methoden, daten, analysen · Jg. 7(1), S. 89-122 DOI: 10.12758/mda.2013.004

Respondent Incentives in a National Face-to-Face Survey

Effects on Outcome Rates, Sample Composition and Fieldwork Efforts

Der Einsatz von Befragten-Incentives in einer bundesweiten face-to-face-Umfrage

Effekte auf Ausschöpfung, Stichprobenzusammensetzung und Feldarbeitsaufwand

Michael Blohm und Achim Koch

Abstract

Nonresponse is an ongoing challenge for survey research. In the German General Social Survey (ALLBUS) 2010, an experiment was set up to test the effect of respondent incentives on outcome rates, sample composition and fieldwork efforts. A random subsample of target persons was offered a monetary incentive of €10 to be paid upon completion of the interview. The other part of the sample acted as a control group receiving no incentive. The incentive used in ALLBUS 2010 led to an increase in the response rate, mainly by improving the cooperation rate. It did not change the sample composition in a major way. Concerning fieldwork efforts, a slight reduction was observed: In the incentive condition, a given number of interviews was achieved with a lower number of contact attempts than in the no incentive condition.

Zusammenfassung

Nonresponse ist eine Herausforderung für die gesamte Umfrageforschung. In der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) wurde im Jahr 2010 ein Experiment durchgeführt, um die Effekte von Befragten-Incentives auf die Ausschöpfungsquote, die Stichprobenzusammensetzung und den in der Feldarbeit notwendigen Aufwand zu untersuchen. Einer Zufallsstichprobe der Zielpersonen wurde für die Teilnahme an der Umfrage ein monetäres Incentive in Höhe von 10 € angeboten. Der verbleibende Teil der Zielpersonen fungierte als Kontrollgruppe und erhielt kein Incentive. Das verwendete Incentive führte zu einer höheren Kooperationsrate und einer höheren Ausschöpfung in der Experimentalgruppe. Die Stichprobenzusammensetzung unterschied sich nicht wesentlich zwischen Experimental- und

Kontrollgruppe. In der Experimentalgruppe konnten die Interviewer eine gegebene Zahl von Interviews mit weniger Kontaktversuchen erzielen als in der Kontrollgruppe.

Keywords: Incentive, nonresponse, bias, sample composition, fieldwork effort

1 Introduction[1]

Declining response rates are a continuing problem for household surveys in many Western countries (Atrostic, Bates, Burt, & Silberstein, 2001; de Leeuw & de Heer, 2002; Dixon & Tucker, 2010; Kohut, Keeter, Doherty, Dimock, & Christian, 2012). As a consequence, low response rates are no longer exceptional in household surveys in Europe. For instance, in the fifth round of the European Social Survey (ESS), which was fielded in 2010, 4 out of the 27 participating countries had a response rate below 50%.[2] In the fourth wave of the European Values Study (EVS), fielded in 2008, 11 out of 47 countries/regions had a response rate of 50% or less.[3] In the third European Quality of Life Survey (EQLS), fielded in 2011, 17 out of 27 countries had a maximum response rate of 50%, among which 8 countries with a response rate even below 40% (Eurofound, 2012).

Various attempts have been made to increase response rates or at least to halt downward trends in response rates. These include, for instance, an increased number of contact attempts, the use of advance letters or the provision of incentives to sample persons to encourage survey participation (Groves et al., 2004). The use of respondent incentives in order to increase response has a long tradition in mail surveys. More recently, however, the use of incentives has also become more common in face-to-face surveys (Kulka, Eyerman, & McNeeley, 2005; Singer, 2002). The ESS is a case in point. In this repeated cross-national survey, the number of countries using some kind of a respondent incentive has been continuously increasing over the first five survey rounds. Whereas in the first round, which was fielded in the years 2002/2003, only 7 out of the 16 permanent countries used a respondent incentive, this number rose to 8, 10, 11 and 13 countries, respectively, in rounds 2 to 5.[4]

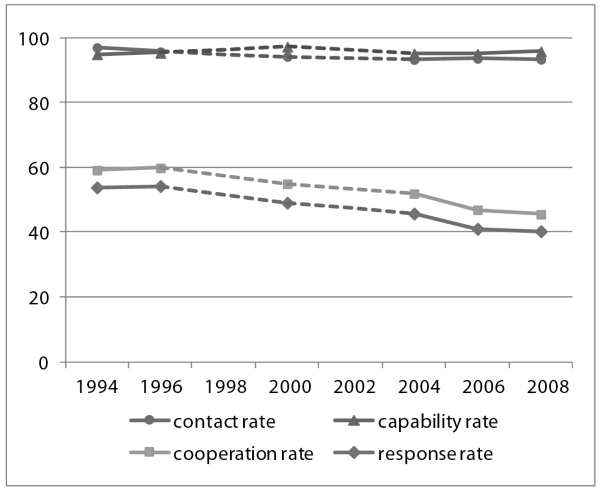
1 The authors would like to thank Dorothée Behr, Henning Best, Christof Wolf and the reviewers for providing constructive comments on a former draft of the paper.

2 Information from ESS website: http://ess.nsd.uib.no/ess/round5/deviations.html (03.04.2013).

3 Personal communication from Evelyn Brislinger, GESIS (28.03.2013).

4 Whether or not the use of an incentive actually increased the response rate in the relevant ESS countries (and, if yes, to what extent), unfortunately remains an unanswerable question, since apart from one exception (Phelps, 2008) the effects of introducing

Figure 1 Contact, capability, cooperation and response rates in ALLBUS 1994 – 2008 (in %)



The German General Social Survey (ALLBUS) has also faced an increase in nonresponse in the past years. The ALLBUS is a biennial face-to-face survey of the adult population, covering a wide range of topics and aiming at charting the long-term trends in attitudes and behaviour in Germany (http://www.gesis.org/en/allbus). Between 1994 and 2008 the response rate of ALLBUS decreased from 54% to 40%. The main reason for this decline was a drop in the cooperation rate (see Figure 1).[5] In contrast, the contact rate remained broadly stable on a high level. Own analyses of data from the ALLBUS contact forms show that the latter result can be attributed to an increase in the number of contact attempts over time.

In this context, ALLBUS 2010 included an experiment that was set up to investigate whether or not the provision of a respondent incentive might help to counteract the downward trend in the response rate. Using a respondent incentive seemed to be a promising measure insofar as an incentive will primarily help

Footnote continued from previous page

incentives were not measured in the ESS (or at least, the results of relevant experiments were not (yet) published).

5 For a definition of the outcome rates, see Appendix 1. In order to ensure comparability over time, ALLBUS 1998 and 2002 were excluded, since in both years the survey was fielded by a different survey organisation.

to reduce refusals (Singer, 2002).[6] The experiment in ALLBUS 2010 was designed to help to clarify three issues. First, the experiment should provide an answer to the question whether respondent incentives increase cooperation and response rates in a face-to-face survey like ALLBUS. Second, the experiment should provide evidence on whether respondent incentives affect sample composition. Third, the experiment should give some indication of whether the use of a respondent incentive affects fieldwork efforts and survey costs.[7]

Thus far, the literature on incentive effects is less comprehensive with regard to face-to-face surveys than with regard to mail surveys (see section 2). There is evidence that incentives help to increase response rates also in face-to-face surveys. However, there is less (and often only mixed) evidence as regards the effects of incentives on sample composition and survey costs. This is unfortunate since in the past decade a shift in the focus of nonresponse research has occurred (Singer, 2006). It is increasingly acknowledged that a concentration on response rates is inadequate; the focus has to be directed towards nonresponse bias. Regarding the use of incentives the crucial question is whether incentives reinforce or counteract tendencies to underrepresent certain subgroups of the population. Respondent incentives can help to decrease nonresponse bias if they disproportionately attract target persons whose response propensity is below average otherwise.[8] Furthermore, the cost implications of various survey design features (like the number of contact attempts or refusal conversion activities) have gained more attention in recent years. Regarding respondent incentives the question is whether the obvious increase in costs associated with their use will (at least partly) be absorbed by a reduction in fieldwork efforts necessary to complete the interviews.

In the ALLBUS 2010 experiment, we wanted to test an incentive appropriate for the use in future survey rounds – given the specific design of the ALLBUS and the available budget. It was decided to investigate the effect of a €10 conditional cash incentive. A random subsample of the ALLBUS 2010 respondents was to receive a conditional monetary incentive of €10, and the rest of respondents was to act as a control group receiving no incentive. In the survey literature (see

6 Up to that point, ALLBUS had used a respondent incentive only once. In 2002, all ALLBUS respondents received a commemorative coin worth €10. At that time, no efforts were made to measure the effect of the incentive experimentally.

7 Note that in the present paper we are not concerned with the potential effects incentives may have on measurement, i.e., whether or not they change the response behaviour of respondents.

8 As far as the ALLBUS is concerned, for instance, an issue is whether an incentive might help to reduce the underrepresentation of lower educated people that had been observed in previous rounds of the survey (Wasmer, Scholz, & Blohm, 2010).

section 2), unconditional ('prepaid') incentives are often described as being preferable to conditional ('promised') incentives. In survey practice, however, conditional incentives play an important role in face-to-face surveys. In the fifth round of the ESS, for instance, six countries offered a conditional cash incentive or a shopping voucher to respondents, whereas only two countries provided an unconditional cash incentive to all target persons. Likewise, in the EVS 2008, eight countries used a conditional cash incentive and only one country an unconditional cash incentive.

Several reasons led us to the decision to use a conditional cash incentive in the ALLBUS experiment. First, in face-to-face surveys (in contrast to mail and telephone surveys) an interviewer is present who can pass the money to the respondent immediately after the interview has been completed. The presence of an interviewer and the short time span between the conduct of an interview and the handover of the incentive should induce respondents to trust that a promised incentive will actually be delivered to them (Singer, 2002). Second, the costs of a conditional incentive will be lower than the costs of an unconditional incentive of the same value, particularly if the expected response rate is rather low as it is in ALLBUS. Third, for government-financed surveys like ALLBUS, the use of a conditional incentive is easier to justify, both vis-à-vis the funders and the target persons of the survey. In contrast, the use of an unconditional incentive is more difficult to justify, as some people might see such an incentive as a waste of tax money.[9]

Implementing the experiment in the regular ALLBUS survey provides a solid base for the decision as to whether or not such an incentive should be used in future survey rounds of ALLBUS. In addition, the question is whether the present results will also be relevant for other surveys. Of course, the effect of an incentive (on response rates, sample composition and fieldwork efforts) can vary, depending on the concrete conditions in a survey, like the topic and target population of the survey, the survey mode, the burden imposed on respondents, the implemented call schedule and the refusal conversions efforts being made, etc.. According to Groves (2008), it has to be noted "... that the marginal effects of incentives are a function

9 In 2010, the German part of the Survey of Health, Ageing and Retirement in Europe (SHARE) fielded a large-scale experiment using unconditional monetary incentives worth €10, €20 and €40 (Börsch-Supan, Krieger, & Schröder, 2013). In each incentive condition, 6-7% of the target persons called the telephone hotline (compared to 1.5% of the target persons in the condition with no incentive). Most of them were unhappy with the inclusion of cash in the advance letter and questioned the legitimacy of the survey. The researchers responsible for the experiment also point out: "One complaint to a member of the Bundestag reached high levels at the Federal Ministry of Education and Research, the funders of SHARE in Germany." (Börsch-Supan et al., 2013: 11)

of the base protocol to which they are being compared." This limits the generalisability of the results of any single incentive experiment.[10] As a consequence, we have to acknowledge that the results of the present experiment will be primarily relevant for face-to-face surveys similar to ALLBUS. ALLBUS represents a certain type of survey, namely cross-sectional high-quality face-to-face surveys of the general population covering a broad range of topics. This type of survey continues to play an important role in social sciences all over the world, both as national general social surveys (Smith, Kim, Koch, & Park, 2006) or as cross-national surveys like the ESS or the EVS.

In the following, we first summarise the previous research on incentives in face-to-face surveys and formulate our working hypotheses. We then describe the design and implementation of the ALLBUS 2010 incentive experiment. In the results section, the effect of the incentive on contact, capability, cooperation and response rates is analysed. In addition, we investigate whether the incentive differentially attracted certain demographic subgroups and whether the response distributions and survey results for selected variables differ between respondents in the incentive and no incentive condition. Finally, we touch upon the cost issues related to incentives and examine whether the incentive used reduced the amount of interviewer effort required to complete a case. We discuss our results against the backdrop of possible practical consequences for future surveys like ALLBUS.

2 Previous research and working hypotheses

Incentives have been used in mail surveys for a long time, and their positive effect on response rates is well documented (Church, 1993). Regarding telephone and face-to-face surveys one could expect incentives to be of lower importance than in self-administered mail surveys. In interviewer-administered surveys, the interviewer plays a vital role in persuading people to participate in the survey. This limits the potential effect of an incentive. Singer, van Hoewyk, Gebler, Raghunathan, and McGonagle (1999; see also Singer, 2002) summarise what is known from the experimental literature about the effects of respondent incentives in interviewer-mediated surveys. Their meta-analysis is based on 39 experiments conducted in telephone and face-to-face surveys in the United States and Canada. In the following, we describe their major findings and add results from several other face-

10 In order to achieve a solid base for the decision to use an incentive in a particular study, Singer and Kulka (2001) recommend to always pretest the specific incentive plan with the target population and the survey methods to be employed.

to-face studies that were either published only in recent years or conducted in other countries than the United States or Canada.

Singer et al. (1999) found that paying an incentive is an effective tool for increasing the response rate in face-to-face and telephone surveys. However, the effects were smaller than in mail surveys. On average, each dollar of incentive paid increased the response rate by about one third of a percentage point. Gifts were less effective in raising cooperation than money (even when the value of the incentive was controlled for). Both prepaid and promised incentives produced a significant increase in response rates (with prepaid incentives yielding higher response rates than promised incentives). Incentives were effective both in low and high burden surveys. The effect of incentives was larger in studies in which the response rate without an incentive was low. Incentives seem to work primarily by reducing refusals; there is no evidence that incentives are helpful in making contact with respondents (Singer, 2002).[11]

The results of several other face-to-face studies are more or less in line with Singer's main conclusions. Both promised incentives (Eyerman, Bowman, Butler, & Wright, 2005; Lynn, Thomson, & Brook, 1998; Lynn, 2001; Nicolaas, 2004) and prepaid incentives (Börsch-Supan et al., 2013; Wetzels, Schmeets, van den Brakel, & Feskens, 2008) increase response rates, mainly by boosting the level of cooperation. Direct comparisons of the effects of unconditional and conditional incentives on response rates are rare. Only a few experiments were conducted outside the United States and Canada, and the evidence they provide is inconclusive. Scherpenzeel and Toepoel (2012) found that unconditional incentives worked better than conditional incentives in the recruitment for a new online household panel. In this study, contact was made by either telephone or face-to-face and then led to a short recruitment interview (CATI or CAPI, respectively). Contrary to this result, neither a prepaid nor a promised incentive raised the cooperation rate in the first wave of a face-to-face panel survey in Germany (Castiglioni, Pforr, & Krieger, 2008). However, having kept the incentive conditions the same in waves 2 and 3, it turned out that the cumulative response rate of the panel was lowest for the no incentive group and highest for the conditional incentive group. Similarly, Phelps (2008) found in an incentive experiment implemented in the United Kingdom in round 3 of the ESS that a conditional cash incentive (£10) led to a higher cooperation and response rate than an unconditional incentive in the form of stamps (worth £6).

11 Barón, Breunig, Cobb-Clark, Gørgens, and Sartbayeva (2009), however, report for their study that incentives improved the probability of making contact with individuals with heavy exposure to the income support system.

Regarding the effects of incentives on sample composition, the available information is limited and the results are mixed (Singer et al., 1999; Singer, 2002). Whereas in a number of studies no differences in sample composition between the incentive and the no incentive condition could be observed, there were also a few studies that found that incentives encouraged participation particularly among the groups who would otherwise be underrepresented in the survey, like low education, low income or minority respondents. All these groups have a lower socio-economic status, and it can be assumed that the perceived value of a given incentive will be higher for them than for groups who are better off. In addition, it was found that incentives can act as a motive to participate in a survey for groups in which other stimulating factors like interest in the survey topic or a sense of civic obligation are absent. At the same time, there was no indication that incentives reduce the motivation to participate among those groups who are more interested in the survey topic or more altruistic. In the Detroit Area Study, for instance, Groves, Singer, and Corning (2000) found that the effect of an incentive on the participation rate in a follow-up mail survey was higher for persons scoring low on an index of community/political involvement than for persons scoring high on the same index.[12] The results concerning differences for socio-economic groups as well as differences for persons with different levels of community and political involvement can be seen in accordance with leverage-saliency theory of survey participation (Groves et al., 2000). In both cases, we can assume that the leverage of incentives (i.e., the importance of incentives for the decision to participate) is higher for one group (with lower socio-economic status and lower civic involvement, respectively) than for its counterpart. Whether this actually translates into a difference in respondent cooperation will depend on the leverage of other survey features (like the survey topic, sponsor, etc.) and how salient the different survey features were made to the target persons during the presentation of the survey request.

More recent studies continue to provide mixed evidence regarding sample composition effects. Eyerman et al. (2005) report no effect of incentives on the demographic structure of their samples. In a slightly different analysis of the same data, Wright, Bowman, Butler, and Eyerman (2005) observed a larger effect of incentives for younger age groups and also an interaction effect with population density. Wetzels et al. (2008) found hardly any impact of incentives on response distributions of demographic variables (including age and household income). In a telephone survey, Barón et al. (2009) observed an equalising effect of incentives

12 It has repeatedly been shown for the United States that people doing volunteer work or, more generally, showing larger political and social engagement, are more willing to agree to participate in a survey (Abraham, Helms, & Presser, 2009; Kohut et al., 2012).

on response rates across different socio-economic groups. In a meta-analysis of studies, including self-administered, face-to-face and telephone surveys, Groves and Peytcheva (2008) found no evidence that incentives reduced the nonresponse bias in the statistics of interest.

Regarding fieldwork efforts and survey costs, Singer (2002) reports a number of examples in which incentives were cost-effective measures not only in mail but also in face-to-face surveys. For instance, in a field experiment in the National Adult Literacy Survey in the United States, the average costs per interview were lower when a $20 incentive was used than in the no incentive condition (Berlin et al., 1992). Likewise, Scherpenzeel and Toepoel (2012) estimate that the costs per registered household for an online panel survey were lowest when a prepaid €10 incentive was used (compared both to the no incentive condition, and to all other conditions of prepaid/promised incentives). Several other face-to-face surveys not included in the summary of Singer looked at the mean number of interviewer visits that were required either to complete an interview or to achieve a final disposition code for all eligible addresses (Eyerman et al., 2005; Lynn et al., 1998; Nicolaas & Stratford, 2005). In each case it was found that less effort was required by interviewers to elicit respondent cooperation when an incentive was provided. However, as Singer (2002) points out, one has to be aware of the fact that the cost-effectiveness of an incentive will vary depending on the specific conditions in a survey.

All results taken together, we can conclude that the findings of field experiments on incentives in face-to-face surveys show that incentives help to increase response rates in face-to-face surveys. The increase in response is mainly achieved by increasing the cooperation rate. Regarding the effects of incentives on sample composition the evidence is mixed. Whereas some analyses found that incentives particularly attract certain subgroups of the population (like younger people, people with a lower socio-economic status or people with a lower community involvement), other analyses did not.[13] Finally, the literature indicates that providing an

13 Finding out more about the reasons for these divergent results requires a closer look at the different surveys, the experiments and the incentives used. For example, it seems plausible that the higher effect of incentives on young people, which was observed in the study of Wright et al. (2005), might be a result of the fact that a $20 or $40 cash incentive is particularly attractive to young people. In contrast, in the study from Wetzels et al. (2008) no differential incentive effect for age was found. It is easily comprehensible that offering a few stamps (worth €2, €4 or €8), as it was done in this study, is not an incentive that particularly attracts young people. A profound review of incentive effects in face-to-face surveys would have to take into account all potential survey features which might interact with the way incentives work. Given the large number of these potential features and the small number of available empirical studies, such a review does not seem promising for the time being.

incentive in face-to-face surveys can help to reduce fieldwork efforts by reducing the number of contact attempts necessary to complete an interview.

Considering the main results of the relevant literature, we set up the following working hypotheses for the incentive experiment implemented in ALLBUS 2010:

Effects of incentives on outcome rates

H1. A higher cooperation and response rate will be achieved in the incentive condition than in the no incentive condition. We do not expect a difference in the contact and capability rate.

Effects of incentives on sample composition

H2. The sample of completed interviews will contain more people with lower socioeconomic status in the incentive condition than in the no incentive condition.

H3. The sample of completed interviews will contain more people with low involvement or low interest in community, society and politics in the incentive condition than in the no incentive condition.

Effects of incentives on fieldwork efforts

H4. A lower number of contact attempts will be required to finalise a case / to complete an interview in the incentive condition than in the no incentive condition.

3 Study design

3.1 The sample

We use data from the German ALLBUS 2010 survey.[14] The ALLBUS General Social Survey is a face-to-face survey, fielded every second year since 1980. Its goal is the long-term monitoring of attitudes, behaviour and social structure. Sampling and fieldwork is done by a commercial survey organisation. In 2010, TNS Infratest Social Research (Munich) was responsible for data collection (Wasmer, Scholz, Blohm, Walter, & Jutz, 2012).

14 GESIS – Leibniz Institute for the Social Sciences: ALLBUS 2010 - German General Social Survey. GESIS, Cologne, Germany, ZA4610 Data File version 1.0.0. (2011-05-30), doi:10.4232/1.10445.

ALLBUS uses a national area probability sample of non-institutionalised adults in Western and Eastern Germany, with some oversampling in Eastern Germany. In 2010, the sample was drawn in two stages. In the first stage, 149 communities (including 162 Primary Sampling Units, PSUs) were selected. In the second stage, 40 addresses of individuals were randomly selected from the lists of residents for every PSU. The gross sample of ALLBUS 2010 consisted of 6,480 addresses (162 x 40 addresses) in total.

ALLBUS 2010 was fielded as a CAPI survey. The average length of the interview was about 70 minutes. Fieldwork started at the end of May 2010 and ended at the beginning of November 2010. The incentive experiment was part of the main fielding period which lasted until the beginning of August 2010. In the main phase, 205 interviewers were deployed. The interviewers had to make at least four contact attempts to the target persons, spread over different days of the week and different times of the day. The response rate at the end of the main phase was 24.5%. At the end of the re-issue period a final response rate of 34.6% could be achieved.

3.2 The ALLBUS 2010 incentive experiment

When the use of an incentive is planned in a survey, many practical decisions have to be made (Kulka et al., 2005). These decisions refer, for instance, to the type (e.g., monetary vs. non-monetary) and the value of the incentive, to the timing of payment of the incentive (prepaid vs. promised) and to the groups of persons who should be offered the incentive (e.g., all target persons vs. only persons living in areas with low response rates).

For the ALLBUS 2010 experiment, it was decided to study the effect of a €10 conditional cash incentive offered to all target persons. €10 is a modest amount, thus limiting the risk of the incentive being perceived as a payment for time.[15] To avoid confounding area effects, sample members were randomly assigned to treatments within PSUs. In each PSU, 16 addresses were assigned to the €10 incentive condition, and 24 addresses were assigned to the control group (no incentive).

15 In Germany, an incentive worth €10 might be characterised as a token of appreciation rather than a compensation for time spent for most target persons. For groups who are financially less well off, however, the incentive might be attractive also as a form of payment. In Germany, there is no uniform statutory minimum wage that applies to all employees. The various forms of industry-specific minimum wages are sometimes below €10 per hour (see the information provided by the German Federal Statistical Office:
https://www.destatis.de/DE/ZahlenFakten/GesamtwirtschaftUmwelt/VerdiensteArbeitskosten/Mindestloehne/Tabellen/MindestlohnDeutschland.html?nn=50678 (02.04.2012)).

As a rule, all addresses from a PSU were allocated to a single interviewer. This means that each interviewer normally worked in both incentive conditions. Thus, the design also controls for interviewer effects (similar: Lynn et al., 1998; Lynn, 2001; Nicolaas, 2004).

The incentive was announced in the advance letter sent to all target persons by the fieldwork department of the survey organisation a few days before the fielding period started. Thus, all target persons were informed about the incentive in a standardised and direct way. The interviewers knew from the colour of the contact forms which addresses had been promised an incentive and which had not. There was no specific instruction for the interviewers with respect to how prominently they should mention the incentive when approaching the respondents.

3.3 Analysis plan

In order to test our hypotheses (see section 2), we first analyse differences in contact, capability, cooperation and response rates between the treatment conditions (H1).

In a second step, we analyse outcome rates for demographic subgroups using data from the sampling frame. The sampling frame of ALLBUS includes the following variables for each target person (i.e., for both respondents and non-respondents): age, gender, citizenship, size of community and region. We use these variables to examine whether incentives have a differential effect on certain population subgroups and may thus alter sample composition (no specific hypothesis formulated).

Subsequently, we use the ALLBUS data set of completed interviews and look at response distributions and survey results for indicators of low socio-economic status (H2) and low involvement in community, society and politics (H3). For both areas of interest the ALLBUS data set provides several indicators.

The next step we take is to look at different fieldwork efforts. We analyse whether the number of contact attempts is lower in the €10 incentive than in the no incentive condition (H4). The required data come from the contact forms the interviewers had to fill in.

All analyses refer to the results of the main fielding period. In all analyses, the sampling design is taken into account. This refers both to the geographical clustering of respondents (communities) and to the unequal selection probabilities for respondents from Eastern and Western Germany. The analyses were performed using stata 12.1 and IBM SPSS 20.

Table 1 Outcome rates ALLBUS 2010 by use of incentive

	N	Contact rate (%)	N	Capability rate (%)	N	Cooperation rate (%)	N	Response rate (%)
No incentive	3,585	82.3	2,951	91.2	2,690	30.8	3,585	23.1
€10	2,356	82.9	1,954	92.1	1,800	34.9	2,356	26.7
Total	5,941	82.6	4,905	91.5	4,490	32.4	5,941	24.5
$F_{\text{design-based}}$		0.431		0.966		7.092		8.197
(df 1, df 2)		(1, 149)		(1, 149)		(1, 149)		(1, 149)
p-value		.513		.327		.009		.005

Notes: The definition of the outcome rates is provided in Appendix 1. The columns labelled 'N' contain the sample size (= 100%) to which the contact, capability, cooperation and response rate refer. $F_{\text{design-based}}$ *is a variant of the second-order Rao-Scott adjusted chi-square statistic.*

4 Results

4.1 Effect of incentives on outcome rates

Table 1 shows outcome rates for the ALLBUS survey 2010 at the end of the main fielding period (before re-issuing took place). At this stage, the response rate was 24.5%.[16] The use of the €10 incentive had a positive effect on the response rate in ALLBUS 2010. In the incentive group, the response rate is 3.6 percentage points higher than in the control group (p=.005). This is mainly achieved by a higher cooperation rate in the €10 condition: the cooperation rate is 34.9% in the incentive condition and 30.8% in the no incentive condition (p=.009). As expected, there were no significant differences in the contact and capability rate between the experimental conditions. Table 1 therefore shows clear support for our first hypothesis.

4.2 Differential effects for subgroups

With data from the sampling frame on respondents and non-respondents we can test whether incentives are equally effective in attracting different subgroups of the population. We ran logistic regression models with the four survey outcome variables as dependent (dichotomous) variables. The explanatory variables included

16 This is a rather low interim rate, even for the ALLBUS. In the three preceding survey rounds of ALLBUS, the respective rate had been 7 to 11 percentage points higher.

the frame variables gender (female vs. male), citizenship (German vs. non-German), region (Western vs. Eastern Germany), age (in years) and size of community (number of inhabitants, 7 categories), plus the treatment condition (€10 incentive vs. no incentive).

We started with a model including only the main effects of the frame variables and the treatment condition. Table 2 shows that, according to the values of the Pseudo R^2, 'contact' and 'capability' can be better explained by the frame variables than 'cooperation' and 'response'. Turning to the effects of the individual frame variables, we find that females have a higher propensity of being contacted and a lower propensity of being capable to provide an interview than males. Older people are easier to contact than younger people. On the other hand they are less capable of providing an interview. German citizenship has a positive effect on 'capability' and 'response'. This is mainly due to the fact that an insufficient command of the German language is more widespread among non-Germans than among Germans. Regarding region, we do not observe an effect for any of the outcome variables. In contrast, size of community plays a significant role for several outcome variables. Target persons living in larger communities are more difficult to contact, and less cooperative than persons living in smaller communities. As a consequence, the chance of getting an interview is lower in larger communities than in smaller communities. These results are similar to results already observed in previous analyses (Koch, 1997). In addition to the effects of the frame variables, we also find a positive effect from the incentive. As in the bivariate case, the positive effect of the €10 incentive only refers to 'cooperation' and 'response' (and not to 'contact' and 'capability').

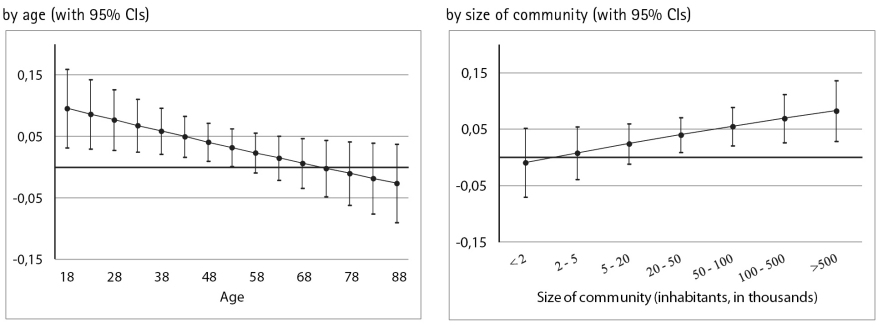
However, one has to be careful when reporting and interpreting parameter estimates of logistic regressions. Logit models are not very sensitive to explicitly specified interaction terms. The reason for this is that even without modelling interaction terms, logit models estimate model-inherent interactions (conditional effects) (Nagler, 1994; Best & Wolf, 2010). In addition, one has to bear in mind that the estimated levels of significance only refer to the latent variables; they do not directly indicate changes in probabilities, which is our main interest (Ai & Norton, 2003). As a consequence, Best and Wolf (2010, 2012) suggest analysing the relationship between independent and dependent variables using a graphic representation and to use the averaged effect as a measure (Average Marginal Effect,

Table 2 Logistic regression of contact, capability, cooperation and response on frame variables and incentives

	Contact		Capability		Cooperation				Response			
	Model 1		Model 1		Model 1		Model 2 (with interactions)		Model 1		Model 2 (with interactions)	
	ß	SE	ß	SE	ß	SE	ß	SE	ß	SE	ß	SE
Female (*male*)	.278***	(.066)	-.246*	(.116)	.102	(.073)	.099	(.072)	.129+	(.068)	.126+	(.068)
German (*non-German*)	.169	(.144)	2.240***	(.185)	.087	(.154)	.083	(.155)	.462**	(.150)	.462**	(.150)
Western Germany (*Eastern Germany*)	.066	(.142)	.19	(.175)	.133	(.114)	.139	(.115)	.158	(.109)	.0163	(.109)
Age, centred	.020***	(.002)	-.047***	(.005)	-.003	(.002)	.000	(.003)	-.003+	(.002)	-.000	(.002)
Size of community, inhabitants	-.240***	(.044)	-.030	(.040)	-.067*	(.032)	-.097**	(.037)	-.119***	(.029)	-.141***	(.035)
€10 incentive (*no incentive*)	.033	(.072)	.174	(.128)	.186*	(.073)	-.112	(.166)	.190**	(.070)	-.024	(.153)
Age, centred*incentives							-.008*	(.004)			-.007+	(.003)
Size of community*incentives							.073+	(.037)			.052	(.034)
constant	2.330***	(.277)	0.783*	(.303)	-.777**	(.239)	-.655**	(.248)	-1.331***	(.227)	-1.245***	(.234)
N LL	5745	-2502.1	4749	-1187.7	4351	-2702.3	4351	-2698.1	5745	-3136.0	5745	-3133.0
AIC	0.873		0.503		1.245		1.244		1.094		1.093	
Pseudo R2	0.055		0.125		0.005		0.007		0.013		0.014	

Notes: Reference category in parentheses. $^{+} = p < .10$, $^{*} = p < .05$, $^{**} = p < .01$, $^{***} = p < .001$

Figure 2 Average Marginal Effects of incentives on the probability of 'cooperation'



AME[17]). In the present case, these were the AMEs of using a respondent incentive on the probability of 'cooperation' and 'response', respectively.

Figure 2 depicts the AMEs for the probability of 'cooperation' and the respective confidence intervals for age and size of community (again, the analyses take the sampling design into account). The positive effect of a respondent incentive for younger age groups can clearly be seen. The effect is significantly different from zero for persons aged 50 years or younger. In contrast, for persons aged 50 years or older the effect is no longer significant. Similarly, the effect of the incentive is dependent on the size of the community. Only for persons living in communities with more than 20,000 inhabitants we find a significant effect of the incentive on 'cooperation'. Appendix 2 includes the respective figures for 'response'. The structure of the results is quite similar to the results for 'cooperation'; the effects, however, are weaker. Generally, we can conclude that the interaction effects are not very strong; they improve the model fit only marginally (see the AIC in Table 2). The graphic representation, however, helps to identify the subgroups for which the use of an incentive has a positive effect that is significantly different from zero.

Taken together, our analyses provide some evidence for weak differential effects of incentives on 'cooperation'. The crucial point is that with respect to nonresponse bias the results work in opposite directions for the two variables 'age' and 'size of community'. Regarding urbanicity, the incentive seems to help to reduce an existing imbalance in cooperation propensities, whereas concerning age the incen-

17 To arrive at Average Marginal Effects (AMEs) for every observation, the marginal effects are computed and then averaged (Long, 1997). AMEs measure the mean change in P(Y=1) as one independent variable X_k increases while all other variables are kept constant (Bartus, 2005).

Table 3 Socio-demographic variables and use of incentives

	N	$F_{design\ based}$	(df 1, df 2)	p-value
Gender	1,455	.044	(1, 147)	.835
Age	1,453	2.174	(4.8, 699.1)	.059
Size of community (inhabitants)	1,455	1.876	(5.4, 791.2)	.091
Citizenship (German/non-German)	1,454	.150	(1, 147)	.699
Region (Western/Eastern Germany)	1,455	.380	(1, 147)	.538

tive seems to introduce an imbalance in cooperation propensities in the first place. When it comes to 'response', however, both effects are attenuated.

4.3 Differences in response distributions in the realised sample

Another possibility to assess differential effects of incentives is to look at the response distributions in the realised sample. Since the incentives in our experiment were offered to a random subsample of the respondents, there should be no significant differences in variable distributions between respondents who received an incentive and those who did not. As long as we can preclude that the incentive had an effect on measurement (by altering the responses provided during the interview), any significant difference in variable distributions indicates that the incentive is more effective in recruiting certain subgroups than others.

The merit of this approach is that a wide array of variables can be examined for potential effects of incentives. In a first step, we repeated the analyses with the same variables as were used in the analyses of the sampling frame data (section 4.2). For each variable, we ran bivariate association tests. The present results (see Table 3; the detailed cross tabulations are provided in Appendix 3) follow the same pattern as the results for the analyses above. For age and size of community, we observe small differences between the incentive and the no incentive condition that reach borderline significance (p=.059 and p=.091). Gender, citizenship and region do not show any difference.

In a second step, we analysed differences for a range of variables indicating the socio-economic situation of the respondent in order to investigate whether H2 (more people with a lower socio-economic status in the incentive condition) comes true. The variables included both objective (e.g., education and income) and subjective measures (e.g., the subjective evaluation of the own economic situation or subjective social class). Table 4 contains the results of the bivariate association

tests for each variable (for the detailed cross tabulations see Appendix 3). The results provide only limited evidence that incentives bring people into the respondent pool who are socio-economically less well off. For two variables the difference between the treatment and the control condition is significant ($p<.050$). The proportion of people living in a rented flat is slightly higher in the €10 condition than in the control group (40.3% vs. 34.1%). Also, the proportion of students is higher in the subsample where the €10 incentive was offered (5.9% vs. 3.6%). However, in a multiple regression analysis with age and size of community as control variables both associations are attenuated and fail (student, $p=.138$) or nearly fail the usual p-level for significance (tenant of a flat, $p=.098$).[18] The other objective and subjective measures of the socio-economic situation of the respondent are not related to the treatment condition. We find, for instance, no difference with regard to the educational level of respondents, their income or their subjective and objective social class placement.

The situation looks similar for all indicators of involvement in community, society and politics. We postulated a higher proportion of people with low involvement in the incentive condition (H3). However, the response distributions for both the subjective and objective variables in ALLBUS referring to this topic do not dif fer between the no incentive and the €10 incentive condition (see Table 5; for the detailed cross tabulations see Appendix 3). For instance, we find no indication that persons less interested in politics are more attracted by an incentive to participate in ALLBUS than persons highly interested in politics. Similarly, persons with a smaller personal network or non-members of associations cannot be found more often among the group who received an incentive.

Concluding this section we should note that, in addition to the abovementioned variables, we looked at a wide range of other variables for differences in response distributions between the two treatment conditions. The results of these analyses did not provide any further evidence for systematic differences between respondents who received an incentive and those who did not. This holds for both the demographic and the non-demographic variables included in ALLBUS.

18 We performed multiple logistic regressions with the treatment condition (no incentive (0) vs. €10 incentive (1)) as dependent variable. Tenant of a flat and student were both individually used as independent variables, and age and size of community were included as control variables.

Table 4 Socio-economic variables and use of incentives

	N	$F_{design\ based}$	(df 1, df 2)	p-value
Education	1,436	.369	(2.0, 288.4)	.688
HH equivalence income (quintiles, in €)	1,161	1.801	(3.8, 564.2)	.130
Personal income (quintiles, in €)	1,264	.408	(3.9, 575.3)	.799
Tenant of a flat	1,438	2.958	(2.9, 432.6)	.033
Student	1,451	4.123	(1, 147)	.044
Unemployed	1,451	1.269	(1, 147)	.262
Economic situation (subj.)	1,449	1.348	(3.8, 564.0)	.252
Social class (subj.)	1,426	.691	(3.0, 435.1)	.556
Social class (Goldthorpe/Terwey)[19]	1,279	.650	(9.6, 1417.6)	.766

Table 5 Involvement in community, society and politics and use of incentives

	N	$F_{design\ based}$	(df 1, df 2)	p-value
Interest in politics	1,455	.551	(3.9, 566.7)	.692
Voted last election	1,451	.819	(1, 147)	.367
Size of personal network	1,382	.090	(2.9, 426.3)	.962
No. of association memberships	1,455	.053	(3.0, 435.3)	.983
Member of a party/a union	1,455	.174	(1, 147)	.678

4.4 Incentives and fieldwork efforts

The number of contact attempts interviewers have to make either to complete an interview or to achieve a final outcome status for the nonresponding cases is a strong determinant of fieldwork costs in face-to-face surveys. This assertion holds in particular for the number of in-person contact attempts. In our fourth hypothesis, we postulated that fewer contact attempts are needed when an incentive is promised. Table 6 provides no support for this hypothesis. There are no significant differences in the average number of contact attempts between the €10 incentive and the no incentive condition. Both for the gross sample and for the net sample of completed interviews, and both for the total number of contact attempts (in

19 For details see Terwey and Baltzer (2011).

Table 6 Average number of contact attempts by use of incentives

Mode of contact attempt		No incentive	€10
In person + by phone	Gross sample	3.06 (3,585)	3.04 (2,356)
	Net sample (interviews)	3.49 (827)	3.37 (628)
In person only	Gross sample	2.44 (3,585)	2.45 (2,356)
	Net sample (interviews)	2.62 (827)	2.58 (628)

Note: N in parentheses

person + by phone)[20] and the in-person contact attempts only, the results are more or less the same in the two treatment conditions.[21] Thus, in this perspective we do not find an indication that a €10 incentive helps to reduce interviewer efforts.

The situation looks different when we take the higher response rate into account that was achieved in the incentive condition. When we compute the total number of contact attempts for the gross sample and divide it by the number of completed interviews in both treatment conditions, we find that the total effort exerted is lower when an incentive was used (Table 7). The average number of contact attempts per completed interview (including attempts on respondents and non-respondents) is 13.25 in the control group and 11.42 in the €10 treatment group. This means that for a given number of completed interviews on average 14% fewer contact attempts were needed when a €10 incentive was used. A reduction of a similar magnitude is observed when we look at personal contact attempts only (on average 10.57 vs. 9.20 attempts).

Unfortunately, we do not know what this reduction in efforts meant for the actual survey costs in ALLBUS. An estimation of the cost saving requires knowing about the costs of an interviewer visit. The survey organisation should be able to estimate these costs. However, a survey sponsor usually will not be in the position to receive this information. For this reason, a simple thought experiment might shed some light on this issue.

20 Interviewers were allowed to get in touch with the target person by telephone in order to make an appointment for an interview. However, the interviews had to be completed face-to-face. Since evidence exists that respondent cooperation can be more easily achieved in a face-to-face than in a telephone contact (Blohm, Hox, & Koch, 2007), the general recommendation for the ALLBUS interviewers was to make contact with target persons by a personal visit. As can be seen from Table 6, interviewers followed this recommendation: Most contact attempts were made in person.

21 Note that the results refer to the main fielding period only. The average number of contact attempts are higher when we also include the additional attempts interviewers made during the re-issue phase.

Table 7 Total number of contact attempts per completed interview by use of incentives

Mode of contact attempt		No incentive	€10
In person + by phone	Total # of contact attempts to all units of the gross sample	10,961	7,170
	# of completed interviews	827	628
	Average # of contact attempts per completed interview	13.25	11.42
In person only	Total # of contact attempts to all units of the gross sample	8,745	5,780
	# of completed interviews	827	628
	Average # of contact attempts per completed interview	10.57	9.20

The ALLBUS survey usually aims for 3,500 completed interviews. From the above analyses we know that in the condition with no incentive 10.57 personal contact attempts are made on average per completed interview (taking into account the total number of contact attempts on the complete gross sample). When we use a conditional incentive of €10, the same number of interviews can be completed with 9.20 contact attempts on average. The introduction of a €10 incentive for the whole sample of ALLBUS thus would mean a reduction in the total number of personal contact attempts by 4,795 attempts (= 3,500 x 1.37). On the other hand, the use of a conditional monetary incentive of €10 would cause costs of €35,000 (= 3,500 x 10€; not taking into account any handling costs). Accordingly, we might conclude that the introduction of the incentive would pay for itself if the average costs of a personal contact attempt amount to €7.30 or more (= €35,000 / 4,795).[22]

5 Summary and conclusions

In ALLBUS 2010, an experiment was set up to test the effect of respondent incentives on outcome rates, sample composition and fieldwork efforts. A random subsample of target persons was offered a monetary incentive of €10 to be paid upon completion of the interview. The other part of the sample acted as a control group and received no incentive. In the following, we briefly summarise and discuss the

22 Note that we arrive at a very similar result when we use the total number of contact attempts in the main and re-issue phase for the calculation.

results of the experiment, separately for the three areas of outcome rates, sample composition and fieldwork efforts.

(1) Outcome rates

The monetary incentive used in ALLBUS 2010 led to an increase in the response rate (3.6 percentage points), mainly by improving the cooperation rate. This positive effect occurred at a rather low response rate level. In addition, we should point out that this response rate difference was observed at the end of the main fielding period, before re-issuing efforts were undertaken. Both factors suggest that the effect size we found is probably an upper bound for the size of an effect of a (modest) monetary incentive in a survey like ALLBUS. We might expect a smaller effect in surveys with higher response rates (Singer, 2002). We might also expect that in surveys implementing a re-issue phase, additional interviewer efforts in motivating previously reluctant target persons will tend to decrease the effect of an incentive. Fieldwork in ALLBUS 2010 included a re-issue phase. However, our data do not allow to assess in a rigorous manner whether the latter hypothesis is true or not, since in the re-issue period of ALLBUS 2010 an incentive of €20 was offered to all remaining target persons. After the finalisation of fieldwork, the response rate of ALLBUS 2010 was 34.6%, around 10 percentage points higher than at the end of the main fielding period. Despite the fact that the €20 incentive was offered to all sampling units in the re-issue phase, still a difference in the response rate of 2.9 percentage points could be observed between the original €10 incentive group and the control group. This difference is only a little smaller than the difference at the end of the main fielding period.

Our results refer to a conditional monetary incentive of a modest amount. They do not preclude that monetary incentives of a higher value lead to a more substantial increase in cooperation and response in surveys similar to ALLBUS. Eyerman et al. (2005) provide some empirical evidence that this is actually the case as regards conditional incentives. Likewise, Boersch-Supan et al. (2013) report that an increase in the value of an unconditional incentive (€10 - €20 - €40) led to higher cooperation rates in an almost linear fashion. In contrast to these results, Scherpenzeel and Toepoel (2012) found no difference in response rates when comparing an incentive worth €10 with an incentive worth €20 or €50 (including both promised and prepaid incentives). The different results illustrate that the effects of incentives may vary depending on the concrete survey context.

(2) Sample composition

The incentive used in ALLBUS 2010 did not change sample composition in a major way. We found that the incentive was somewhat more attractive to younger people and to people living in urban areas. Also, the incentive condition contained a slightly higher number of students and tenants of a flat. The latter results might be interpreted as a clue that a €10 incentive disproportionately attracts people who are financially less well off. At the same time, however, we have to acknowledge that other, more direct indicators of socio-economic status (like income) did not differ at all. One might speculate whether a larger monetary incentive might be followed by larger differences in sample composition regarding indicators of socio-economic status.

We also hypothesised that persons with a lower interest in politics and lower involvement in community and society would disproportionately be attracted by an incentive. The idea behind this was that interest in politics or involvement in community might indicate some kind of feeling of civic duty or a norm of cooperation that would motivate people with these predispositions to disproportionately participate in the ALLBUS survey. People with a low interest in politics or a low involvement in community, on the other hand, would be more attracted by an extrinsic benefit like an incentive. The results of our experiment provide no support for this hypothesis. A possible explanation might be that the variables we looked at are not well suited to capture the underlying concepts of 'civic duty' or 'norms of cooperation'. In addition, we might also ask whether 'norms of cooperation' are more relevant for motivating people to participate in surveys in the United States than in Germany. In any case, the widespread absence of differences in response distributions between the incentive and the no incentive condition in ALLBUS indicates that the incentive did not introduce any imbalance in sample composition.

(3) Fieldwork efforts

Decisions on survey design features typically have to be made against the backdrop of a limited project budget. The use of incentives, on the one hand, increases costs. These costs include both the pure costs of the incentives and the associated costs of their handling and administration. The crucial question is whether incentives, on the other hand, help to save costs by reducing the fieldwork efforts interviewers have to exert. This will be the case when an incentive makes (at least some) sample members more readily willing to participate.

In the ALLBUS experiment, we found that using an incentive allowed the survey organisation to achieve a certain number of interviews with a lower number

of contact attempts than when no incentive was used. From the outside it was not possible for us to provide a reliable estimation what this reduction in efforts meant for the actual survey costs in ALLBUS. Apart from that, one has to be aware that this cost saving first and foremost is a benefit for the survey organisation. In survey practice, it will often be questionable whether the survey organisation passes the cost savings associated with the use of a respondent incentive to the survey sponsor.

To conclude, the conditional monetary incentive of €10 used in ALLBUS 2010 had a small positive effect on the cooperation and response rate achieved. There is hardly any indication that the incentive altered the sample composition of the ALLBUS survey. Regarding survey costs, the incentive led to a small reduction in the fieldwork efforts the interviewers had to exert.

For the decision on the future use of a respondent incentive in ALLBUS, the willingness of the survey organisation to share the cost savings associated with the use of an incentive with the survey sponsor will play an important role. If the survey organisation is willing to pass on the cost savings, a higher response rate might be achieved for a similar price – which would be a clear argument in favour of using an incentive in upcoming survey rounds of ALLBUS. In case the survey organisation will not pass the cost savings to the survey sponsor, we might say that the decision will depend on the answer to the following question: Are you willing to accept an additional cost burden (3,500 interviews x €10 = €35,000) in order to achieve an increase in response by 3-4 percentage points? A well-considered answer to this question has to weigh the incentive costs against the costs of other survey features (like, e.g., better training of interviewers) that are also deemed adequate to increase response rates or to improve data quality in general (Rodgers, 2011).

In addition, for an informed choice we might also wish to take into account the potential effects of incentives on measurement. In the present paper, we did not investigate whether or not the €10 incentive changed the response behaviour of the ALLBUS 2010 respondents. Incentives can affect data quality, for instance, by changing the level of item nonresponse or by altering the expression of opinions. Up till now there exists only limited evidence on the effects of incentives on measurement. Some studies reveal a positive effect, others a negative effect. More recent studies often found no association between receiving an incentive and data quality (Singer & Ye, 2013). Given the rather modest value of the incentive used in the present study, large effects on measurement do not seem to be very likely.

Nevertheless, only future research can show whether this assumption is actually true or not.

Beyond ALLBUS, the present results should be relevant for other surveys as well. A very conservative approach would request that each survey considering the use of incentives should implement its own experiment to find out about the potential effects. We do not think that this claim has to be asserted in such a strict manner. The results of a particular experiment like the present one provide a reference point for other studies that follow a similar design and take place in a comparable social context. Thus, social surveys of the general population that are fielded face-to-face and cover a broad range of topics might profit from the present results. As we elaborated in the introduction, there is a large number of surveys all over the world for which this description applies.

References

Abraham, K. G., Helms, S., & Presser, S. (2009). How Social Processes Distort Measurement: The Impact of Survey Nonresponse on Estimates of Volunteer Work in the United States. *American Journal of Sociology*, 114(4), 1129-1165.

Ai, C., & Norton, E. C. (2003). Interaction Terms in Logit and Probit Models. *Economics Letters*, 80(1), 123-129.

The American Association for Public Opinion Research (2011). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition*. AAPOR.

Atrostic, B. K., Bates, N., Burt, G., & Silberstein, A. (2001). Nonresponse in U.S. Government Household Surveys: Consistent Measures, Recent Trends, and New Insights. *Journal of Official Statistics*, 17(2), 209-226.

Barón, J. D., Breunig, R. V., Cobb-Clark, D., Gørgens, T., & Sartbayeva, A. (2009). Does the Effect of Incentive Payments on Survey Response Rates Differ by Income Support History? *Journal of Official Statistics*, 25(4), 483-507.

Bartus, T. (2005). Estimation of marginal effects using margeff. *The Stata Journal*, 5(3), 309-329.

Berlin, M., Mohadjer, L., Waksberg, J., Kolstad, A., Kirsch, I., Rock, D., & Yamamoto, K. (1992). An Experiment in Monetary Incentives. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 393-398.

Best, H., & Wolf, C. (2010). Logistische Regression. In C. Wolf & H. Best (Eds.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (pp. 827-854). Wiesbaden: VS Verlag für Sozialwissenschaften.

Best, H., & Wolf, C. (2012). Modellvergleich und Ergebnisinterpretation in Logit- und Probit-Regressionen. *Kölner Zeitschrift für Soziologie und Sozialwissenschaften*, 64(2), 377-395.

Blohm, M., Hox, J., & Koch, A. (2007). The Influence of Interviewers' Contact Behavior on the Contact and Cooperation Rate in Face-to-Face Household Surveys. *International Journal of Public Opinion Research*, 19(1), 97-111.

Börsch-Supan, A., Krieger, U., & Schröder, M. (2013). *Respondent Incentives, Interviewer Training and Survey Participation*. SHARE Working Paper 12-2013, Munich Center for the Economics of Aging (MEA).

Castiglioni, L., Pforr, K., & Krieger, U. (2008). The Effect of Incentives on Response Rates and Panel Attrition: Results of a Controlled Experiment. *Survey Research Methods*, 2(3), 151-158.

Church, A. H. (1993). Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis. *Public Opinion Quarterly*, 57(1), 62-79.

de Leeuw, E., & de Heer, W. (2002). Trends in Household Survey Nonresponse. A Longitudinal and International Comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey Nonresponse* (pp. 41-54). New York: John Wiley & Sons.

Dixon, J., & Tucker, C. (2010). Survey Nonresponse. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of Survey Research.* Second Edition (pp. 593-630). Bingley: Emerald.

Eurofound (2012). *Third European Quality of Life Survey – Quality of life in Europe: Impacts of the crisis.* Publications Office of the European Union, Luxembourg.

Eyerman, D., Bowman, K., Butler, C., & Wright, D. (2005). The differential impact of incentives on refusals: Results from the 2001 national household survey on drug abuse incentive experiment. *Journal of Economic and Social Measurement*, 30(2-3), 157-169.

Groves, R. M. (2008). Verbal statement during a "Seminar On Survey Respondent Incentives: Research and Practice", held on March 10 by The Council of Professional Associations on Federal Statistics. Retrieved September 08, 2011, from http://www.copafs.org/semi nars/research_and_practice.aspx

Groves, R. M., & Couper, M. P. (1998). *Nonresponse in Household Interview Surveys.* New York: John Wiley & Sons.

Groves, R. M., Fowler, F. J. Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey Methodology.* Hoboken, New Jersey: John Wiley & Sons.

Groves, R. M., & Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias. A Meta-Analysis. *Public Opinion Quarterly*, 72(2), 167-189.

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage Saliency Theory of Survey Participation. *Public Opinion Quarterly*, 64(3), 299-308.

Koch, A. (1997). Teilnahmeverhalten beim ALLBUS 1994. Soziodemographische Determinanten von Erreichbarkeit, Befragungsfähigkeit und Kooperationsbereitschaft. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 49(1), 98-122.

Kohut, A., Keeter, S., Doherty, C., Dimock, M., & Christian, L. (2012). *Assessing the Representativeness of Public Opinion Surveys*. The Pew Research Center.

Kulka, R. A., Eyerman, J., & McNeeley, M. E. (2005). The use of monetary incentives in federal surveys on substance use and abuse. *Journal of Economic and Social Measurement*, 30(2-3), 233-249.

Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, California: Sage Publications.

Lynn, P. (2001). The Impact of Incentives on Response Rates to Personal Interview Surveys: Role and Perceptions of Interviewers. *International Journal of Public Opinion Research*, 13(3), 326-336.

Lynn, P., Thomson, K., & Brook, L. (1998). An Experiment with Incentives on the British Social Attitudes Survey. *Survey Methods Centre Newsletter*, 18(2), 12-14.

Nagler, J. (1994): Scobit: An Alternative Estimator to Logit and Probit. *American Journal of Political Science,* 38(1), 230-255.

Nicolaas, G. (2004). The use of incentives to motivate "hard to get" households on the National Travel Survey. *Survey Methods Newsletter*, 22(Spring), 19-27.

Nicolaas, G., & Stratford, N. (2005). A Plea for the Tailored Use of Respondent Incentives. In C. van Dijkum, J. Blasius, & C. Durand (Eds.), *Recent developments and applications in social research methodology* (CD-ROM proceedings of the RC33 Sixth International Conference on Social Science Methodology, Workshop No. 321). Amsterdam: Budrich Verlag.

Phelps, A. (2008). *UK Incentive Experiment.* Paper presented at the ESS Field Directors Meeting, Mannheim, January 2008.

Rodgers, W. L. (2011). Effects of Increasing the Incentive Size in a Longitudinal Study. *Journal of Official Statistics*, 27(2), 279-299.

Scherpenzeel, A. & Toepoel, V. (2012). Recruiting a Probability Sample for an Online Panel. Effects of Contact Mode, Incentives, and Information. *Public Opinion Quarterly*, 76(3), 470-490.

Singer, E. (2002). The Use of Incentives to Reduce Nonresponse in Household Surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey Nonresponse* (pp. 163-177). New York: John Wiley & Sons.

Singer, E. (2006). Introduction – Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, 70(5), 637-645.

Singer, E., & Kulka, R. A. (2001). Paying Respondents for Survey Participation. In M. V. Ploeg, R. A. Mofitt, & C. F. Citro (Eds.), *Studies of Welfare Populations: Data Collection and Research Issues* (pp. 105-128). Committee on National Statistics, National Research Council.

Singer, E., van Hoewyk, J., Gebler, N., Raghunathan, T., & McGonagle, K. (1999). The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys. *Journal of Official Statistics*, 15(2), 217-230.

Singer, E., & Ye, C. (2013). The Use and Effects of Incentives in Surveys. *The ANNALS of the American Academy of Political and Social Science*, 645, 112-141.

Smith, T. W., Kim, J., Koch, A., & Park, A. (2006). Social-Science Research and the General Social Surveys. *Comparative Sociology*, 5(1), 33-43.

Terwey, M., & Baltzer, S. (2011): *Variable Report ALLBUS / Allgemeine Bevölkerungsumfrage der Sozialwissenschaften 2010*. ZA-Nr. 4610. GESIS-Variable Reports 2011/02. Köln: GESIS – Leibniz Institut für Sozialwissenschaften. doi:10.4232/1.10760

Wasmer, M., Scholz, E., & Blohm, M. (2010). *Konzeption und Durchführung der "Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften" (ALLBUS) 2008.* GESIS-Technical Reports 2010|04. Mannheim: GESIS – Leibniz Institut für Sozialwissenschaften.

Wasmer, M., Scholz, E., Blohm, M., Walter, J., & Jutz, R. (2012). *Konzeption und Durchführung der "Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften" (ALLBUS) 2010.* GESIS-Technical Reports 2012|12. Mannheim: GESIS – Leibniz Institut für Sozialwissenschaften.

Wetzels, W., Schmeets, H., van den Brakel, J., & Feskens, R. (2008). Impact of Prepaid Incentives in Face-to-face Surveys: A Large-Scale Experiment with Postage Stamps. *International Journal of Public Opinion Research*, 20(4), 507-516.

Wright, D., Bowman, K., Butler, D., & Eyerman, J. (2005). Non-response bias from the national household survey on drug abuse incentive experiment. *Journal of Economic and Social Measurement*, 30(2-3), 219-231.

Addresses of authors

Michael Blohm
GESIS – Leibniz Institute for the Social Sciences
P.O. Box 122155
68072 Mannheim
Germany
E-Mail: michael.blohm@gesis.org

Achim Koch
GESIS – Leibniz Institute for the Social Sciences
P.O. Box 122155
68072 Mannheim
Germany
E-Mail: achim.koch@gesis.org

Appendix 1: Definition of outcome rates

In household surveys like ALLBUS, survey participation is a sequential process, including several steps that have to be passed successfully in order to achieve a completed interview (Groves & Couper, 1998). First, the target persons have to be contacted. Second, the target persons' capability to respond has to be assessed. Third, target persons who have been successfully contacted and who have proven able to respond have to be motivated to actually provide the interview. Accordingly, we analyse four separate outcome variables: the contact rate, the capability rate, the cooperation rate and the response rate. We rely on the following procedures and assumptions when calculating the rates:

- All rates are respondent-level rates.
- There are no cases of unknown eligibility.
- "Partial" interviews are not accepted; they are classified as invalidated interviews.
- The outcome "Not able" also includes "cases not fully processed/invalidated interviews" in addition to the main codes "permanently ill/disabled" and "language difficulties".

Our outcome rates compare to the AAPOR Standard Definitions (2011) in the following way:

Contact rate: AAPOR CON 3

- Contacts excepting frame errors
  $(N_{interviews} + N_{not\ able} + N_{refused}) / (N_{interviews} + N_{not\ contacted} + N_{not\ able} + N_{refused})$

Capability rate: no AAPOR definition available

- Capability excepting frame errors and noncontacts
  $(N_{interviews} + N_{refused}) / (N_{interviews} + N_{not\ able} + N_{refused})$

Cooperation rate: AAPOR COOP 3

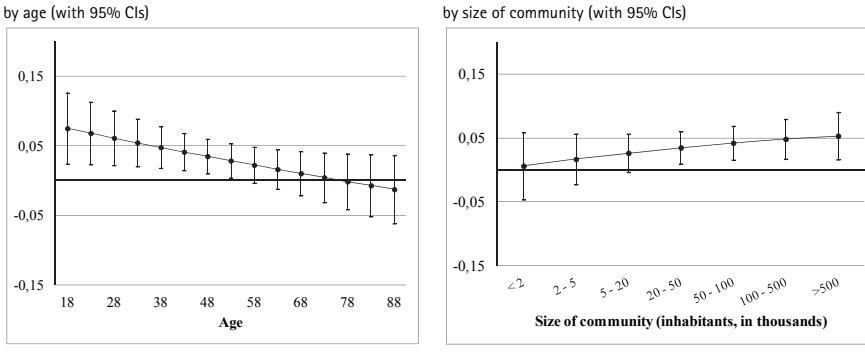
- Cooperation excepting frame errors, noncontacts and cases unable to respond
  $(N_{interviews}) / (N_{interviews} + N_{refused})$

Response rate: AAPOR RR 5

- Response excepting frame errors
  $(N_{interviews}) / (N_{interviews} + N_{not\ contacted} + N_{not\ able} + N_{refused})$

Appendix 2: Average Marginal Effects

Figure A2: Average Marginal Effects of incentives on the probability of 'response'



Appendix 3: Cross tabulation of ALLBUS 2010 variables by incentive

Notes: All cross tabulations are based on weighted counts. F (design based) is a variant of the second-order *Rao-Scott adjusted chi-square* statistic.

Table A3.1: Socio-demographic variables and use of incentives (column percent)

	No incentive	€10 incentive	$F_{design\ based}$	(df 1, df 2)	p-value
Gender			.044	(1, 147)	.835
Male	45.7	46.3			
Female	54.3	53.7			
N	*827*	*628*			
Age			2.174	(4.8, 699.1)	.059
18-29	15.9	18.2			
30-39	13.2	13.7			
40-49	18.3	22.2			
50-59	19.3	14.1			
60-69	16.1	16.4			
70 +	17.2	15.3			
N	*826*	*627*			
Size of community (inhabitants)			1.876	(5.4, 791.2)	.091
– 1,999	9.2	6.9			
2,000 – 4,999	15.1	13.5			
5,000 – 19,999	25.1	24.0			
20,000 – 49,999	14.9	17.3			
50,000 – 99,999	6.6	8.7			
100,000 – 499,999	18.9	17.3			
500,000 +	10.2	12.3			
N	*827*	*628*			
Citizenship			.150	(1, 147)	.699
German	95.4	94.9			
Non-German	4.6	5.1			
N	*826*	*628*			
Region			.380	(1, 147)	.538
Western Germany	82.7	81.7			
Eastern Germany	17.3	18.3			
N	*827*	*628*			

Table A3.2: Socio-economic variables and use of incentives (column percent)

	No incentive	€10 incentive	$F_{design\ based}$	(df 1, df 2)	p-value
Education			.369	(2.0, 288.4)	.688
Certificate of Secondary Education (*Hauptschule*)	37.8	36.9			
Advanced Technical College Entrance Qualification (*Mittlere Reife & Fachhochschulreife*)	33.3	32.2			
University Entrance Qualification/Adv. Tech. College Degree/ University Degree (*Abitur/ Fachhochschul-/Universitätsabschluss*)	28.8	30.9			
N	*817*	*619*			
HH equivalence income (quintiles, in €)			1.801	(3.8, 564.2)	.130
– 849	15.8	20.8			
850 – 1149	20.2	19.6			
1150 – 1459	23.1	17.7			
1460 – 1979	20.3	20.5			
1980 +	20.7	21.3			
N	*653*	*509*			
Personal income (quintiles, in €)			.408	(3.9, 575.3)	.799
None	20.4	22.6			
– 499	21.4	21.0			
500 – 999	22.6	20.1			
1000 – 1499	16.8	16.5			
1500 +	18.9	19.9			
N	*716*	*548*			
Tenant / owner			2.958	(2.9, 432.6)	.033
Tenant of flat	34.1	40.3			
Tenant of house	5.2	3.6			
Owner of flat	8.4	5.8			
Owner of house	52.3	50.3			
N	*821*	*618*			
Student			4.123	(1, 147)	.044
No	96.4	94.1			
Yes	3.6	5.9			
N	*825*	*626*			
Unemployed			1.269	(1, 147)	.262
No	94.9	96.0			
Yes	5.1	4.0			
N	*825*	*626*			

	No incentive	€10 incentive	$F_{design\ based}$	(df 1, df 2)	p-value
Economic situation (subj.)			1.348	(3.8, 564.0)	.252
Very good	3.2	4.9			
Good	49.8	50.8			
Partly good/partly bad	34.1	30.6			
Bad	10.3	9.9			
Very bad	2.7	3.9			
N	*823*	*627*			
Social class (subj.)			.691	(3.0, 435.1)	.556
Lower class	3.1	3.5			
Working class	26.9	24.3			
Middle class	59.3	59.4			
Upper middle class and upper class	10.6	12.7			
N	*817*	*609*			
Social class (Goldthorpe/Terwey)[23]			.650	(9.6, 1417.6)	.766
Higher service class	14.1	14.0			
Lower service class	28.1	27.7			
Non-manual, clerical	7.9	8.7			
Proprietor, 2-49 employees	3.4	2.4			
Self-empl., 1/no employee	3.9	4.7			
Farmer, smallholder	1.9	1.4			
Foreman, technician	8.4	8.9			
Skilled worker	14.3	12.3			
Semi-/unskilled worker	11.4	13.8			
Agricultural worker	1.8	0.9			
Other non-manual worker	4.8	5.1			
Farmer, co-op farmer	0.1	0.1			
N	*732*	*547*			

23 For details see Terwey and Baltzer (2011).

Table A3.3: Involvement in community, society and politics and use of incentives (column percent)

	No incentive	€10 incentive	$F_{design\ based}$	(df 1, df 2)	p-value
Interest in politics			.551	(3.9, 566.7)	.692
Very strong	10.4	9.7			
Strong	22.3	23.5			
Medium	41.6	40.7			
Weak	18.3	17.0			
Not at all	7.4	9.1			
N	*827*	*627*			
Voted last election			.819	(1, 147)	.367
No	18.7	20.6			
Yes	81.3	79.4			
N	*825*	*626*			
Size of personal network			.090	(2.9, 426.3)	.962
0	7.9	7.8			
1	13.4	13.4			
2	16.3	17.4			
3 or more persons	62.3	61.5			
N	*792*	*590*			
No. of association memberships			.053	(3.0, 435.3)	.983
0	43.3	43.8			
1	31.8	32.2			
2	14.8	14.1			
3 or more	10.1	9.9			
N	*827*	*628*			
Member of a party / a union			.174	(1, 147)	.678
No	84.0	83.2			
Yes	16.0	16.8			
N	*827*	*628*			