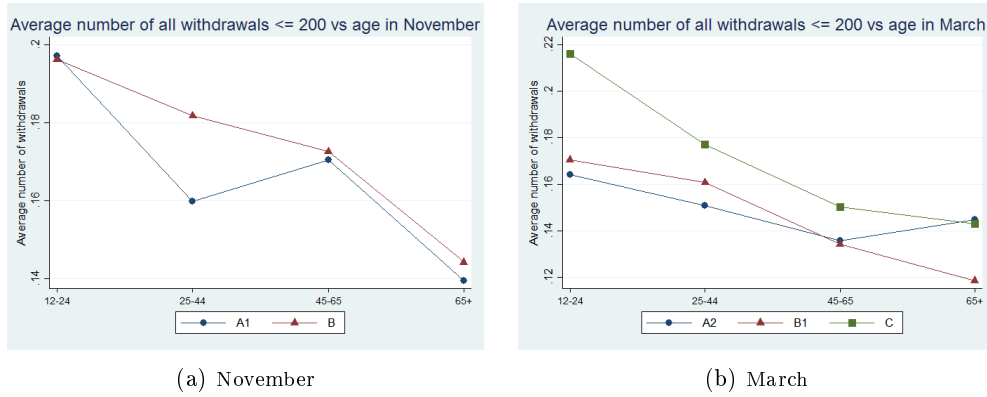
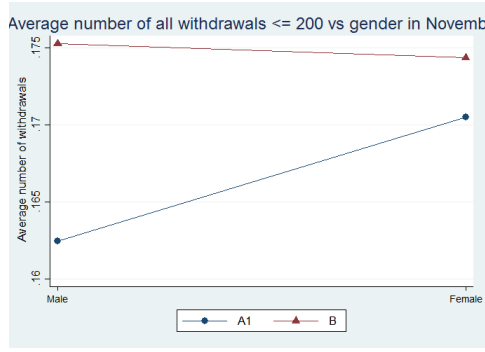
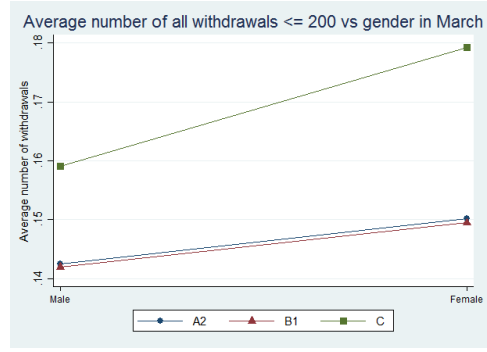


Figure 11: Average number of withdrawals per person per day in November and March per group and age category. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

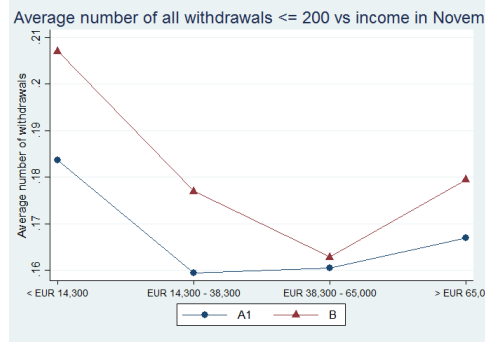


(a) November

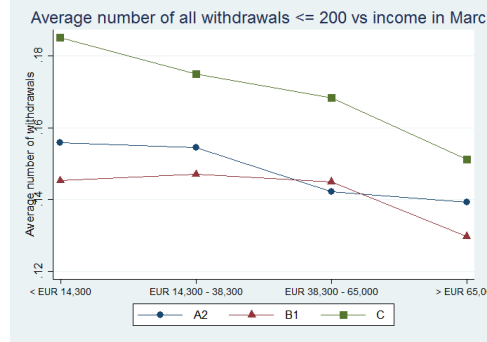


(b) March

Figure 12: Average number of withdrawals per person per day in November and March per group and gender. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.



(a) November



(b) March

Figure 13: Average number of withdrawals per person per day in November and March per group and income class. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

Table 13: Significance test on the difference of means for number and value of total point-of-sale payments between groups, for age, district, education level, ethnicity, gender, and income class

Number	A1 vs B	B1 vs C	A2 vs C	A2 vs B1	Value	A1 vs B	B1 vs C	A2 vs C	A2 vs B1
Age									
12-24	0.9988 (0.3180)	-0.5451 (0.5858)	0.0407 (0.9675)	0.5590 (0.5762)		0.4507 (0.6522)	0.2525 (0.8007)	1.2755 (0.2023)	0.9307 (0.3521)
25-44	-2.7416** (0.0061)	-1.1715 (0.2415)	-2.3346** (0.0196)	-1.3048 (0.1920)		0.1238 (0.9015)	-2.1084** (0.0351)	-2.0881** (0.0368)	0.0055 (0.9956)
45-64	-1.8918* (0.0586)	-0.5381 (0.5906)	-0.3278 (0.7431)	0.2199 (0.8260)		0.7053 (0.4806)	-1.8620* (0.0627)	-0.9693 (0.3324)	0.9649 (0.3346)
65+	-0.2600 (0.7949)	2.1059** (0.0353)	0.8259 (0.4089)	-1.3376 (0.1812)		-1.2062 (0.2278)	-1.9183* (0.0552)	-2.0440** (0.0410)	-0.0979 (0.9220)
District									
Big cities + aggl.	-2.0247** (0.0430)	0.1505 (0.8804)	-0.7479 (0.4546)	-0.9404 (0.3472)		-0.8347 (0.4040)	-0.1515 (0.8796)	-0.6971 (0.4858)	-0.5523 (0.5808)
West	-0.1049 (0.9164)	1.6183 (0.1057)	1.5658 (0.1175)	-0.0192 (0.9847)		-1.1342 (0.2567)	-0.7553 (0.4501)	-0.4307 (0.6667)	0.3440 (0.7308)
North	-0.9907 (0.3220)	-1.6537* (0.0985)	-1.7489* (0.0805)	-0.0909 (0.9276)		1.7128* (0.0869)	-2.3843** (0.0172)	-2.016** (0.0440)	0.5311 (0.5954)
East	0.0286 (0.9772)	-0.7591 (0.4479)	0.0176 (0.9860)	0.7718 (0.4403)		2.5820** (0.0099)	-1.2359 (0.2166)	0.0276 (0.9780)	1.2977 (0.1945)
South	-1.4069 (0.1596)	0.7435 (0.4572)	-0.2959 (0.7673)	-1.0663 (0.2864)		-2.3932** (0.0167)	-1.1366 (0.2558)	-1.2740 (0.2027)	-0.1084 (0.9137)
Education									
Primary education	1.2380 (0.2160)	-0.0170 (0.9865)	-0.4301 (0.6672)	-0.3824 (0.7023)		-1.2019 (0.2297)	0.4906 (0.6240)	0.1657 (0.8684)	-0.3075 (0.7586)
Lower voc.	0.3547 (0.7228)	0.3650 (0.7152)	0.4894 (0.6246)	0.1035 (0.9176)		-0.3258 (0.7446)	-0.9145 (0.3605)	0.0878 (0.9300)	1.0168 (0.3093)
Senior sec. voc.	-2.6939** (0.0071)	-0.0247 (0.9803)	-0.3373 (0.7359)	-0.3180 (0.7505)		1.3037 (0.1924)	-1.0061 (0.3144)	-1.4409 (0.1497)	-0.4166 (0.6770)
Higher professional	-0.4360 (0.6628)	-2.0273** (0.0427)	-1.8426* (0.0655)	0.1181 (0.9060)		-0.4431 (0.6577)	-2.7445** (0.0061)	-1.8324* (0.0670)	0.9393 (0.3476)
University	-2.0182** (0.0438)	1.5042 (0.1329)	0.5179 (0.6046)	-0.8492 (0.3960)		-0.5610 (0.5749)	-1.2017 (0.2297)	-0.3391 (0.7346)	0.8167 (0.4143)
Ethnicity									
Native	-1.0998 (0.2715)	-0.4404 (0.6597)	-0.4584 (0.6467)	-0.0114 (0.9909)		0.1213 (0.9035)	-1.8076* (0.0707)	-1.7094* (0.0874)	0.1849 (0.8533)
Foreign	-1.3060 (0.1917)	0.9833 (0.3256)	0.7976 (0.4252)	-0.0699 (0.9443)		-0.9053 (0.3654)	-1.4177 (0.1564)	-0.2477 (0.8044)	1.1713 (0.2416)
Gender									
Male	-1.5581 (0.1192)	-0.2127 (0.8315)	-0.0840 (0.9331)	-0.2127 (0.8315)		2.0012** (0.0454)	-1.0906 (0.2755)	-0.7410 (0.4587)	0.3810 (0.7032)
Female	-1.5819 (0.1137)	1.0314 (0.3024)	-0.1584 (0.8741)	1.0314 (0.3024)		-2.2661** (0.0235)	-2.3145** (0.0207)	-1.6972* (0.0897)	0.6471 (0.5176)
Income									
< EUR 14,300	-0.0745 (0.9406)	-0.0641 (0.9489)	-1.2037 (0.2290)	-1.0959 (0.2735)		-2.1774** (0.0296)	-1.2375 (0.2162)	-1.1586 (0.2469)	0.1278 (0.8983)
EUR 14,300–38,300	-2.5794** (0.0099)	-0.0429 (0.9658)	-0.8194 (0.4126)	-0.7778 (0.4367)		0.2906 (0.7714)	-3.4495** (0.0006)	-3.4970** (0.0005)	0.0487 (0.9612)
EUR 38,300–65,000	-1.9336* (0.0532)	-0.7488 (0.4541)	0.2200 (0.8259)	1.0231 (0.3063)		1.6986* (0.0895)	-0.9631 (0.3356)	-0.1320 (0.8950)	0.8873 (0.3750)
> EUR 65,000	-0.5616 (0.5745)	-1.2387 (0.2157)	-0.8973 (0.3698)	0.2728 (0.7850)		-0.0487 (0.9611)	-1.3132 (0.1893)	0.1948 (0.8456)	1.4832 (0.1382)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of means for number and value of total point-of-sale payments between groups, for age, district, education level, ethnicity, gender, and income class. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

Table 14: Significance test on the difference of means for number and value of point-of-sale debit card payments between groups, for age, district, education level, ethnicity, gender, and income class

Number	A1 vs B	B1 vs C	A2 vs C	A2 vs B1	Value	A1 vs B	B1 vs C	A2 vs C	A2 vs B1
Age									
12-24	0.8626 (0.3885)	-0.1264 (0.8994)	0.3744 (0.7082)	0.4927 (0.6223)		0.8010 (0.4233)	0.2513 (0.8017)	1.2051 (0.2286)	0.8411 (0.4006)
25-44	-2.4259** (0.0153)	-0.2741 (0.7841)	-1.5073 (0.1318)	-1.3197 (0.1870)		1.3991 (0.1619)	-2.3426** (0.0192)	-1.4235 (0.1547)	0.9109 (0.3624)
45-64	-0.9692 (0.3325)	-0.1880 (0.8509)	-0.6196 (0.5356)	-0.4453 (0.6561)		1.0575 (0.2904)	-0.8189 (0.4129)	-0.7454 (0.4561)	0.1176 (0.9064)
65+	1.1832 (0.2368)	0.5099 (0.6102)	0.5458 (0.5852)	0.0145 (0.9884)		-1.6242 (0.1045)	-0.3861 (0.6995)	-1.5741 (0.1157)	-1.1792 (0.2386)
District									
Big cities + aggl.	-1.8458* (0.0651)	0.5314 (0.5952)	-1.3368 (0.1815)	-1.9197* (0.0551)		0.3102 (0.7565)	0.6018 (0.5474)	0.1891 (0.8501)	-0.3956 (0.6925)
West	-0.1698 (0.8652)	2.4562** (0.0141)	1.2786 (0.2011)	-1.2002 (0.2301)		0.1444 (0.8852)	-1.0307 (0.3028)	0.3395 (0.7343)	1.4413 (0.1496)
North	0.2537 (0.7998)	-2.1722** (0.0301)	-1.5997 (0.1099)	0.6063 (0.5445)		1.3676 (0.1717)	-1.6531* (0.0988)	-1.5761 (0.1154)	0.2008 (0.8409)
East	0.5737 (0.5662)	-0.9143 (0.3606)	0.5059 (0.6130)	1.4375 (0.1507)		2.0139** (0.0442)	-0.4653 (0.6418)	-0.7559 (0.4499)	-0.2528 (0.8004)
South	-0.9430 (0.3458)	0.5357 (0.5922)	0.0101 (0.9919)	-0.5492 (0.5829)		-1.4425 (0.1493)	-1.1427 (0.2533)	-1.5888 (0.1123)	-0.4142 (0.6788)
Education									
Primary education	-0.6028 (0.5468)	-0.6512 (0.5151)	-0.0527 (0.9580)	0.6960 (0.4867)		0.3102 (0.7565)	0.6018 (0.5474)	0.1891 (0.8501)	-0.3956 (0.6925)
Lower voc.	0.7306 (0.4651)	-0.6978 (0.4854)	-0.8161 (0.4145)	-0.0674 (0.9462)		0.1444 (0.8852)	-1.0307 (0.3028)	0.3395 (0.7343)	1.4413 (0.1496)
Senior sec. voc.	-0.5055 (0.6132)	0.1928 (0.8471)	0.0164 (0.9869)	-0.1823 (0.8554)		1.3676 (0.1717)	-1.6531* (0.0988)	-1.5761 (0.1154)	0.2008 (0.8409)
Higher professional	-0.3831 (0.7016)	-0.7402 (0.4593)	-0.5547 (0.5791)	0.1580 (0.8745)		2.0139** (0.0442)	-0.4653 (0.6418)	-0.7559 (0.4499)	-0.2528 (0.8004)
University	-1.8839* (0.0599)	1.5910 (0.1120)	0.1104 (0.9121)	-1.5060 (0.1324)		-1.4425 (0.1493)	-1.1427 (0.2533)	-1.5888 (0.1123)	-0.4142 (0.6788)
Ethnicity									
Native	-0.6115 (0.5409)	0.1285 (0.8978)	-0.5660 (0.5714)	-0.6960 (0.4865)		1.0856 (0.2777)	-1.2003 (0.2301)	-0.9777 (0.3283)	0.2741 (0.7840)
Foreign	0.0187 (0.9851)	0.6946 (0.4874)	1.4005 (0.1616)	0.8864 (0.3755)		-0.8146 (0.4154)	-0.8914 (0.3729)	-0.9080 (0.3641)	-0.0833 (0.9336)
Gender									
Male	0.0914 (0.9271)	-0.3566 (0.7214)	-0.3807 (0.7035)	-0.0241 (0.9808)		2.1349** (0.0328)	-1.3116 (0.1897)	-0.4487 (0.6537)	0.9100 (0.3629)
Female	-1.6836* (0.0923)	1.4922 (0.1357)	0.3319 (0.7400)	-1.1888 (0.2346)		-0.6932 (0.4882)	-1.0653 (0.2868)	-1.3470 (0.1781)	-0.2684 (0.7884)
Income									
< EUR 14,300	-0.7427 (0.4579)	0.0536 (0.9573)	-0.2182 (0.8273)	-0.2680 (0.7888)		-0.5429 (0.5874)	-1.3167 (0.1887)	-1.3204 (0.1874)	0.0883 (0.9297)
EUR 14,300–38,300	-2.1245** (0.0337)	-1.1286 (0.2592)	-1.6972* (0.0898)	-0.4995 (0.6174)		1.9375* (0.0528)	-3.0276** (0.0025)	-2.5057** (0.0123)	0.5965 (0.5509)
EUR 38,300–65,000	-0.5192 (0.6037)	0.4389 (0.6608)	0.3584 (0.7201)	-0.0759 (0.9395)		1.1561 (0.2477)	-1.5840 (0.1134)	-0.4774 (0.6331)	1.2013 (0.2298)
> EUR 65,000	0.8676 (0.3858)	-1.4289 (0.1533)	0.1498 (0.8810)	1.6196 (0.1056)		-1.2945 (0.1957)	-0.6870 (0.4923)	-0.4955 (0.6204)	0.1865 (0.8521)

Note: This table reports t -statistics obtained by a two-sample t -test on the difference of means for number and value of point-of-sale debit card payments between groups, for age, district, education level, ethnicity, gender, and income class. * indicates that $p < 0.10$ and ** that $p < 0.05$. P -values in parenthesis. Group A are first-time participants in September, Group B and A1 are first and second-time participants in November, respectively, and Group C, B1, and A2 are first, second, and third-time participants in March, respectively.

Previous DNB Working Papers in 2017

- No. 542 **Jasper de Jong, Marien Ferdinandusse and Josip Funda**, Public capital in the 21st century: As productive as ever?
- No. 543 **Martijn Boermans and Sweder van Wijnbergen**, Contingent convertible bonds: Who invests in European CoCos?
- No. 544 **Yakov Ben-Haim, Maria Demertzis and Jan Willem Van den End**, Fundamental uncertainty and unconventional monetary policy: an info-gap approach
- No. 545 **Thorsten Beck and Steven Poelhekke**, Follow the money: Does the financial sector intermediate natural resource windfalls?

DeNederlandscheBank

EUROSYSTEEM

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl