

The Influence of an Up-Front Experiment on Respondents' Recording Behaviour in Payment Diaries: Evidence from Germany

Tobias Schmidt¹ and Susann Sieber¹

In this article, we analyse the effect of an incentive experiment on German consumers' recording behaviour on the basis of a one-week diary of their point-of-sale expenditure. Part of the experiment, which was carried out shortly before the consumers began filling in their payment diaries, involved consumers rolling a die with a chance of winning either EUR 20 or nothing, that is, they were randomly assigned an incentive. We ask whether respondents' recording behaviour differs depending on whether individuals win or lose. We argue that winners attach a more positive feeling to the survey than losers and therefore show a stronger commitment to the diary. As the incentive experiment is part of a larger experiment to elicit respondents' risk preferences, we also provide evidence on the effect of conducting up-front behavioural experiments in representative surveys. Our results indicate that the outcome of the lottery (rolling of the die) has an impact on the quantity of transactions recorded, but does not affect other aspects of respondents' recording behaviour, such as item nonresponse or rounding. It also has a negligible impact on substantive measures, such as the cash share.

Key words: Incentives; risk experiments; data quality.

1. Introduction

Behavioural economists and psychologists often conduct experiments using convenience samples (e.g., college students). Recently, however, there has been growing interest in embedding behavioural experiments in representative surveys. In many behavioural experiments, participants receive a monetary payoff, which largely depends on the respondents' behaviour, but often contains a random component as well. The interest of behavioural economists lies in the observed behaviour of the participant (e.g., whether a guaranteed payment is preferred over participation in a lottery), and not in the individual payoffs. However, when these experiments are embedded in a standard survey, the payoffs from the experiment can be interpreted as an incentive payment. The payoff may thus have an impact on participants' attitude towards the survey and ultimately affect their reporting behaviour in other parts of the study, for example, the classic questionnaire. Our article provides evidence on this issue. Failure to acknowledge the incentive effects of

¹ Deutsche Bundesbank – Research Centre, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main, Germany. Email: Tobias.Schmidt@bundesbank.de and Susann.Sieber@bundesbank.de.

Acknowledgments: The authors thank Stephanie Eckman, Edith de Leeuw, participants of the ESRA 2015 and 2015 Total Survey Error conferences, as well as four anonymous referees for valuable comments and suggestions.

Disclaimer: The article represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or its staff.

behavioural experiments and the possible biases and measurement error they induce in the collected data might lead to a misinterpretation of the survey results.

To the best of our knowledge, we are the first to link behavioural experiments to the literature on the role of incentives in surveys. Due to the fact that our incentive experiment constitutes part of a behavioural experiment, it differs from the existing literature on survey incentives in two important respects. First, we do not focus on participation incentives, but on incentives randomly assigned to participants between two stages of the interview process, that is between a regular questionnaire and the self-completed diary collecting data on payment transactions. We will therefore be able to assess whether incentives have an effect on respondents' answers, given participation. Second, our participants are aware that other participants receive a different incentive or no incentive at all. Thus, they might be disappointed if they receive nothing or, on the contrary, be very pleased if they obtain the incentive. This positive or negative feeling – in addition to the monetary value of the incentive – should result in more pronounced effects of the incentive on commitment to the survey, measured along various dimensions.

To assess the impact of the incentive on respondents' recording behaviour, we consider different indicators of item nonresponse: our focus will be on transactions that are not reported at all. To be more precise, we study the possible underreporting of cash and low-value transactions. In addition, we look at common measures of data quality: the incidence of incompletely reported transactions and the rounding of transaction amounts. Our analysis fits well in the Total Survey Error literature, as our study sheds light on a nonsampling error (Groves and Lyberg 2010; Groves 2004; Biemer and Lyberg 2003), namely measurement error as a result of a specific design feature.

The basis of our analysis is a behavioural experiment eliciting respondents' risk preferences, which is carried out between a standard questionnaire-type data collection and a self-completed one-week diary of consumers' point-of-sale expenditure. In the first stage of the risk experiment, consumers have the choice between receiving a guaranteed payment of EUR 10 and participating in a lottery with an expected value of EUR 10. Risk-averse consumers will choose the guaranteed payment while more risk-loving consumers will go for the lottery option. Consumers participate in the lottery by rolling a die. Participants either win EUR 20 if they throw a 4, 5, or a 6, or nothing if they throw a 1, 2, or a 3. While economists running these kinds of experiments are mainly interested in whether the respondent opts for the guaranteed payment (risk-averse) over participation in the lottery (risk-loving), we are more interested in the lottery part. We re-use the behavioural experiment to learn something about the effects of incentives on consumers' recording behaviour.

The lottery can be interpreted as an incentive experiment, with the random assignment of an incentive of EUR 20 for the "winners" and nothing for the "losers" resulting from the rolling of the die. In addition to the monetary value, "winning" the game may induce a positive attitude towards the survey. Both the monetary incentive itself and the positive attitude from winning should lead to a higher commitment to the survey and may affect the respondents' recording behaviour in the diary.

In the analysis of incentive effects, we will focus solely on the lottery and restrict the sample to respondents who participated in the lottery. It is necessary to exclude the risk-averse consumers who chose the guaranteed payment in the first stage of the risk

experiment, because the payoff they receive is not assigned randomly, but chosen by the respondents themselves based on their risk preference. Hence, this group of respondents differs systematically from respondents participating in the lottery. It nonetheless makes sense to analyse the behaviour of those risk-loving consumers who chose the lottery option over the guaranteed payment. If the recording behaviour of this selected group differed depending on whether it wins or loses, the survey results could be biased.

We find that the payoff from the lottery seems to have some impact on consumers' commitment to the diary part of the study in that these respondents record more transactions. However, it does not induce a bias on the key qualitative results, such as the share of cash payments. Our results indicate that the monetary incentive, as well as its non-monetary component ("winner" vs. "loser"), given to the participants during the interview, do not lead to biased outcomes in subsequent parts of the survey. This is good news for economists who plan to embed behavioural experiments with a random payoff in representative surveys. The incentives' main effect seems to be the increase in the number of transactions recorded during the first few days of the diary recording period. Survey designers with short diary studies (e.g., one or two days) may thus be able to increase the number of transactions recorded by respondents by paying an additional incentive during the interview, that is, before the diary recording period starts. However, in general, paying an unconditional incentive during an interview or different parts of an interview has very limited impact on the recording behaviour of respondents.

2. Related Literature

Our analysis can be linked to issues related to measurement errors discussed within the Total Survey Error (TSE) framework (see [Groves and Lyberg 2010](#) for an overview). We are mainly concerned with measurement error due to respondents' recording behaviour, including item nonresponse and omissions by respondents (see [Weisberg 2005](#); [Biemer and Lyberg 2003](#)). [Biemer and Lyberg \(2003\)](#) suggest, describing the merits of the TSE framework, that "among the set of alternative designs, the design that gives the smallest total survey error (for a given fixed cost) should be chosen." ([Biemer and Lyberg 2003, 850](#)). Our results can help survey designers wanting to include behavioural experiments in their studies to gain a better understanding of the consequences of the experiment for the total survey error and to make a more informed choice about where in the interview/study process to use behavioural experiments.

Our study is also closely related to the literature on the effects of incentives on data quality and respondents' answering behaviour. From a theoretical point of view, incentives can have an impact on the recording behaviour due to a stronger commitment of the respondent to the survey. It is a well-established fact that responding to survey questions is a complex process that imposes a burden on the interviewee (see, for example, [Jones 2012](#); [Sharp and Frankel 1983](#); [Groves et al. 1992](#)). Various methods to reduce or counter the response burden for interviewees have been proposed (see [Hedlin et al. 2005](#), for business surveys), among them fostering respondents' interest in and commitment to the survey ([Bonke and Fallesen 2010](#); [Davern et al. 2003](#)). In this vein, we argue that providing respondents with an incentive that has both a monetary component (EUR 20) and a non-monetary component ("winner" vs. "loser") will increase their commitment to

the survey and subsequently influence their reporting behaviour. There is some evidence in the literature that this is indeed the case. The study by [Bonke and Fallesen \(2010\)](#) on Danish data comes closest to what we are researching in this article. [Bonke and Fallesen \(2010\)](#) study how different incentives, paid out through a lottery game, explain people's participation rates, choice of survey mode (CATI/CAWI), and data quality in a large-scale Danish survey on time-use and consumption. What makes the study particularly interesting for us is that they investigate respondents' behaviour in the survey diary context. However, in contrast to our study, their participants are not aware that there are lotteries with varying prizes. They find a strong effect of incentives on response rates as well as on mode choice, but no effect *per se* on respondents' behaviour. Neither item nonresponse in the regular questionnaire of their study, nor the number of reported activities or consumed goods and services in the diary differ significantly depending on the incentives provided, if both CATI and CAWI respondents are analysed. They do find some positive effects of incentives with respect to reporting behaviour and data quality for CATI interviewees only. The fact that the impact of incentives on respondents' recording behaviour may be rather limited has also been documented by other researchers. [Davern et al. \(2003\)](#) and [Shettle and Mooney \(1999\)](#) investigate the impact of incentives on classic measures of data quality, such as item nonresponse and the number of edited variables/cases. They find that (prepaid) monetary incentives do not have an impact on data quality. Similarly, [Tzamourani and Lynn \(1999\)](#) show that there is no clear effect of incentives on respondents' recording behaviour, concluding that "... the incentives did not affect the respondents' answers in any way, that is they did not induce bias in the responses." ([Tzamourani and Lynn 1999, 16](#)). [Göritz \(2005\)](#) documents for a web-based survey that if respondents are offered an incentive that is contingent on completing all relevant questions in the questionnaire, their reporting behaviour in terms of the number of omitted questions and other quality indicators does not differ from that of respondents who are not offered an incentive. The same seems to hold true for web-based studies using access panels ([Göritz 2004](#)).

Whether incentives have a positive, negative or no effect at all on respondents' reporting behaviour is nonetheless still an open question. Studies by [Goldenberg and Ryan \(2009\)](#), [Singer et al. \(2000\)](#), [Willimack et al. \(1995\)](#), [James and Bolstein \(1990\)](#), and several of those cited in [Laurie and Lynn \(2009\)](#) found – contrary to the studies cited above – that incentives do have an effect on reporting behaviour and data quality. [Goldenberg and Ryan \(2009\)](#) report that in the US Consumer Expenditure Diary Survey, respondents receiving a pre-paid monetary incentive of USD 20 or USD 40 reported more transactions and also performed better on other indicators of data quality. A similar result is reported by [Goldenberg et al. \(2009\)](#) for the same type of incentives used in the Consumer Expenditure Interview Survey. [Singer et al. \(2000\)](#) show that for some households, paying an incentive reduces item nonresponse. However, the effect is very small: "Only 7 percent is explained by both the demographics and the incentives, and incentives alone explain less than 1 percent of the variance in item non-response." ([Singer et al. 2000, 180](#)). They also find an impact of incentives on the distribution of responses. Respondents receiving an incentive seem to be in a better mood (see also [Schwarz and Clore 1996](#)) and report more optimistic expectations. [Willimack et al. \(1995\)](#) summarise their findings: "In addition, evidence suggests greater response completeness among

responding incentive recipients early in the interview, with no evidence of increased measurement error due to the incentive” (Willimack et al. 1995, 78). James and Bolstein (1990) find that what they call “large” prepaid incentives of USD 2 lead respondents to expend more effort on completing questions in a mail survey. They measure greater effort by the length of the respondents’ answers, the number of comments and number of words written. Interestingly, they also find that large incentives increase “. . . comments that were more favourable towards the survey sponsor” (James and Bolstein 1990, 346), which signals a stronger commitment to the survey. James and Bolstein (1990) cite several older studies (e.g., Godwin 1979, and Shuttleworth 1931) which have also found that respondents receiving monetary incentives have a tendency to provide more comments and more complete responses. A similar result has been found by Goetz et al. (1984).

Given the mixed evidence of incentives on respondents’ recording behaviour, it is difficult to derive a clear hypothesis. However, in contrast to the studies cited above which find no effect of incentives, the monetary incentive in our study is substantial (EUR 20) and also has a non-monetary component (“winner” vs. “loser”). We thus formulate the hypothesis that the incentive will have an effect on respondents’ diary recording behaviour.

Our study also provides evidence on aspects of incentives that have, to the best of our knowledge, not yet been addressed in the literature. The papers cited above mainly deal with conditional and unconditional *participation* incentives, that is, incentives paid to induce the respondents to take part in the surveys in the first place. In our experiment, the incentive is offered to participants on an unconditional basis. We are therefore able to isolate the effect of incentives on recording behaviour without having to worry about confounding effects due to participation choices induced by the incentive.

It is hard to imagine a real-life survey where an incentive similar to ours is paid to respondents during a survey of this kind. Yet, with behavioural experiments incorporating an incentive component becoming more popular in representative surveys, the question of whether these experiments have a (negative) influence on respondents’ recording behaviour is certainly relevant.

3. Data and Variables

In this section we describe the Bundesbank’s Payment Survey and the behavioural experiment we conducted in more detail. We also provide some information about respondent characteristics. Furthermore, we discuss various measures of respondents’ commitment and data quality in a payment diary survey that might be affected by the incentives.

3.1. The Bundesbank’s Payment Survey

In 2014, the Deutsche Bundesbank conducted the third wave of its payment behaviour survey entitled “Payment Behaviour in Germany” (see Deutsche Bundesbank 2015). The survey was carried out by the market research institute MARPLAN on behalf of the Bundesbank. The sample for the survey was drawn using a random-route procedure developed by the Association of German Market and Social Researchers, or ADM for short (see Hoffmeyer-Zlotnik 2003, on random route samples). The face-to-face

interviews with the respondents were conducted between May and July of 2014. The net sample comprises 2,036 persons. The survey is representative for the German population aged 18 and above. Care was taken to ensure that consumers from all 16 federal states were included in the gross sample.

The survey consists of two main parts, a CAPI interview and a drop-off paper and pencil diary. The diary was handed out to participants after completion of the interview. Respondents were also given the option of filling in the diary using a smartphone app, but less than two percent of respondents chose this option. The CAPI interviews took about 30 minutes on average and contained questions on topics such as the ownership and usage of payment instruments, cash withdrawal behaviour, perceived risks of payment instrument usage, and respondents' demographics.

The payment diary collected information on actual transactions over a period of seven days and specifically refers to direct payment transactions at the point-of-sale, that is, all transactions apart from recurring transactions, such as rent payments, insurance premiums, telephone and utility bills. The information collected in the paper and pencil diary includes the euro amount for each transaction, the location where the transaction took place (16 different possible locations including "retail purchases for day-to-day needs", "filling stations", "restaurants", "e-commerce", "payments to private individuals", etc.) and the payment medium used to settle the transaction (cash and a list of eleven cashless payment methods, e.g., debit cards, credit cards, e-payment schemes, payment schemes via mobile phone, contactless card payments). In addition, respondents had to provide information on cash withdrawals in the diary. The diary contains space for up to eight transactions for each day and some spare pages in case more than eight transactions were made in any one day. At the top of each page of the diary, the respondents were asked to fill in the date and then list all transactions pertaining to this date. The printed diary also contains a page with an example of how to fill in the diary, and the interviewers explained the transaction recording procedure to the respondents when they handed over the diary.

Interviews were conducted and diary data collected across the entire survey period and on all days of the week (including weekends and public holidays). Participants were instructed to start the diary on the day following the face-to-face interview. In practice, we see that more respondents started filling in their diaries on Wednesdays (19%) compared to Sundays (9%). To a certain extent, the unequal distribution of diary days might be caused by the fact that the probability of reaching a participant for an interview is higher on some days than on others. However, we cannot rule out that some participants did not closely follow the instructions and put off starting the diary until a day on which they had a transaction to report.

Respondents received participation incentives both for answering the survey and for filling in the diary. After completing the CAPI interview, the interviewers gave the respondents a pen and a notepad to help them make notes on their payment transactions during the day and a package of shredded banknotes as a participation incentive. A monetary incentive of EUR 10 was sent to everyone who answered the payment diary and returned it to the market research institute. All these incentives were paid out to respondents, irrespective of the outcome of the risk experiment. Therefore, they do not confound our analysis.

3.2. Behavioural Experiment

A novel feature of the 2014 survey was a behavioural experiment which is supposed to elicit respondents' risk preferences. This type of experiments is typically used in behavioural economics (see, for example, [Dohmen et al. 2011](#); [Dohmen et al. 2010](#); [Eckel and Grossmann 2002](#); [Charness et al. 2013](#); and [Awel and Azomahou 2015](#)). The aim of including the experiment in the survey was to have a measure for participants' risk preferences, which could then be linked to information provided in other parts of the survey, for example on the adoption of innovative payment instruments.

The risk experiment was performed directly after completing the face-to-face interview and carried out by the interviewers with all respondents. Respondents were told that they had the possibility to take part in a brief experiment which would take about five minutes and would not cost them anything. They were also told that they had the chance to win a cash prize. Out of the 2,036 persons completing the CAPI, 1,952 respondents (almost 96%) decided to take part in the risk experiment. The behavioural experiment was conducted using a representative sample of the population without considerable bias in the group of participants due to non-participation of respondents. This makes the experiment exceptionally valuable for behavioural economists who are usually constrained to conducting such experiments with a narrow subgroup of the population, such as university students (see, for example, [Charness and Gneezy 2010](#)).

The sequence of the behavioural experiment can be seen in [Figure 1](#). Respondents who decided to take part in the experiment were given the choice between receiving a guaranteed payment of EUR 10 or the chance to participate in a lottery draw with an expected value of EUR 10 (50% chance of winning either EUR 20 or nothing at all). The rationale of the experiment is that risk-averse participants will choose the guaranteed payment while risk-loving participants will decide to take part in the lottery. This provides the economist with a rough measure of respondents' risk preference. Out of a total of 1,952 consumers participating in the risk experiment 994 took the guaranteed payment of EUR 10, and 958 consumers chose to participate in the lottery (see bottom line of [Table 1](#)).

To test whether respondents who took part in the lottery and those who chose the guaranteed payment differ in terms of sociodemographic characteristics, we run a series of t-tests. As expected, we find significant differences between the groups. Columns (II), (III), and (VI) in [Table 1](#) reveal that lottery players are younger, more likely to be male and have a higher income on average (for literature on the link between sociodemographic characteristics and risk aversion see, for example, [Eckel and Grossman 2008](#); [Borghans et al. 2009](#); and [Halek and Eisenhauer 2001](#)). In addition, we use the CAPI interview to check whether the two groups differ in their self-assessed risk preferences, their technological literacy and their attitude towards new payment methods. [Tables 1 and 2](#) in the Supplemental data present detailed results by group (available online at: <http://dx.doi.org/10.1515/jos-2017-0021>). Participants taking the guaranteed payment are significantly more risk-averse and prudent than those who chose to participate in the lottery, which supports the validity of the behavioural experiment. What is more, they are less technologically literate (i.e., less likely to use the internet and electronic devices) and are more conservative in their payment behaviour (i.e., less open to payment innovations, less likely to own a credit card or use e-payment schemes).

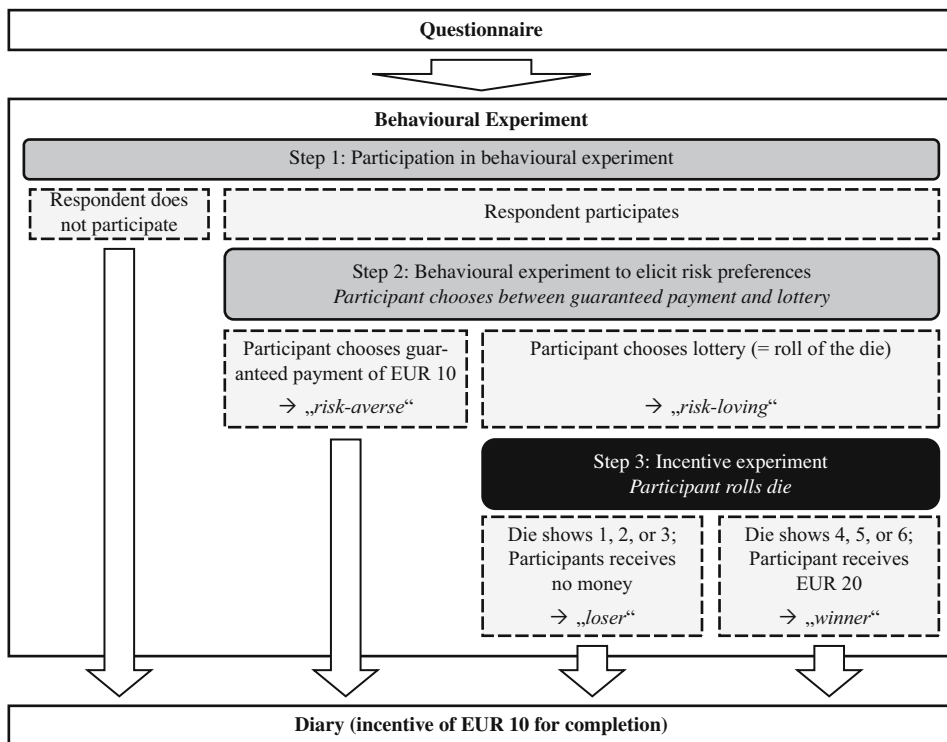


Fig. 1. Sequence of the behavioural experiment.

We also check for interviewer effects in the decision between the guaranteed payment and the lottery (see the lower part of Table 1). Respondents who had female and/or younger interviewers were more likely to choose the lottery option.

3.3. Incentive Experiment

Only the risk-loving participants choosing to participate in the lottery went on to participate in the final step of the risk experiment in which the lottery game was played and the payoffs were determined. The lottery game consisted of rolling a die. Participants either won EUR 20 if they threw a 4, 5, or a 6, or nothing if they threw a 1, 2, or a 3. This last stage of the risk experiment constitutes the incentive experiment, which is the main focus of our research. The payoff amounts resulting from the rolling of the die can be interpreted as an incentive payment. It should be noted, however, that the experiment was not specifically designed to test the effect of incentives on consumers' recording behaviour; instead, we "re-use" the second part of the risk experiment, which allows us to learn something about the effects of incentives on consumers' recording behaviour. In a pure incentive experiment, additional control groups would have been included in the experiment, for example consumers who chose the guaranteed payment would have had to roll the die, etc. This limits the scope of our study. Nonetheless we are able to provide some insight into how recording behaviour and data quality is affected by including a standard behavioural risk experiment in a survey. The "assignment" of incentive payments

Table 1. Sociodemographic characteristics of respondents and interviewers by participation in and outcome of the behavioural experiment.

| Variable | (I) | | (II) | | (III) | | (IV) | | (V) | | (VI) | (VII) |
|-----------------------------------|------------------|-------|--------------------|-------|-------|--|---------|---------|---------|--|---|-------------------------------|
| | All participants | | Guaranteed payment | | Total | | Lottery | | Winners | | Difference guaranteed payment vs. lottery | Difference losers vs. winners |
| | | | | | | | Losers | Winners | | | | |
| Age in years | 46.79 | 47.90 | 45.64 | 46.54 | 45.03 | | | | | | 2.25*** | 1.51 |
| Female | 0.55 | 0.59 | 0.51 | 0.50 | 0.52 | | | | | | 0.08*** | -0.03 |
| Lives without partner | 0.47 | 0.47 | 0.48 | 0.50 | 0.47 | | | | | | -0.02 | 0.03 |
| Household size | 2.22 | 2.19 | 2.25 | 2.25 | 2.24 | | | | | | -0.05 | 0.01 |
| Education (4 groups) | 2.05 | 2.02 | 2.10 | 2.08 | 2.11 | | | | | | -0.08* | -0.03 |
| Household net income (12 groups) | 5.04 | 4.90 | 5.18 | 5.30 | 5.09 | | | | | | -0.28** | 0.21 |
| Individual net income (12 groups) | 3.45 | 3.28 | 3.63 | 3.72 | 3.57 | | | | | | -0.34*** | 0.15 |
| East German | 0.19 | 0.17 | 0.21 | 0.18 | 0.23 | | | | | | -0.04** | -0.05** |
| Interviewer female | 0.52 | 0.49 | 0.55 | 0.53 | 0.57 | | | | | | -0.07*** | -0.04 |
| Age of interviewer | 56.74 | 57.40 | 56.74 | 56.15 | 55.99 | | | | | | 1.34** | 0.15 |
| Number of observations | 1,952 | 994 | 958 | 391 | 567 | | | | | | | |

Notes: Number of observations differs for each variable as some values are missing. **, ***, **** mean difference is significant at the 90%, 95%, or 99% level (two-sided test).

Table 2. Results of estimations on the number of transactions and empty diaries.

| Variable | (I) Total number of TA (count) | (II) Number of TA on day 1 (count) | (III) Number of TA on day 2 (count) | (IV) Empty diary on day 1 (dummy) | (V) Empty diary on day 2 (dummy) |
|-----------------------|---|--|---|--|---|
| Negative Binomial | | | | | |
| WINNER | 0.057* [0.032] | 0.095* [0.051] | 0.091* [0.055] | -0.229** [0.097] | -0.224** [0.094] |
| Control variables | Included (see Tables A1 and A5 for results and definitions) | | | | |
| Observations | 949 | 949 | 949 | 949 | 949 |
| alpha | 0.129 [0.011] | 0.000 [0.000] | 0.000 [0.011] | | |
| Chi2 | 110.55 | 113.44 | 99.21 | 63.50 | 72.76 |
| Pseudo-R ² | | | | 0.067 | 0.073 |

Notes: *, **, *** mean coefficient is significant at the 90%, 95%, or 99% level. Robust standard errors are given in brackets. TA stands for the number of transactions.

to consumers was obviously random if the die was non-biased and interviewers carried out the experiment correctly. However, the share of winners is about ten percentage points higher than expected, at almost 60%. Out of the 958 participants who rolled the die 567 won the EUR 20.

It cannot be ruled out that the interviewers deviated from the instructions and, for example, allowed the respondents to roll the die several times or simply paid out the EUR 20 regardless of the number on the die. If this were the case, this would mean a deviation from an experimental setting with randomly assigned outcomes. It cannot be checked *ex-post* why the realised and expected values do not match. However, in checks carried out by the survey agency after the interviews, respondents confirmed that the interviewers had actually offered the experiment and that they had correctly noted the result of rolling the die. Unfortunately, this does not constitute a thorough check of whether consumers were allowed to roll the die more than once. What is reassuring is that we do not find significant effects of interviewers' gender and age on the outcome of the roll of the die. As an additional check for the existence of interviewer effects, we re-calculated all results presented in Section 4, after eliminating cases from interviewers with very high winning rates. The results did not change.

To make sure that no bias with respect to observable sociodemographics exists between winners and losers, we again run a series of t-tests (for results see columns (IV), (V), and (VII) in Table 1). All but one come up negative, indicating that the composition of the two groups is very similar. We also run probit regressions with the outcome of the roll of the die (win/lose) as the dependent variable and sociodemographic variables as explanatory variables. They broadly support the results of the individual t-tests (see Table 3 in the Supplemental data online at <http://dx.doi.org/10.1515/jos-2017-0021>). In addition, we do not find any significant differences between winners and losers in their self-assessed risk preferences, their technological literacy and their attitude towards new payment methods as stated in the CAPI interview (see Tables 1 and 2 in the Supplemental data online at: <http://dx.doi.org/10.1515/jos-2017-0021>).

In summary, the randomisation of incentives worked well despite the fact that there are too many winners. We provide robustness tests below where we will show that our results prevail even after controlling for sociodemographics in a regression.

3.4. Measures of Respondents' Recording Behaviour and Data Quality

A key decision we had to take is along which dimensions we want to assess respondents' recording behaviour in payment diary surveys. An obvious first choice is to examine unit nonresponse and nonresponse bias (see, for example, Bonke and Fallesen 2010, 24). However, we do not consider these measures here as 951 out of the 958 participants who rolled the die returned a completed diary. Instead, we follow the literature cited above in Section 2 (see, for example, Davern et al. 2003; Shettle and Mooney 1999) and look, among other things, at measures related to item nonresponse. Item nonresponse can come in the form of a missing answer for an individual transaction or a missing transaction. We use the "share of incomplete transactions" as our measure of (classic) item nonresponse of the first type. A transaction is incomplete if any of the required information regarding the transaction is missing.

While classic item nonresponse is easy to detect and measure, transactions which are missing completely are harder to examine. Usually, no reference statistics are available which would allow the researcher to detect underreporting of transactions. A comparison of the total number of transactions or activities reported by consumers with incentives and those without has thus been used as an indirect measure (e.g., [Fricker and Tourangeau 2010](#), [Axhausen et al. 2002](#)). We follow this literature and look at the total number of transactions reported for each day and their structure with respect to the payment method chosen (cash vs. non-cash payments) and transaction size (payment value below EUR 5). Our approach parallels [Fricker and Tourangeau's \(2010\)](#) classification of activities into different types, like basic daily activities and other activities.

Finally, we look at the share of rounded values (see, for example, [Fricker and Tourangeau 2010](#)). The share of rounded values should be an indicator of less commitment to the survey, as providing an exact figure can be bothersome for consumers.

4. Results

The results we present are based on those respondents who took part in the lottery game. The reasons for excluding those participants who opted for the guaranteed payment are discussed in Section 1. Out of the 958 persons who rolled the die, two losers and four winners had to be excluded because they did not return the payment diary. One winner was excluded because this person returned the diary without noting down any transactions. This leaves us with 951 observations on which we base our results. Due to the low number of cases excluded, we are not concerned about any nonresponse bias within the sample of risk-loving participants.

4.1. Number of Transactions

The first indicator we look at is the number of transactions. We find that winners record significantly more transactions, especially on the first two days and the last day of the diary period than respondents who did not win (see [Figure 2](#)). Throughout the seven-day diary week, winners noted down 10.40 transactions (1.49 transactions per day), while losers reported only 9.84 transactions (1.41 transactions per day). The number of reported transactions is highest on days 1, 2, and 3 for winners, and on days 1 and 3 for losers. This implies that both winners and losers show a strong initial commitment to the diary, but the level of commitment appears to be even higher for winners than for losers.

In order to rule out that sociodemographic variables or the day of the week confound the descriptive results presented above (see [Figure 2](#)), we regress the number of transactions recorded on the variable of interest "WINNER" and control variables. [Table 2](#) provides a summary of those regressions (results for control variables reported in [Table A1](#) in the Appendix). Columns (I), (II), and (III) of [Table 2](#) confirm that winners report more transactions in total and on the first two diary days, even after controlling for consumers characteristics, such as age, gender, income and household size as well as the day of the week (where applicable). The estimated coefficient for "WINNER" of 0.057 in column (I) corresponds to a difference in the number of transactions between winners and losers of almost 0.6 (see also [Table A2](#) in the Appendix). For the first and second day of the diary, the difference between winners and losers also remains significant at conventional levels.

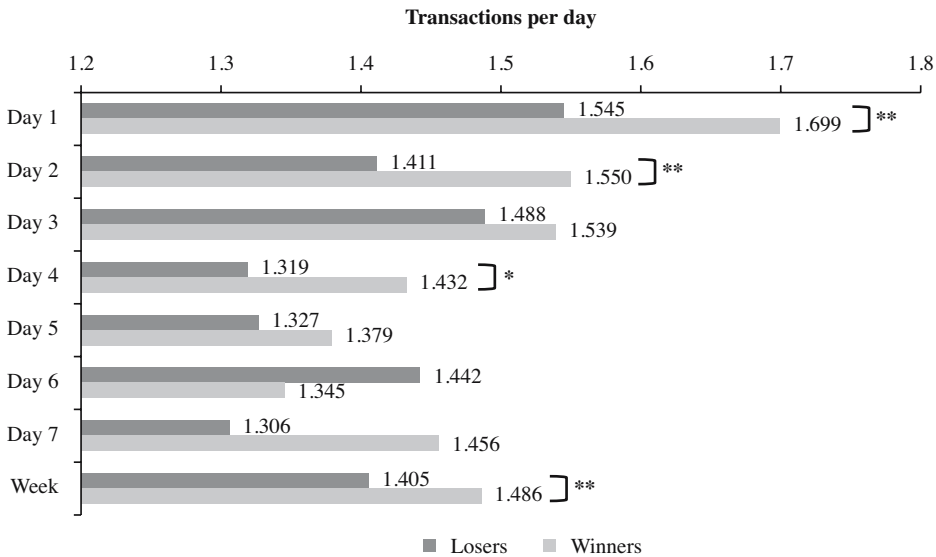


Fig. 2. Average number of transactions per person per diary day.

Notes: *, **, *** mean difference is significant at the 90%, 95%, or 99% level (one-sided t-test).

This result is mainly driven by the fact that the share of diaries without any transactions on days one and two is significantly lower for winners than for participants who did not win the EUR 20 (see columns (IV) and (V) in Table 2).

In order to be able to attribute the differences in the number of recorded transactions to an incentive effect, we have to rule out that the higher number of transactions for winners can be traced back to an income effect. In this case, winners would simply feel “richer” due to receiving the EUR 20 and therefore spend more and have more transactions than if they had not received the money. If the income effect existed, it should be larger for consumers with a low income than for those with a high income. Consequently, the difference in the number of transactions reported by winners and losers should be larger for respondents in the lowest income category compared with those in higher income classes. We test this assumption by including an interaction variable WINNER*INCOME in our regressions on the number of transactions reported and the share of empty diaries, which allows us to identify the effect of winning on the quantity of transactions reported in each income category. Figure 3 gives a graphical presentation of the estimated number of transactions in the diary week for winners and losers in different income categories (see also Tables 4 and 5 in the Supplemental data online <http://dx.doi.org/10.1515/jos-2017-0021>). Winners with a low income report only slightly more transactions than losers in the same income category. The difference is by no means statistically significant. In the middle income category, winners report substantially more transactions than losers (the difference is 1.2 transactions). A Wald test shows that the hypothesis of an equal number of transactions reported by winners and losers in this income group can be rejected at the 95% confidence level. Consequently, the higher number of transactions reported by winners in the estimation without interaction terms can mainly be traced back

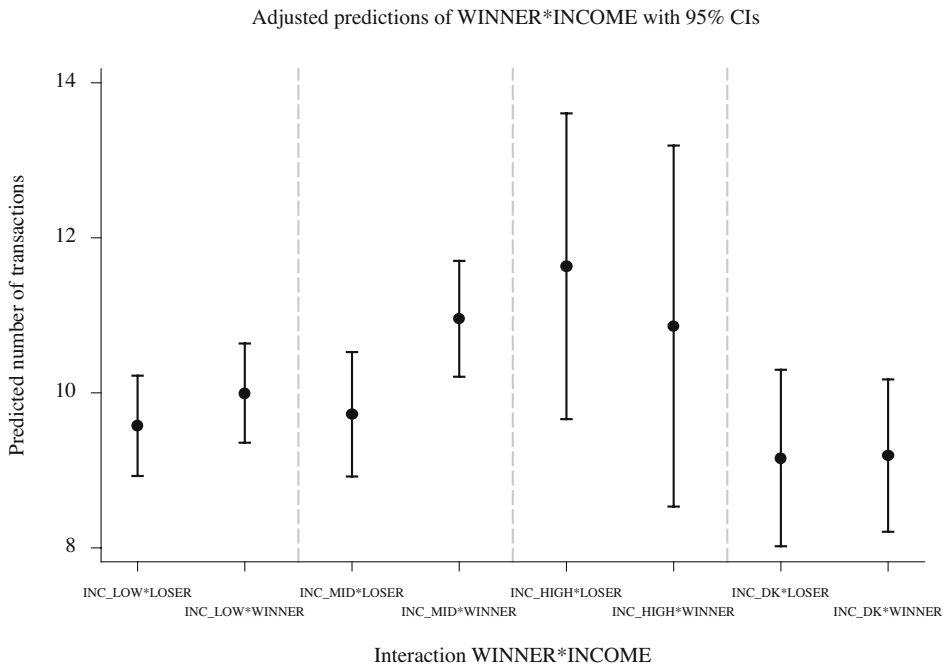


Fig. 3. Predictions of the number of transactions for winners and losers, by income category.

to middle income participants, not to those respondents for whom a potential income effect is expected to be the largest.

As an additional test we look at cash transactions. If winners immediately spent the EUR 20 they receive (in addition to their normal spending), their amount spent (in cash) during the diary reporting week should, all other things being equal, be higher than the amount spent by losers. Using robust regression techniques that down-weight outliers we show that there is no statistically significant difference between winners and losers in total spending and cash spending (see Table A3 in the Appendix). Although the estimated coefficients for winners are positive, they are far below EUR 20. Again, this leads us to conclude that income effects do not drive our results on the quantity of transactions reported in the payment diaries.

To sum up, we find that winners report more transactions than losers, in particular on the first and second diary day. These differences are neither driven by the sociodemographics of the respondents, nor by income effects.

A difference in the number of reported transactions does not affect the substantial outcomes of the survey or response quality unless the additional transactions reported differ systematically from the other transactions. To shed some light on this issue we analyse the structure of payments for each day for the two groups of consumers below.

4.2. Type and Size of Transactions

To investigate whether winners report specific or special payments on days one and two, we first look at cash transactions as a specific type of transaction. Days one and two exhibit

the highest average number of cash transactions for winners (see Figure A1 in the Appendix). However, day one also sticks out for losers. Since cash transactions follow a similar pattern to the total number of transactions described in Subsection 4.1, looking at cash shares is more informative. The cash share is computed by first calculating the share of cash transactions for each individual and diary day and then taking the average across all individuals. Figure 4 shows that the difference with respect to the cash share between those who win and those who lose is insignificant on each of the seven diary days as well as for the whole diary week. Moreover, it is sometimes positive and sometimes negative, and alternates for the first two days.

In addition, we do not find significant differences in the reported cash shares between day one and all other days of the diary week – neither for winners nor for losers. Looking at these factors together, there is no evidence of a disproportional reporting of cash payments on days one and two for winners.

Another classic categorisation of transactions is by their size (see, for example, Bagnall et al. 2016). There is evidence that small-value transactions (below EUR 5) are underreported in diary surveys (Jonker and Kosse 2013). A high number of small-value transactions and, in particular, a high share of this type of transaction among all transactions are signs of good data quality and would suggest a strong commitment to the survey. In our payment diary, the share of transactions below EUR 5 does not differ significantly between winners and losers (see Table 3). The differences observed between winners and losers are insignificant overall and on all individual days. Similarly, the differences in average transaction amounts between winners and losers are insignificant for each diary day and for all days taken together (see Table 3).

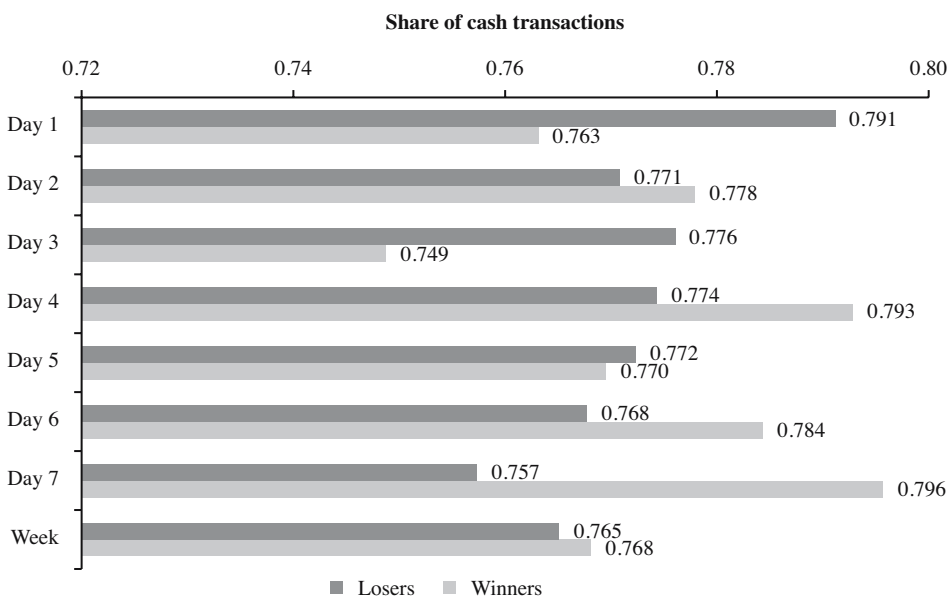


Fig. 4. Average share of cash transactions per person per day.

Notes: Calculated for individuals with more than zero transactions on a given day. *, **, *** mean difference significant at the 90%, 95%, or 99% level (two-sided *t*-test).

Table 3. Average share of small-value transactions below EUR 5 and average transaction amount (in EUR) per person per diary day.

| | Share of small-value TA | | | Average TA amount | | |
|------------|-------------------------|---------|-------------------------------|-------------------|---------|-------------------------------|
| | Losers | Winners | Difference losers vs. winners | Losers | Winners | Difference losers vs. winners |
| Day 1 | 0.247 | 0.226 | -0.021 | 28.48 | 24.51 | -3.97 |
| Day 2 | 0.217 | 0.225 | 0.009 | 25.29 | 25.16 | -0.13 |
| Day 3 | 0.195 | 0.222 | 0.027 | 25.53 | 26.97 | 1.44 |
| Day 4 | 0.225 | 0.210 | -0.015 | 25.90 | 24.07 | -1.83 |
| Day 5 | 0.195 | 0.207 | 0.013 | 24.22 | 23.98 | -0.24 |
| Day 6 | 0.194 | 0.239 | 0.046 | 23.99 | 23.72 | -0.27 |
| Day 7 | 0.216 | 0.221 | 0.005 | 26.42 | 29.28 | 2.86 |
| Diary week | 0.206 | 0.222 | 0.016 | 26.08 | 26.98 | 0.90 |

Notes: Calculated for individuals with more than zero transactions on a given day. *, **, *** mean difference significant at the 90%, 95%, or 99% level (two-sided t-test). TA is short for transactions.

In order to make sure that differences between winners and losers in the type and size of payments reported are not driven by sociodemographic factors, we run a series of regressions with payment type and payment size as dependent variables, the outcome of the roll of the die (winner or loser) as an independent variable and various sociodemographics as control variables. The regression analysis confirms the descriptive results: being a winner does not have a significant effect on the share of cash payments, the share of small-value transactions and the average transaction amount (see columns (I) to (III) in Table A4 in the Appendix).

All these results indicate that the higher quantity of transactions reported on the first days does not induce a bias in substantive diary results.

4.3. Incomplete Transaction Data and Rounding

In the previous sections we looked at item nonresponse with respect to completely missing transactions. In this section we focus on the classic item nonresponse measures, that is missing information on recorded transactions. In the payment diary, respondents were asked to answer several questions concerning each individual transaction. Besides stating the amount, the payment instrument used and the location where the transaction took place, respondents were also asked about various circumstances that might have influenced their payment choice. Up to eight variables pertain to one transaction. Given that all the requested information requires the respondent to read the diary carefully and to remember the actual transaction precisely, this can be bothersome for some respondents. Similarly, reporting precise amounts in the diary can be a burden for some consumers. They have to remember the exact amount or keep the receipt, for example. Thus, many people round the reported transaction values in payment diaries. In the 2014 Bundesbank payment diary, we see that, on average, about one quarter (29%) of all transactions per person per day are rounded to the next whole euro.

Table 4. Average share of incomplete transactions and average share of rounded transaction values per person per diary day.

| | Share of incomplete TA | | | Share of rounded TA values | | |
|------------|------------------------|---------|-------------------------------|----------------------------|---------|-------------------------------|
| | Losers | Winners | Difference losers vs. winners | Losers | Winners | Difference losers vs. winners |
| Day 1 | 0.319 | 0.282 | -0.036 | 0.237 | 0.258 | 0.022 |
| Day 2 | 0.311 | 0.322 | 0.012 | 0.278 | 0.267 | -0.011 |
| Day 3 | 0.379 | 0.329 | -0.050 | 0.294 | 0.316 | 0.023 |
| Day 4 | 0.376 | 0.342 | -0.034 | 0.271 | 0.306 | 0.035 |
| Day 5 | 0.385 | 0.345 | -0.039 | 0.356 | 0.292 | -0.064** |
| Day 6 | 0.347 | 0.386 | 0.039 | 0.306 | 0.290 | -0.016 |
| Day 7 | 0.423 | 0.376 | -0.047 | 0.284 | 0.302 | 0.018 |
| Diary week | 0.368 | 0.337 | -0.031 | 0.287 | 0.293 | 0.006 |

Notes: Calculated for individuals with more than zero transactions on a given day.

*, **, *** mean difference is significant at the 90%, 95%, or 99% level (two-sided t-test). TA is short for transactions.

Table 4 shows that, on average, around one in three transactions contains missing information. There are slightly fewer incomplete transactions for winners (33.7%) than for losers (36.8%). Both winners and losers report more incomplete transactions as the diary week progresses. However, the shares of incomplete transactions do not differ significantly overall or for any of the individual days of the diary between winners and losers.

The picture looks similar for the rounding of values. Table 4 shows that the rate of rounded transaction values is only marginally higher for consumers who win the EUR 20. Consumers winning and those losing in the experiment both show significantly increasing rates of rounded values as the diary progresses, under-scoring the value of this measure as a quality indicator. Overall, the difference between the two groups with respect to rounded values is not significant, with the exception of day 5.

Again, regressions with our two quality indicators as dependent variables, the outcome of the lottery (winner or loser) as an independent variable and sociodemographics as control variables confirm that being a winner does not have a significant effect on the share of incompletely reported transactions and rounded transaction values (see columns (IV) and (V) of Table A4 in the Appendix).

5. Conclusions

In this article, we analyse the effect of an incentive experiment on German consumers' recording behaviour in a one-week diary based on their point-of-sale expenditure. As part of the experiment, participants were asked to roll a die and had the chance of winning EUR 20 depending on the number thrown on the die. We interpret this as a random assignment of a monetary incentive, where, in contrast to most other incentive experiments, the persons receiving the incentive (winner) know that they received a higher incentive than some of the other participants. The experiment itself could thus stimulate a positive sentiment in winners and a negative sentiment in losers towards the survey and the

diary in general. Consequently, we expect winners to exhibit a greater commitment to the diary than losers, and thus a differential reporting behaviour.

We measure respondents' recording behaviour using several indicators: the number of transactions recorded, the share of cash transactions, the share of low-value transactions (below EUR 5), the share of rounded transactions values, and the share of transactions with incomplete information.

We find that the overall impact of paying an unconditional incentive during an interview or different parts of an interview on the recording behaviour is rather limited. Our results indicate that the outcome of the game has an impact on the quantity of transactions recorded, but does not affect other aspects of respondents' recording behaviour, such as item nonresponse or rounding. It also has a negligible impact on substantive measures, such as the cash share. The low variation between the two groups with respect to the cash share and the share of small-value transactions indicates that the key findings from the diary, such as the overall cash share of point-of-sale transactions and the share of transactions within certain value ranges, are not biased by the outcome of the experiment.

Our results have implications beyond research on incentive effects. To the best of our knowledge, we are the first to look into the effects of behavioural or psychological experiments with monetary rewards carried out during a representative national survey. The incentive experiment we examine was an integral part of a behavioural experiment eliciting respondents' risk inclination, which was performed on a representative sample of the adult population. Incorporating behavioural experiments in such surveys is a rather new and promising approach considering that these experiments were previously often carried out with a non-representative part of the population (e.g., students). Up to now, little is known about how such experiments influence participants' attitude towards the survey to which the experiment is linked. Our research indicates that including an experiment does not negatively affect the survey's data quality. Survey designers can thus include these types of experiments without having to worry about increasing the total survey error of their study.

We were mainly concerned with the incentive effects of the behavioural experiment, both monetary and non-monetary. Whether other features of the experiment also play a role for respondents' reporting behaviour is still an open issue. It is feasible to assume that the design and administration of the experiment by the interviewers could potentially confound and even counteract the incentive effects. More research on these issues is necessary.

Appendix

Table A1. Results of estimations on the number of transactions and empty diaries.

| Variable | (I) | | (II) | | (III) | | (IV) | | (V) | |
|-------------|----------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|------------------|
| | Total number of TA (count) | Number of TA on day 1 (count) | Number of TA on day 1 (count) | Number of TA on day 2 (count) | Number of TA on day 2 (count) | Empty diary on day 1 (dummy) | Empty diary on day 2 (dummy) | Empty diary on day 1 (dummy) | Empty diary on day 2 (dummy) | |
| | Negative Binomial | | | | | | | | | |
| WINNER | 0.057* [0.032] | 0.095* [0.051] | 0.091* [0.055] | 0.091* [0.055] | 0.091* [0.055] | -0.229** [0.097] | -0.229** [0.097] | -0.229** [0.097] | -0.229** [0.097] | -0.224** [0.094] |
| AGE_25_34 | 0.051 [0.061] | -0.045 [0.111] | 0.124 [0.126] | 0.124 [0.126] | 0.124 [0.126] | -0.035 [0.221] | -0.035 [0.221] | -0.035 [0.221] | -0.035 [0.221] | 0.006 [0.208] |
| AGE_35_44 | 0.232*** [0.064] | 0.223** [0.110] | 0.274** [0.124] | 0.274** [0.124] | 0.274** [0.124] | -0.166 [0.228] | -0.166 [0.228] | -0.166 [0.228] | -0.166 [0.228] | -0.159 [0.212] |
| AGE_45_54 | 0.205*** [0.062] | 0.127 [0.109] | 0.211* [0.124] | 0.211* [0.124] | 0.211* [0.124] | -0.075 [0.224] | -0.075 [0.224] | -0.075 [0.224] | -0.075 [0.224] | 0.051 [0.205] |
| AGE_55_64 | 0.193*** [0.069] | 0.124 [0.118] | 0.246* [0.136] | 0.246* [0.136] | 0.246* [0.136] | -0.110 [0.236] | -0.110 [0.236] | -0.110 [0.236] | -0.110 [0.236] | -0.051 [0.224] |
| AGE_65+ | 0.208*** [0.093] | 0.160 [0.159] | 0.219 [0.166] | 0.219 [0.166] | 0.219 [0.166] | 0.126 [0.289] | 0.126 [0.289] | 0.126 [0.289] | 0.126 [0.289] | -0.227 [0.271] |
| FEMALE | 0.022 [0.033] | 0.004 [0.053] | -0.002 [0.056] | -0.002 [0.056] | -0.002 [0.056] | 0.036 [0.103] | 0.036 [0.103] | 0.036 [0.103] | 0.036 [0.103] | -0.001 [0.099] |
| HH_SIZE_2 | 0.031 [0.040] | 0.135** [0.062] | -0.015 [0.069] | -0.015 [0.069] | -0.015 [0.069] | 0.007 [0.120] | 0.007 [0.120] | 0.007 [0.120] | 0.007 [0.120] | -0.062 [0.115] |
| HH_SIZE_3 | -0.044 [0.045] | 0.019 [0.078] | -0.029 [0.080] | -0.029 [0.080] | -0.029 [0.080] | 0.089 [0.151] | 0.089 [0.151] | 0.089 [0.151] | 0.089 [0.151] | -0.211 [0.149] |
| HH_SIZE_4+ | 0.010 [0.048] | 0.127* [0.074] | 0.012 [0.082] | 0.012 [0.082] | 0.012 [0.082] | 0.012 [0.156] | 0.012 [0.156] | 0.012 [0.156] | 0.012 [0.156] | -0.079 [0.146] |
| EDU_MEDIUM | 0.097** [0.041] | 0.032 [0.063] | 0.105 [0.067] | 0.105 [0.067] | 0.105 [0.067] | 0.128 [0.122] | 0.128 [0.122] | 0.128 [0.122] | 0.128 [0.122] | -0.155 [0.113] |
| EDU_HIGH | 0.247*** [0.048] | 0.155* [0.080] | 0.204** [0.082] | 0.204** [0.082] | 0.204** [0.082] | 0.147 [0.154] | 0.147 [0.154] | 0.147 [0.154] | 0.147 [0.154] | -0.296** [0.147] |
| EDU_UNI | 0.269*** [0.057] | 0.213** [0.092] | 0.234** [0.094] | 0.234** [0.094] | 0.234** [0.094] | 0.153 [0.170] | 0.153 [0.170] | 0.153 [0.170] | 0.153 [0.170] | -0.121 [0.169] |
| INC_MID | 0.061 [0.039] | 0.086 [0.060] | 0.042 [0.071] | 0.042 [0.071] | 0.042 [0.071] | -0.169 [0.123] | -0.169 [0.123] | -0.169 [0.123] | -0.169 [0.123] | -0.076 [0.115] |
| INC_HIGH | 0.140* [0.076] | -0.031 [0.113] | 0.283** [0.112] | 0.283** [0.112] | 0.283** [0.112] | -0.256 [0.225] | -0.256 [0.225] | -0.256 [0.225] | -0.256 [0.225] | -0.468* [0.243] |
| INC_DK | -0.070 [0.049] | -0.064 [0.081] | -0.009 [0.089] | -0.009 [0.089] | -0.009 [0.089] | -0.092 [0.162] | -0.092 [0.162] | -0.092 [0.162] | -0.092 [0.162] | -0.101 [0.160] |
| REGION_EAST | 0.007 [0.039] | -0.037 [0.064] | 0.188*** [0.066] | 0.188*** [0.066] | 0.188*** [0.066] | 0.069 [0.120] | 0.069 [0.120] | 0.069 [0.120] | 0.069 [0.120] | -0.091 [0.118] |

Table A1. Continued

| Variable | (I) | (II) | (III) | (IV) | (V) |
|-----------------------|----------------------------|-------------------------------|-------------------------------|------------------------------|------------------------------|
| | Total number of TA (count) | Number of TA on day 1 (count) | Number of TA on day 2 (count) | Empty diary on day 1 (dummy) | Empty diary on day 2 (dummy) |
| Negative Binomial | | | | | |
| OCC_TRAIN | 0.046 [0.089] | 0.162 [0.150] | -0.003 [0.171] | -0.174 [0.299] | 0.175 [0.273] |
| OCC_WORK | 0.047 [0.058] | 0.195** [0.090] | 0.002 [0.100] | -0.173 [0.171] | -0.110 [0.163] |
| OCC_OTHER | -0.041 [0.079] | 0.069 [0.122] | -0.022 [0.132] | 0.265 [0.214] | 0.017 [0.209] |
| Probit | | | | | |
| SUNDAY | n.a. | -0.683*** [0.123] | -0.644*** [0.124] | 0.461** [0.185] | 0.987*** [0.199] |
| TUESDAY | n.a. | -0.030 [0.098] | -0.254** [0.105] | -0.037 [0.178] | 0.488** [0.201] |
| WEDNESDAY | n.a. | 0.189** [0.083] | -0.164 [0.103] | -0.530*** [0.174] | 0.308 [0.203] |
| THURSDAY | n.a. | 0.036 [0.093] | -0.040 [0.096] | -0.218 [0.175] | 0.401** [0.190] |
| FRIDAY | n.a. | 0.089 [0.096] | 0.055 [0.091] | -0.449** [0.193] | -0.083 [0.208] |
| SATURDAY | n.a. | 0.110 [0.097] | -0.023 [0.106] | -0.142 [0.184] | 0.266 [0.208] |
| CONSTANT | 1.933*** [0.083] | 0.009 [0.156] | 0.092 [0.173] | -0.496 [0.309] | -0.584** [0.315] |
| Observations | 949 | 949 | 949 | 949 | 949 |
| Alpha | 0.129 [0.011] | 0.000 [0.000] | 0.000 [0.011] | | |
| Chi2 | 110.55 | 113.44 | 99.21 | 63.50 | 72.76 |
| Pseudo-R ² | | | | 0.067 | 0.073 |

Notes: *, **, *** mean coefficient is significant at the 90%, 95% or 99% level. Robust standard errors are given in brackets. TA is short for transactions.

Table A2. Predicted number of events for winners and losers.

| | Winner | | | Loser | | |
|-------------------------------|----------|----------------|-------------------------|----------|----------------|-------------------------|
| | Estimate | Standard error | 95% Confidence interval | Estimate | Standard error | 95% Confidence interval |
| Total number of TA (count) | 10.278 | 0.219 | [9.850; 10.707] | 9.712 | 0.227 | [9.267; 10.158] |
| Number of TA on day 1 (count) | 1.637 | 0.052 | [1.534; 1.739] | 1.489 | 0.060 | [1.371; 1.606] |
| Number of TA on day 2 (count) | 1.493 | 0.052 | [1.391; 1.595] | 1.363 | 0.059 | [1.247; 1.479] |
| Empty diary on day 1 (dummy) | 0.159 | 0.016 | [0.128; 0.189] | 0.221 | 0.022 | [0.178; 0.264] |
| Empty diary on day 2 (dummy) | 0.194 | 0.017 | [0.161; 0.228] | 0.262 | 0.023 | [0.217; 0.307] |

Notes: Effects are estimated at the mean of other regressor variables. TA is the number of transactions.

Table A3. Results of estimations on the total amount spent and cash spent during the week.

| Variable | (I) | | (II) | | (III) | | (IV) | |
|-----------------------|------------------------------|-------------------|------------------------------|-------------------|----------------------------|-------------------|----------------------------|-------------------|
| | Amount spent during the week | Robust regression | Amount spent during the week | Median regression | Cash spent during the week | Robust regression | Cash spent during the week | Median regression |
| WINNER | 7.577 [8.545] | | 6.165 [11.267] | | 4.885 [5.093] | | 3.254 [6.525] | |
| AGE_25_34 | 35.377* [19.717] | | 40.220* [23.229] | | 22.607* [11.751] | | 27.625*** [10.308] | |
| AGE_35_44 | 44.650** [20.169] | | 53.603** [23.732] | | 30.675** [12.021] | | 39.167*** [10.420] | |
| AGE_45_54 | 65.811*** [19.602] | | 75.322*** [24.665] | | 43.954*** [11.683] | | 49.040*** [10.760] | |
| AGE_55_64 | 72.762*** [20.964] | | 90.105*** [29.363] | | 49.914*** [12.494] | | 57.896*** [13.322] | |
| AGE_65+ | 65.277** [25.480] | | 78.520** [32.524] | | 51.374*** [15.186] | | 63.344*** [20.820] | |
| FEMALE | 15.948* [8.931] | | 5.238 [10.968] | | 8.052 [5.323] | | 13.623** [6.754] | |
| HH_SIZE_2 | 23.566** [10.584] | | 37.773*** [14.456] | | 6.358 [6.308] | | 6.594 [8.274] | |
| HH_SIZE_3 | 22.915* [13.080] | | 18.883 [14.365] | | 0.355 [7.796] | | 4.181 [8.803] | |
| HH_SIZE_4+ | 63.653*** [13.212] | | 65.488*** [18.954] | | 13.985* [7.875] | | 15.518 [11.564] | |
| EDU_MEDIUM | 17.525* [10.461] | | 16.515 [11.609] | | 2.071 [6.235] | | 3.371 [7.866] | |
| EDU_HIGH | 34.942*** [13.200] | | 42.575** [18.413] | | 0.869 [7.867] | | 1.599 [9.562] | |
| EDU_UNI | 24.665** [14.879] | | 44.655* [23.233] | | - 11.994 [8.868] | | - 16.198 [14.304] | |
| INC_MID | 51.947*** [10.545] | | 65.548*** [14.344] | | 17.137*** [6.285] | | 18.461** [8.203] | |
| INC_HIGH | 60.137*** [19.076] | | 103.160*** [35.659] | | 20.552* [11.370] | | 29.417* [16.961] | |
| INC_DK | 6.001 [14.159] | | 4.950 [19.316] | | - 2.832 [8.439] | | - 9.530 [9.956] | |
| REGION_EAST | - 5.813 [10.623] | | - 7.188 [12.260] | | - 3.208 [6.332] | | - 11.125 [6.941] | |
| OCC_TRAIN | - 26.970 [26.068] | | - 17.163 [33.842] | | - 14.440 [15.536] | | - 14.494 [20.049] | |
| OCC_WORK | 17.757 [14.953] | | 21.403 [21.428] | | - 4.688 [8.912] | | - 1.848 [17.771] | |
| OCC_OTHER | - 41.508** [19.566] | | - 22.648 [21.224] | | - 19.830* [11.662] | | - 22.173 [20.399] | |
| CONSTANT | 72.786*** [24.876] | | 48.143 [30.731] | | 61.947* [14.826] | | 49.489** [20.091] | |
| Observations | 949 | | 949 | | 949 | | 949 | |
| Pseudo-R ² | | | 0.072 | | | | 0.042 | |

Notes: Median regression with bootstrapped standard errors (1,000 repetitions).

Table A4. Results of estimations on various transaction data features.

| Variable | (I) | | (II) | | (III) | | (IV) | | (V) |
|----------------|-------------------|--|-------------------|--|-------------------|--|------------------------|--|---------------------|
| | Share of cash TA | | Share of small TA | | Average TA amount | | Share of incomplete TA | | Share of rounded TA |
| | OLS | | OLS | | OLS | | OLS | | OLS |
| WINNER | 0.006 [0.015] | | 0.012 [0.014] | | 1.258 [2.189] | | -0.022 [0.023] | | 0.006 [0.014] |
| AGE_25_34 | 0.007 [0.034] | | -0.063** [0.030] | | 0.964 [3.639] | | -0.002 [0.049] | | -0.028 [0.034] |
| AGE_35_44 | 0.024 [0.034] | | -0.043 [0.031] | | 1.310 [2.915] | | -0.026 [0.050] | | -0.046 [0.034] |
| AGE_45_54 | 0.002 [0.034] | | -0.088** [0.031] | | 4.727 [3.250] | | 0.026 [0.049] | | -0.026 [0.033] |
| AGE_55_64 | 0.033 [0.036] | | -0.105** [0.033] | | 6.893 [4.307] | | 0.078 [0.054] | | -0.040 [0.035] |
| AGE_65+ | 0.042 [0.044] | | -0.096** [0.039] | | 4.340 [5.255] | | 0.108* [0.065] | | -0.020 [0.041] |
| FEMALE | 0.026* [0.016] | | -0.016 [0.015] | | -6.283** [3.164] | | -0.001 [0.023] | | -0.035** [0.015] |
| HH_SIZE_2 | -0.050*** [0.018] | | -0.029* [0.017] | | 8.825** [4.270] | | 0.024 [0.028] | | 0.021 [0.017] |
| HH_SIZE_3 | -0.043* [0.023] | | 0.006 [0.022] | | 4.171** [1.826] | | 0.033 [0.034] | | 0.044** [0.022] |
| HH_SIZE_4+ | -0.078*** [0.023] | | -0.062*** [0.020] | | 8.172*** [2.724] | | 0.091*** [0.035] | | 0.079*** [0.022] |
| EDU_MEDIUM | -0.020 [0.017] | | 0.032* [0.017] | | 2.503 [1.716] | | -0.080*** [0.029] | | -0.001 [0.018] |
| EDU_HIGH | -0.071*** [0.024] | | 0.027 [0.023] | | 3.760 [5.173] | | -0.108*** [0.034] | | 0.007 [0.022] |
| EDU_UNI | -0.076*** [0.027] | | 0.044* [0.022] | | 1.710 [3.331] | | -0.118*** [0.039] | | -0.030 [0.022] |
| INC_MID | -0.020 [0.019] | | -0.037** [0.017] | | 2.941 [2.185] | | 0.015 [0.027] | | 0.033** [0.017] |
| INC_HIGH | -0.012 [0.033] | | 0.002 [0.032] | | 2.216 [3.473] | | 0.005 [0.052] | | 0.003 [0.030] |
| INC_DK | -0.037 [0.026] | | 0.021 [0.023] | | 4.410 [3.461] | | 0.010 [0.038] | | -0.011 [0.023] |
| REGION_EAST | -0.001 [0.018] | | 0.027 [0.017] | | -2.783 [3.091] | | -0.115*** [0.025] | | -0.016 [0.017] |
| OCC_TRAIN | -0.014 [0.045] | | 0.065 [0.045] | | 6.136 [13.135] | | -0.008 [0.065] | | 0.050 [0.042] |
| OCC_WORK | -0.036 [0.026] | | -0.001 [0.021] | | -1.923 [3.351] | | 0.046 [0.039] | | 0.023 [0.022] |
| OCC_OTHER | 0.050 [0.033] | | 0.045 [0.033] | | -7.881** [3.753] | | 0.116** [0.054] | | -0.004 [0.035] |
| CONSTANT | 0.831*** [0.042] | | 0.280*** [0.038] | | 20.154*** [4.926] | | 0.347*** [0.063] | | 0.284*** [0.041] |
| Observations | 949 | | 949 | | 949 | | 949 | | 949 |
| R ² | 0.081 | | 0.077 | | 0.030 | | 0.066 | | 0.049 |

Notes: *, **, *** mean coefficient is significant at the 90%, 95% or 99% level. Robust standard errors are given in brackets. TA is short for transactions.

Table A5. Construction of regression variables.

| Variable name | Type | Description |
|---------------|-------|--|
| WINNER | Dummy | Outcome of roll of the die. One, if person wins EUR 20. |
| AGE_18_24 | Dummy | One, if respondent is aged 18 to 24. Reference category. |
| AGE_25_34 | Dummy | One, if respondent is aged 25 to 34. |
| AGE_35_44 | Dummy | One, if respondent is aged 35 to 44. |
| AGE_45_54 | Dummy | One, if respondent is aged 45 to 54. |
| AGE_55_64 | Dummy | One, if respondent is aged 55 to 64. |
| AGE_65+ | Dummy | One, if respondent is aged 65 or above. |
| FEMALE | Dummy | Gender of respondent. One, if gender is female. |
| HH_SIZE_1 | Dummy | Number of persons living in respondent's household (including children). One, if household size is one. Reference category. |
| HH_SIZE_2 | Dummy | Number of persons living in respondent's household (including children). One, if household size is two. |
| HH_SIZE_3 | Dummy | Number of persons living in respondent's household (including children). One, if household size is three. |
| HH_SIZE_4+ | Dummy | Number of persons living in respondent's household (including children). One, if household size is four or more. |
| EDU_LOW | Dummy | Educational attainment of respondent. One, if education is low (no educational degree (yet), lower secondary education of less than 10 years) or if education is "other/don't know". Reference category. |
| EDU_MEDIUM | Dummy | Educational attainment of respondent. One, if respondent has secondary education of at least 10 years. |
| EDU_HIGH | Dummy | Educational attainment of respondent. One, if respondent has upper secondary education (= qualification for entering university). |
| EDU_UNI | Dummy | Educational attainment of respondent. One, if respondent has university degree (including university of applied sciences). |
| INC_LOW | Dummy | Respondent's personal monthly net income. One, if income is less than EUR 1,500. Reference category. |
| INC_MID | Dummy | Respondent's personal monthly net income. One, if income is between EUR 1,500 and EUR 3,000. |
| INC_HIGH | Dummy | Respondent's personal monthly net income. One, if income is more than EUR 3,000. |
| INC_DK | Dummy | Respondent's personal monthly net income. One, if "don't know/no answer". |
| REGION_EAST | Dummy | Respondent's region of residence. One, if region is East Germany. |

Table A5. Continued

| Variable name | Type | Description |
|---------------|-------|--|
| OCC_HOME | Dummy | Respondent's current occupation. One, if respondent is not working or working at home (pensioner, homemaker). Reference category. |
| OCC_TRAIN | Dummy | Respondent's current occupation. One, if respondent is in training (student, apprentice, volunteer in federal volunteer service ("Bundesfreiwilligendienst")). |
| OCC_WORK | Dummy | Respondent's current occupation. One, if respondent is working outside the home (employee, public servant, self-employed person). |
| OCC_OTHER | Dummy | Respondent's current occupation. One, if unemployed or "other/don't know". |
| SUNDAY | Dummy | One, if transaction takes place on Sunday. |
| MONDAY | Dummy | One, if transaction takes place on Monday. Reference category. |
| TUESDAY | Dummy | One, if transaction takes place on Tuesday. |
| WEDNESDAY | Dummy | One, if transaction takes place on Wednesday. |
| THURSDAY | Dummy | One, if transaction takes place on Thursday. |
| FRIDAY | Dummy | One, if transaction takes place on Friday. |
| SATURDAY | Dummy | One, if transaction takes place on Saturday. |

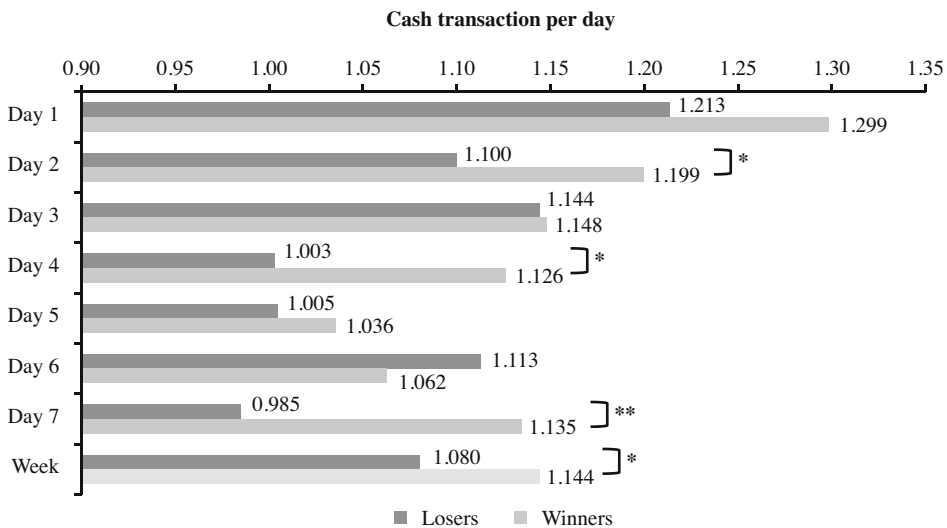


Fig. A1. Number of cash transactions per person per diary day. Unauthenticated
Download Date | 1/22/18 1:23 PM

6. References

- Awel, Y.M. and T.T. Azamhou. 2015. "Risk Preference or Financial Literacy? Behavioural Experiment on Index Insurance Demand." UNU-MERIT Working Papers 2015-005, Maastricht. Available at: <http://www.merit.unu.edu/publications/working-papers/abstract/?id=5626> (accessed August 2016).
- Axhausen, K.W., A. Zimmermann, S. Schönfelder, G. Rindsfuser, and T. Haupt. 2002. "Observing the Rhythms of Daily Life: A Six-Week Travel Diary." *Transportation* 29(2): 95–124. Doi: <http://dx.doi.org/10.1023/A:1014247822322>.
- Bagnall, J., D. Bounie, K.P. Huynh, A. Kosse, T. Schmidt, S.D. Schuh, and H. Stix. 2016. "Consumer Cash Usage: A Cross-Country Comparison with Payment Diary Survey Data." *International Journal of Central Banking* 12(4): 1–62. Available at: <http://www.ijcb.org/journal/ijcb16q4a1.htm> (accessed December 2016).
- Biemer, P.P. and L.E. Lyberg. 2003. *Introduction to Survey Quality*. New Jersey: John Wiley & Sons.
- Bonke, J. and P. Fallesen. 2010. "The Impact of Incentives and Interview Methods on Response Quantity and Quality in Diary- and Booklet-Based Surveys." *Survey Research Methods* 4(2): 91–101. Doi: <http://dx.doi.org/10.18148/srm/2010.v4i2.3614>.
- Borghans, L., J.J. Heckman, B.H. Golsteyn, and H. Meijers. 2009. "Gender Differences in Risk Aversion and Ambiguity Aversion." *Journal of the European Economic Association* 7(2–3): 649–658. Doi: <http://dx.doi.org/10.1007/s11238-016-9565-9>.
- Charness, G. and U. Gneezy. 2010. "Portfolio Choice and Risk Attitudes: An Experiment." *Economic Inquiry* 48(1): 133–146. Doi: <http://dx.doi.org/10.1111/j.1465-7295.2009.00219.x>.
- Charness, G., U. Gneezy, and A. Imas. 2013. "Experimental Methods: Eliciting Risk Preferences." *Journal of Economic Behavior & Organization* 87: 43–51. Doi: <http://dx.doi.org/10.1016/j.jebo.2012.12.023>.
- Davern, M., T.H. Rockwood, R. Sherrod, and S. Campbell. 2003. "Prepaid Monetary Incentives and Data Quality in Face-to-Face Interviews: Data from the 1996 Survey of Income and Program Participation Incentive Experiment." *Public Opinion Quarterly* 67(1): 139–147. Doi: <http://dx.doi.org/10.1086/346012>.
- Deutsche Bundesbank. 2015. *Payment Behaviour in Germany in 2014. Third Study on the Utilisation of Cash and Cashless Payment Instruments*. Frankfurt am Main: Deutsche Bundesbank. Available at: http://www.bundesbank.de/Redaktion/EN/Downloads/Publications/Studies/payment_behaviour_in_germany_in_2014.pdf?__blob=publicationFile (accessed December 2015).
- Dohmen, T., A. Falk, D. Huffman, and U. Sunde. 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" *The American Economic Review* 100(3): 1238–1260. Doi: <http://dx.doi.org/10.1257/aer.100.3.1238>.
- Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G.G. Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences." *Journal of the European Economic Association* 9(3): 522–550. Doi: <http://dx.doi.org/10.1111/j.1542-4774.2011.01015.x>.

- Eckel, C.C. and P.J. Grossman. 2002. "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk." *Evolution and Human Behavior* 23(4): 281–295. Doi: [http://dx.doi.org/10.1016/S1090-5138\(02\)00097-1](http://dx.doi.org/10.1016/S1090-5138(02)00097-1).
- Eckel, C.C. and P.J. Grossman. 2008. "Men, Women and Risk Aversion: Experimental Evidence." In *Handbook of Experimental Economics Results* 1(113), edited by C.R. Plott and V.L. Smith. 1061–1073. New York: Elsevier.
- Fricker, S. and R. Tourangeau. 2010. "Examining the Relationship Between Nonresponse Propensity and Data Quality in Two National Household Surveys." *Public Opinion Quarterly* 74(5): 934–955. Doi: <http://dx.doi.org/10.1093/poq/nfq064>.
- Godwin, R.K. 1979. "The Consequences of Large Monetary Incentives in Mail Surveys of Elites." *Public Opinion Quarterly* 43(3): 378–387. Doi: <http://dx.doi.org/10.1086/268528>.
- Goetz, E.G., T.R. Tyler, and F.L. Cook. 1984. "Promised Incentives in Media Research: A Look at Data Quality, Sample Representativeness and Response Rates." *Journal of Marketing Research* 21(2): 148–154. Doi: <http://dx.doi.org/10.2307/3151697>.
- Goldenberg, K., D. McGrath, and L. Tan. 2009. "The Effects of Incentives on the Consumer Expenditure Interview Survey." Paper presented at the 64th Annual Conference of the American Association for Public Opinion Research, May 14 – 17, 2009, Hollywood, Florida. Available at: <https://stats.bls.gov/osmr/pdf/st090100.pdf> (accessed December 2015).
- Goldenberg, K. and J. Ryan. 2009. "Evolution and Change in the Consumer Expenditure Surveys: Adapting Methodologies to Meet Changing Needs." Paper presented at the NBER Summer Institute 2009 – Conference on Research on Income and Wealth, July 13, 2009, Cambridge, Massachusetts. Available at: https://www.bls.gov/cex/research_papers/pdf/Evolution-and-Change-in-the-Consumer-Expenditure-Surveys-Adapting-Methodologies-to-Meet-Changing-Needs.pdf (accessed December 2015).
- Görizt, A.S. 2004. "The Impact of Material Incentives on Response Quantity, Response Quality, Sample Composition, Survey Outcome, and Cost in Online Access Panels." *International Journal of Market Research* 46(3): 327–345. Available at: https://www.mrs.org.uk/ijmr_article/article/79180 (accessed January 2016).
- Görizt, A.S. 2005. "Contingent versus Unconditional Incentives in WWW-Studies." *Metodolosky Zvezki* 2(1): 1–14. Available at: <http://mrvar.fdv.uni-lj.si/pub/mz/mz2.1/goritz.pdf> (accessed January 2016).
- Groves, R.M., R.B. Cialdini and M.P. Couper. 1992. "Understanding the Decision to Participate in a Survey." *Public Opinion Quarterly* 56(4): 475–495. Doi: <http://dx.doi.org/10.1086/269338>.
- Groves, R.M. 2004. *Survey Errors and Survey Costs*. New Jersey: John Wiley & Sons.
- Groves, R.M. and L. Lyberg. 2010. "Total Survey Error: Past, Present, and Future." *Public Opinion Quarterly* 74(5): 849–879. Doi: <http://dx.doi.org/10.1093/poq/nfq065>.
- Halek, M. and J.G. Eisenhauer. 2001. "Demography of Risk Aversion." *Journal of Risk and Insurance* 68(1): 1–24. Doi: <http://dx.doi.org/10.2307/2678130>.
- Hedlin, D., T. Dale, G. Haraldsen, and J. Jones (eds.). 2005. "Developing Methods for Assessing Perceived Response Burden. Research Report." Stockholm: Statistics Sweden, Oslo: Statistics Norway, and London: Office for National Statistics. Available at: <http://ec.europa.eu/eurostat/documents/64157/4374310/10-DEVELOPING-METHODS-FOR>

- [ASSESSING-PERCEIVED-RESPONSE-BURDEN.pdf/1900efc8-1a07-4482-b3c9-be88ee71df3b](#) (accessed November 2015).
- Hoffmeyer-Zlotnik, J.H. 2003. "New Sampling Designs and the Quality of Data." In *Developments in Applied Statistics*. Metodoloski zvezki 19, edited by A. Ferligoj and A. Mrvar. 205–217. Ljubljana: FDV.
- James, J.M. and R. Bolstein. 1990. "The Effect of Monetary Incentives and Follow-Up Mailings on the Response Rate and Response Quality in Mail Surveys." *Public Opinion Quarterly* 54(3): 346–361. Doi: <http://dx.doi.org/10.1086/269211>.
- Jones, J. 2012. "Response Burden: Introductory Overview Lecture." Paper presented at the 4th International Establishment Surveys Conference, June 11–14, 2012, Montreal, Canada. Available at: <https://ww2.amstat.org/meetings/ices/2012/papers/302289.pdf> (accessed July 2016).
- Jonker, N. and A. Kosse. 2013. "Estimating Cash Usage: The Impact of Survey Design on Research Outcomes." *De Economist* 161(1): 19–44. Doi: <http://dx.doi.org/10.1007/s10645-012-9200-2>.
- Laurie, H. and P. Lynn. 2009. "The Use of Respondent Incentives on Longitudinal Surveys." In *Methodology of Longitudinal Surveys*, edited by P. Lynn, 205–234. Chichester, UK: John Wiley & Sons.
- Schwarz, N. and G.L. Clore. 1996. "Feelings and Phenomenal Experiences." In *Social Psychology: Handbook of Basic Principles*, edited by E.T. Higgins and A. Kruglanski, 433–465. New York: Guilford.
- Sharp, L.M. and J. Frankel. 1983. "Respondent Burden: A Test of Some Common Assumptions." *Public Opinion Quarterly* 47(1): 36–53. Doi: <http://dx.doi.org/10.1086/268765>.
- Shettle, C. and G. Mooney. 1999. "Monetary Incentives in US Government Surveys." *Journal of Official Statistics* 15(2): 231–250. Available at: <http://www.jos.nu/Articles/abstract.asp?article=152231> (accessed January 2016).
- Shuttleworth, F.K. 1931. "A Study of Questionnaire Technique." *Journal of Educational Psychology* 22(9): 652–658. Doi: <http://dx.doi.org/10.1037/h0074591>.
- Singer, E., J. Van Hoewyk, and M. Maher. 2000. "Experiments with Incentives in Telephone Surveys." *Public Opinion Quarterly* 64(2): 171–188. Doi: <http://dx.doi.org/10.1086/317761>.
- Tzamourani, P. and P. Lynn. 1999. "The Effect of Monetary Incentives on Data Quality – Results from the Social Attitudes Survey 1998 experiment." CREST Working Paper 73, Oxford, UK. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.573.5965&rep=rep1&type=pdf> (accessed January 2016).
- Weisberg, H.F. 2005. *The Total Survey Error Approach. A Guide To The New Science Of Survey Research*. Chicago/London: The University of Chicago Press.
- Willimack, D.K., H. Schuman, B.E. Pennell, and J.M. Lepkowski. 1995. "Effects of a Prepaid Nonmonetary Incentive on Response Rates and Response Quality in a Face-to-Face Survey." *Public Opinion Quarterly* 59(1): 78–92. Doi: <http://dx.doi.org/10.1086/269459>.

Received January 2016

Revised January 2017

Accepted January 2017