SSOAR

Open Access Repository
www.ssoar.info

Are incentive effects on response rates and nonresponse bias in large-scale, face-to-face surveys generalizable to Germany? Evidence from ten experiments

Pforr, Klaus; Blohm, Michael; Blom, Annelies G.; Erdel, Barbara; Felderer, Barbara; Fräßdorf, Anna; Hajek, Kristin; Helmschrott, Susanne; Kleinert, Corinna; Koch, Achim; Krieger, Ulrich; Kroh, Martin; Martin, Silke; Saßenroth, Denise; Schmiedeberg, Claudia; Trüdinger, Eva-Maria; Rammstedt, Beatrice

Postprint / Postprint
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Pforr, K., Blohm, M., Blom, A. G., Erdel, B., Felderer, B., Fräßdorf, A., ... Rammstedt, B. (2015). Are incentive effects on response rates and nonresponse bias in large-scale, face-to-face surveys generalizable to Germany? Evidence from ten experiments. *Public Opinion Quarterly*, *79*(3), 740-768. https://doi.org/10.1093/poq/nfv014

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-58656-1

This is a pre-copyedited, author-produced PDF of an article accepted for publication in
Public Opinion Quarterly, Vol. 79, 2015, No. 3, pp. 740-768 following peer review.
The version of record is available online at: https://doi.org/10.1093/poq/nfv014
Page numbers have been adjusted to the publishers version, whereby this postprint is fully quotable.

ARE INCENTIVE EFFECTS ON RESPONSE RATES AND NONRESPONSE BIAS IN LARGE-SCALE, FACE-TO-FACE SURVEYS GENERALIZABLE TO GERMANY?

EVIDENCE FROM TEN EXPERIMENTS

KLAUS PFORR*
MICHAEL BLOHM
ANNELIES G. BLOM
BARBARA ERDEL
BARBARA FELDERER
MATHIS FRÄßDORF
KRISTIN HAJEK
SUSANNE HELMSCHROTT
CORINNA KLEINERT
ACHIM KOCH
ULRICH KRIEGER
MARTIN KROH
SILKE MARTIN
DENISE SAßENROTH
CLAUDIA SCHMIEDEBERG
EVA-MARIA TRÜDINGER
BEATRICE RAMMSTEDT

---

* Klaus Pforr is a researcher at GESIS-Leibniz-Institute for the Social Sciences (GESIS), Mannheim, Germany. Michael Blohm is a researcher at GESIS, Mannheim, Germany. Annelies G. Blom is an assistant professor at the School of Social Sciences, University of Mannheim, Mannheim, Germany, and principal investigator for the German Internet Panel at the Collaborative Research Center “Political Economy of Reforms” (SFB 884), University of Mannheim, Mannheim, Germany. Barbara Erdel is a researcher at the Institute for Employment Research (IAB) of the German Federal Employment Agency, Nuremberg, Germany. Barbara Felderer is a researcher at the University of Mannheim, Mannheim, Germany. Mathis Fräßdorf is a senior researcher at the German Institute for Economic Research (DIW Berlin), Berlin, Germany. Kristin Hajek is a researcher in the Department of Sociology, University of Munich, Munich, Germany. Susanne Helmschrott is a researcher at the Collaborative Research Center “Political Economy of Reforms” (SFB 884), University of Mannheim, Mannheim, Germany. Corinna Kleinert is a professor of sociology with special focus on longitudinal education research at the University of Bamberg, Bamberg, Germany, and the Leibniz Institute for Educational Trajectories, Berlin, Germany. Achim Koch is a researcher at GESIS, Mannheim, Germany. Ulrich Krieger is a researcher at the Collaborative Research Center “Political Economy of Reforms” (SFB 884), University of Mannheim, Mannheim, Germany. Martin Kroh is Deputy Director of the Socio-Economic Panel Study (SOEP) at DIW Berlin and a professor in the Department of Social Sciences at Humboldt-Universität zu Berlin, Berlin, Germany. Silke Martin is a researcher at GESIS, Mannheim, Germany. Denise Saßenroth

Abstract In survey research, a consensus has grown regarding the effectiveness of incentives encouraging survey participation across different survey modes and target populations. Most of this research has been based on surveys from the United States, whereas few studies have provided evidence that these results can be generalized to other contexts. This paper is the first to present comprehensive information concerning the effects of incentives on response rates and nonresponse bias across large-scale surveys in Germany. The context could be viewed as a critical test for incentive effects because Germany's population is among the most survey-critical in the world, with very low response rates. Our results suggest positive incentive effects on response rates and patterns of effects that are similar to those in previous research: The effect increased with the monetary value of the incentive; cash incentives affected response propensity more strongly than lottery tickets do; and prepaid incentives could be more cost effective than conditional incentives. We found mixed results for the effects of incentives on nonresponse bias. Regarding large-scale panel surveys, we could not unequivocally confirm that incentives increased response rates in later panel waves.

Survey researchers have been increasingly concerned with decreasing response rates, a change that has been reported in developed countries over the past several decades (Atrostic et al. 2001; de Leeuw and de Heer 2002; Brick and Williams 2013). Decreasing response rates can lead to biased estimates if the

---

is a senior researcher at DIW Berlin, Berlin, Germany. Claudia Schmiedeberg is a researcher in the Department of Sociology, University of Munich, Munich, Germany. Eva-Maria Trüdinger is a researcher in the Department of Political Science and Political Sociology, University of Stuttgart, Stuttgart, Germany. Beatrice Rammstedt is Scientific Director of the Department of Survey Design and Methodology, GESIS, Mannheim, Germany, and a professor of psychological assessment, survey design, and methodology at the University of Mannheim, Mannheim, Germany. The authors thank Katrin Auspurg, Henning Best, Mick Couper, Stephanie Eckman, Rainer Schnell, and three anonymous reviewers for helpful comments. This work was supported by the German Federal Ministry of Education and Research (BMBF) [01UW1205]. ALLBUS is conducted by GESIS and receives continued funding from the federal government and the German Länder, through their Joint Science Conference. The German Internet Panel is the central data collection (project Z1) of SFB 884 at the University of Mannheim and is funded by the German Research Foundation (DFG). From 2009 through 2013, the NEPS adult study was funded by the BMBF and run by an interdisciplinary network led by the University of Bamberg. The PASS panel survey is funded by the Federal Ministry of Labor and Social Affairs (BMAS) and run by the IAB. The pairfam experiment was conducted in the context of the DFG Priority Programme “Beziehungs- und Familienpanel” [SPP 1161], funded by the DFG. PIAAC, the Programme for the International Assessment of Adult Competencies, in Germany has been funded by the BMBF [PIA001] with the participation of the BMAS. The SOEP has received continued funding from the federal government and the State of Berlin, through the Joint Science Conference, since 2002. Before that it was primarily funded through the DFG.
*Address correspondence to Klaus Pforr, GESIS–Leibniz Institute for the Social Sciences, B2, 1, 68159 Mannheim, Germany; e-mail: klaus.pforr@gesis.org.

nonresponse is not at random (Rubin 1976). Even when nonresponse is not selective, increasing the sample size as a direct countermeasure incurs higher costs. To increase survey response, several methods have been developed, such as advance letters, special contacting procedures, interviewer training, and various forms of incentives (Groves and Couper 1998; Groves et al. 2009; Schoeni et al. 2013).

This paper contributes to the existing research regarding incentive effects on government-sponsored, large-scale, face-to-face surveys by providing the first comprehensive overview of incentive experiments conducted on surveys in Germany. We focus on face-to-face surveys because they tend to be the primary mode of data collection for large-scale social surveys in the United States and Europe. The current literature indicates ample evidence concerning incentive effects in large-scale, face-to-face-surveys (see Singer [2002] and Singer and Ye [2013] for reviews). However, on closer consideration, the amount of evidence for these surveys is limited. Even Singer (2002, 176) qualified the generalizability of her results: “many of the findings are based on one or a few experiments...[and] a great deal of specification and replication is needed.” This deficiency has not improved much since then, and Singer and Ye (2013, 135) found “only one post-2002 report of an incentive experiment carried out in a cross-sectional face-to-face study.”

The question of the cross-national transferability of incentive effects has been raised before in the literature (e.g., Singer et al. 1999; Couper and de Leeuw 2003; Blom, Jäckle, and Lynn 2010). From a theoretical perspective, incentive effects could depend on cultural and socio-demographic circumstances, as different degrees of resistance to surveys must be overcome through other fieldwork efforts (e.g., Johnson et al. 2002; Cantor, O'Hare, and O'Connor 2008). There is also empirical evidence for cross-national differences in fieldwork effects in general (e.g., Nicoletti and Buck 2004) and particularly regarding incentives (e.g., Mutti et al. 2014).

Large-scale, face-to-face surveys in Germany constitute a critical test of incentive effect theories, as response rates are very low in Germany compared with international standards (Stoop et al. 2010). Independent of Germany's low response rate, Singer et al. (1999) found lower incentive effects in face- to-face surveys.[1] A replication in these surveys in Germany could confirm the

[1] Although Singer et al. (1999) found higher incentive effects for lower response rates, the reanaly¬sis by Gelman, Stevens, and Chan (2003) showed that these effects might have followed from the sample of experiments. The average response rate without incentives was approximately 60 percent. An inherently nonlinear relationship exists between response propensities and response rates. Thus, the incentive effects on response rates could be smaller for experiments with response rates of more than 60 percent than for experiments with response rates of less than 60 percent, even if the incen¬tive effect on response propensity is constant for all response rates. Therefore, because response rates in Germany are much lower than in the United States (as depicted in table 2), incentive effects on response rates might be smaller than in the United States. The combination of expected smaller incentive effects in face-to-face surveys in general and the expected smaller incentive effects in Germany in particular justifies the presented experimental data as a critical case.

theories regarding incentive effects. A replication in these surveys could also strengthen the legitimacy of using incentives in large-scale, face-to-face surveys, which has been scrutinized not only in Germany but also recently in the United States (Marketing Research Association 2015; Fienberg 2013; Pierson 2013). We examine the effects of incentives across 10 experiments implemented in the following eight German surveys: (1) the "German General Social Survey" (ALLBUS) (Koch and Wasmer 2004); (2) the "German Internet Panel" (GIP) (Blom, Gathmann, and Krieger forthcoming); (3) the adult panel of the "National Educational Panel Study" (NEPS) (Allmendinger et al. 2011); (4) the German Family Panel (pairfam) (Huinink et al. 2011); (5) the panel study "Labor Market and Social Security" (PASS) (Trappmann et al. 2010); (6) the German implementation of the "Programme for the International Assessment of Adult Competencies" (PIAAC) (Rammstedt 2013); (7) the German section of the "Survey of Health, Aging, and Retirement" (SHARE) (Börsch-Supan and Jürges 2005); and (8) the "Socio-Economic Panel" (SOEP) (Wagner, Frick, and Schupp 2007). To examine incentive effects on response rates and nonresponse bias, we observe the response rates for the cross-sectional surveys and the first waves of the panel surveys, unless noted otherwise. Additionally, we consider differences in the distributions of specific socio-demographic variables across experimental groups. We investigate the following three aspects of incentives: (1) the specific incentive form, that is, cash compared with nonmonetary incentives; (2) the monetary value of the incentive; and (3) prepaid, compared with conditional, incentives.
The remainder of this paper begins with a review of the literature regarding incentive effects. Next, we describe the incentive practices in the surveys listed above, the examined experiments, and our analytical approach. Subsequently, we present the experimental results. Finally, we summarize the results and discuss further research directions.

State of the Literature and Hypotheses

INTERNATIONAL RESULTS CONCERNING INCENTIVE EFFECTS
For cross-sectional surveys, Singer et al. (1999) show, in a review of existing studies, that incentives increase response rates across all survey modes. In addition, they show that this effect increases with the monetary value of the incentive. In contrast, Martin, Abreau, and Winters (2001) and Scherpenzeel and Toepoel (2012) find no significant increase in response rates with increasing incentives. For mail surveys, Church (1993) finds diminishing marginal returns of incentive value on response rates. Prepaid incentives lead to higher response rates than conditional incentives, for face-to-face surveys as well as mail and CATI surveys (see also Singer et al. 1999; Scherpenzeel and Toepoel 2012). Furthermore, cash incentives have a stronger effect on response

propensity than gifts, lottery tickets, or charitable donations (see also Singer et al. 1999; Simmons and Wilmot 2004). For web surveys, Göritz (2006) finds significant incentive effects on response rates but no effects of specific incentive characteristics.

Fewer experimental studies exist of panel surveys, which indicate similar results (see Laurie and Lynn [2009] for a review). Incentives yield higher response rates (Mack et al. 1998; Castiglioni, Pforr, and Krieger 2008; Zagorsky and Rhoton 2008). This effect increases with the monetary value of the incentive (Booker, Harding, and Benzeval 2011). Booker, Harding, and Benzeval (2011) also show that prepaid incentives, as well as cash incentives, affect response propensity more strongly than conditional incentives and non-monetary incentives. However, Castiglioni, Pforr, and Krieger (2008) find that conditional incentives lead to higher retention rates than prepaid incentives. Overall, there is "no evidence on the relative effectiveness of possible combinations over waves" (Laurie and Lynn 2009, 209). There is evidence that incentives positively affect retention rates in later waves, that increased incentives in later waves increase response rates, that lowering incentives in later waves does not reduce retention rates, and that incentives do not affect nonresponse bias (e.g., Goldenberg, McGrath, and Tan 2007; Laurie 2007; Jäckle and Lynn 2008).

Experimental studies also show that incentives can affect the sample composition. Singer et al. (1999, 225) state that incentives "may induce participation on the part of groups that would otherwise be underrepresented in the survey." That is, incentives can reduce sample selection bias. In their literature review, Simmons and Wilmot (2004) conclude that persons and households with low incomes, low education, and dependent children and young respondents and minority ethnic groups are more susceptible to incentives than other respondents (for theoretical analysis, see also Philipson [1997]).

INCENTIVES IN GERMANY

As in many industrialized countries, large-scale surveys in Germany suffer from decreasing response rates (e.g., Schnell 1997; Stoop et al. 2010). In response to this problem, large-scale German surveys have adopted incentive strategies. However, at the same time, there is continuing debate over whether incentives are the best strategy for German surveys. There is also ongoing debate over whether the results in the international literature are applicable to face-to-face surveys in Germany because most of the experimental evidence has come from the United States and the UK. Studies examining incentive effects in Germany have concentrated on mail surveys (e.g., Berger et al. 2009; Stadtmüller 2009; Becker and Mehlkop 2011) or web surveys (e.g., Göritz 2006, 2010). For example, Schupp (2012) recommends that rather than using incentives, surveyors should concentrate on appealing to respondents' sense of civic duty (see also Schnell 2012; Börsch-Supan, Krieger, and Schröder 2013). To our knowledge, incentive effects on German face-to-face surveys have been examined only by Castiglioni, Pforr, and Krieger (2008), Blohm and Koch (2013), Börsch-Supan,

Krieger, and Schröder (2013), and Schröder et al. (2013), and there has been no attempt to provide a general picture thus far.

HYPOTHESES

The question of cross-national differences in incentive effects has been raised before in the literature. From a theoretical perspective, Groves and Couper (1998), Hox and de Leeuw (2002), and Johnson et al. (2002) present reasons for lower response propensities in Germany than in the United States. Following the results of Singer et al. (1999), incentive effects are higher if response rates are lower; this observation could lead to the hypothesis that the incentive effects in Germany are greater than in the United States. If the results of Singer et al. are neglected, the differences in the mean propensities alone could result in different incentive effects on response rates, even if incentives affect the propensities equally, because the relationship between response propensity and response rate is necessarily nonlinear. Cantor, O'Hare, and O'Connor (2008) note that prepaid incentive effects can depend on differential address availability. In contrast to area- and address-based sampling frames in the United States, sampling frames for large-scale, face-to-face surveys in Germany are drawn mostly from resident registers. Therefore, prepaid incentives are personalized and can be tailored depending on the a priori known household size. Empirically, Mutti et al. (2014) find that in the ITC Four-Country Study, respondents from Australia more often complete the survey without cashing the checks used as prepaid incentives, compared to respondents from Canada and the United States. Additionally, whereas Bosnjak and Tuten (2003) find a positive effect of conditional PayPal incentives on response rates for a US sample, Göritz, Wolff, and Goldstein (2008) find a negative effect on response rates for almost identical incentives in Germany. Overall, this reasoning justifies examining whether incentive effects on response rates and nonresponse bias found in studies in the United States and the UK can be replicated in Germany.

We therefore build on the existing research and comprehensively examine whether the incentive effects found in face-to-face surveys in the United States and the UK and those using various modes also apply to large-scale, face-to- face surveys in Germany. Considering the lack of strong theoretical arguments for either a positive or negative difference between Germany and the known studies, we expect to find results similar to those in the international literature. Regarding the effects of incentives on response rates in cross-sectional surveys and in first waves and refreshment samples of panel studies, we expect that (1) incentives increase overall response rates; (2) incentive effects increase with the monetary value offered; (3) cash incentives result in higher response rates than non-monetary incentives; and (4) prepaid incentives have a stronger effect on response rates than conditional incentives.

Considering panel surveys, we expect that (5) an increase in incentives in later waves positively affects response rates in the wave of introduction; (6) continuously offered incentives lead to continuously increased retention rates; and (7) a decrease in incentives in later waves does not decrease retention rates.

Regarding nonresponse bias, we would expect to find socio-economically deprived respondents to be more susceptible to incentives than other respondents. However, because of data restrictions, we can examine only whether incentives affect bias in selected general socio-demographic variables, as described in the Methods section. For cross-sectional samples, we expect that (8) no heterogeneous susceptibility can be found for incentives regarding general socio-demographic variables. For panel studies, we expect that (9) incentives do not affect nonresponse bias in later waves.

Data and Methods

OVERVIEW OF STUDIES

Table 1 provides general descriptive information concerning the studies under consideration.[2] For large-scale, face-to-face surveys in Germany, the sample design and survey organization are particularly important because they largely determine the fieldwork strategies and thus the nonresponse.[3] All of the surveys considered here are government sponsored and conducted either by the "Infas Institute for Applied Social Sciences" (Infas) or "TNS Infratest Sozialforschung" (Infratest).

[2] For all studies, the sponsor and funding agency are identical and displayed in the acknowledg¬ments. The conductors are displayed in table 1. The populations of all studies are shown in table 1. The geographic location is Germany for all studies but the pairfam and the PIAAC pilots. For pairfam the geographic area is the combined area of Bremen, Chemnitz, Mannheim, and Munich. For PIAAC the area is the combined area of the states Hamburg, Schleswig-Holstein, Bayern, Sachsen, and Thüringen.

[3] See Smith (1978) and Diekmann (2011) for house effects. See Häder and Gabler (2003) for sample designs in Germany. The origins of the sample frames are shown in table 1. A short description of the ADM procedure can found in Häder and Gabler (2003). GIP and PASS use a modified version of the ADM procedure, as they create a list of all households in the selected geographical units. Regarding the sample design, the sample for ALLBUS is drawn as a two-stage random person sample from the resident register. Sample units from new states of Germany are oversampled. The sample for GIP is drawn as a three-stage random sample with an area frame of residential units (modified "ADM" procedure). For NEPS, the pilot study sample, which is analyzed in this paper, is a random sample from the unused addresses of the refreshment and augmentation samples of the first NEPS wave. It is drawn as two-stage random samples from the resident register. For pairfam, the pilot study, which is analyzed here, is drawn as a single-stage random sample from the resident register of four cities in Germany. The sample is stratified by the three birth cohorts. The cities are a convenience sample. For PASS, the sample for the first wave consists of two separately drawn subsamples. The first subsample is a two-stage sample of households that receive unemployment benefits, drawn from the recipient register of the Federal Employment Agency. The second subsample is a two-stage random sample of households, drawn from a list of residential units, with stratification by social status. For PIAAC, the sample for the pilot study, which is analyzed here, is a two-stage random sample drawn from the resident register. For SHARE, the sample of the first wave is drawn as a two-stage random sample from the resident register (Klevmarken, Hesselius, and Swensson 2005). For SOEP, the sample for the first waves in 1984 and all refreshment samples of the general population (e.g., Sample E in 1998, Sample F in 2000, Sample H in 2006, Sample I in 2009, Sample J in 2011, and Sample K in 2012) are drawn as two-stage random household samples with an area frame of residential units ("ADM" procedure)

Table 1. Overview of Investigated Surveys

Survey	Focus	Panel/RCS[a]	Waves	Mode[b]	Population	Sample frame[c]	Survey org.
ALLBUS	General	RCS	Since 1980, two years interval	CAPI	Adults residing in priv. HH[d]	Resident register	Infratest
GIP[e]	Reforms	Panel	Recruitment in 2012	CAPI	Age 16-75 residing in priv. HH[d]	ADM	Infratest
NEPS[f]	Education	Panel	Since 2009, one year interval	CATI/CAPI	Birth cohorts 1956-86, in priv. HH[d]	Resident register	Infas
pairfam[g]	Family	Panel	Since 2008, one-year interval	CAPI/CASI	Birth cohorts 1971-73, 81-83, 91-93, in priv. HH[d]	Resident register	Infratest
PASS	Labor-market reforms	Panel	Since 2007, one year interval	CATI/CAPI	Adults residing in priv. HH[d]	Unempl. benefit recipient register / ADM	Infratest[h]
PIAAC[i]	Competencies,[j] education	CS	2011-2012	CAPI/CASI/ SAQ	Age 16-65, residing in priv. HH[d]	Resident register	Infratest

Continued

Table 1. Continued

Survey	Focus	Panel/RCS[a]	Waves	Mode[b]	Population	Sample frame[c]	Survey org.
SHARE[k]	Health, aging	Panel	Since 2004, two-year interval	CAPI	Persons in HH[d] with at least one German speaker aged 50+	Resident register	Infas
SOEP	General	Panel	Since 1984, one year interval	CAPI/PAPI/ SAQ	HH[d] in priv. residencies	ADM[l]	Infratest

[a]RCS: repeated cross-section; CS: cross-section.
[b]CAPI: computer-assisted personal interview, CASI: computer-assisted self-interview, CATI: computer-assisted telephone interview, SAQ: self-administered questionnaire.
[c]ADM: area-frame of residential units.
[d]HH: household.
[e]Recruitment interview of GIP, online fieldwork bi-monthly.
[f]Only cohort 6 of NEPS is considered (Allmendinger et al. 2011 ; Kleinert et al. 2011).
[g]Only incentives for “anchor” persons are shown.
[h]Waves 1-3 were conducted by Infratest, and wave 4 was conducted by Infas.
[i]Only the German part of PIAAC is considered (Zabal et al. 2014).
[j]Self-administered assessment.
[k]Only the German part of SHARE is considered (de Luca and Lipps 2005; Lynn et al. 2013).
[l]Subsamples of SOEP covering special populations use registers.

Table 1 also indicates the methodological differences among these studies: PIAAC is a cross-sectional survey, whereas ALLBUS is a repeated crosssectional survey. All of the other surveys are panels. Regarding the sampling, GIP and SOEP are area-based household samples,[4] PASS uses a register-based sample and a sample based on a list of residential units, and all of the other studies are register-based samples. The survey mode varies as well: NEPS and PASS use a mixed CAPI/CATI design, and pairfam uses a mixed CAPI/CASI design. SOEP used PAPI as the default mode until the late 1990s and CAPI thereafter, allowing for SAQ in experienced panel households.[5] All of the other studies use a CAPI mode. In GIP, respondents are recruited by CAPI but then are interviewed online for the panel.
We selected this group of surveys for our inquiry because all of these surveys conduct face-to-face interviews, they can be considered large-scale studies with respect to their sample sizes and relevance for the German social science community, and the fieldwork is conducted by a professional survey organization. For comparison, table 2 shows the harmonized response rates of the most recent samples of the cross-sectional surveys and the most recent refreshment samples of the panel surveys.
For the fourth wave of SHARE-Germany, which is the wave considered in this paper, no response rates are available because the fieldwork did not proceed as planned (Börsch-Supan, Krieger, and Schröder 2013). The remaining response rates are calculated as RR1 rates, following the definitions of AAPOR (2011). In addition to the incentive experiments analyzed in this paper, the studies could use different incentive strategies, as shown in table 3.
All of the surveys examined in this paper have adopted incentive strategies. The SOEP survey began with lottery coupons in its first wave in 1984. Cash incentives of commemorative coins were first used in ALLBUS 2002. Pure- cash incentives were adopted in 2009 in NEPS.[6] Currently, most of the studies use conditional cash incentives, rather than other forms of incentives. The monetary value ranges from €5 to €50.

EXPERIMENTAL DESIGNS

In the ALLBUS survey of 2010, two experiments were conducted (Wasmer et al. 2012; Blohm and Koch 2013). The first experiment was implemented during the main fieldwork period of the survey. The pool of addresses was split randomly into

[4] Two of the three migrant samples of SOEP are based on person registers.
[5] In the SOEP, personal interviews are conducted whenever possible (Hanefeld 1987; Haisken-De New and Frick 2005). Since 1998, SOEP has been gradually replacing PAPI (personal paper and pencil interviewing) with CAPI (computer-assisted personal interviewing) as the predominant mode of data collection. If respondents in the old samples A though H refuse to participate in the personal interview, the fieldwork organization offers mailed questionnaires as a means of refusal conversion. Hence, some experienced panel households may also use self-administered mailed questionnaires (SAQ).
[6] Pure-cash incentives were introduced in 2008 in the German portion of the European Social Survey (Stoop et al. 2010; Keil and van Deth 2012).

Table 2. Response Rates (RR1)

Survey	Year	Response rate (%)
ALLBUS	2010	33.1
GIP	2012	42.9
NEPS	2011/12	33.1
pairfam	2008/09	34.3
PASS[a]	2011	28.2
PIAAC	2011/12	53.3
SOEP	2011[b]	33.1

Note.—The response rates are calculated as RR1 rates following the definitions of AAPOR (2011).
[a]Only the refreshment sample in wave 4 of PASS.
[b]Wave 1 of refreshment sample J (Siegel, Huber, and Bohlender 2012).

an experimental group of 2,592 addresses and a control group of 3,888 addresses. In the experimental condition, respondents received €10 in cash, conditional on participation and announced in an advance letter, whereas respondents in the control group received no incentive. The second experiment was implemented in a separate address pool and was issued to the field in the second half of the fielding period. Here, the first experimental group of 972 addresses was offered €20 in cash, conditional on participation, and the second experimental group of 972 addresses was offered €10 in cash, conditional on participation. In both conditions, the incentives were announced in an advance letter. To avoid confounding area and interviewer effects with the incentive effect, for both experiments, the treatment conditions were randomly assigned in primary sampling units, and the interviewers worked addresses from both treatment conditions. The interviewers knew the treatment condition of each respondent.

The GIP panel survey conducted an incentive experiment during its recruitment survey in 2012. The gross address sample for the experiment consisted of 3,900 household addresses allocated to interviewers during the first fieldwork phase. The first experimental group of 1,464 households received €5 in prepaid cash, which was mailed with the advance letter. The second experimental group of 2,436 households received €10 in cash, conditional on participation. Here, the incentive was announced in the advance letter. Addresses for which the prior address listing had yielded no name on the doorbell or mailbox were excluded from the experiment and received €10 in cash, conditional on participation. The value of the incentives was chosen such that the overall costs of each incentive condition were approximately equal, assuming a 50 percent response rate (AAPOR RR2). Cases were randomly allocated to the experimental groups. Interviewers worked across both incentive conditions and were informed regarding the condition to which a household belonged.

In the pilot study of NEPS wave 1 in 2009 (Infas 2009), the incentive amount varied experimentally. The gross sample of the pilot study was split randomly into two groups. The experimental group (N = 861) received €10, conditional

Table 3. Incentives Used in Normal Operations

Survey	Year	Incentive
ALLBUS	2002	€10 commemorative coin, cond.
	2012	€10 cash, cond.
GIP[a]	2012	€10 cash, cond., €5 cash prepaid, see table 5
NEPS[b]	2007-2008	Lottery coupon w/social sponsor, cond. for CATI mode, €15 cash, cond. in CAPI mode
	2009-2010	€10-€50 cash,[c] cond.
	2010-2011	€25 cash, cond.
	2011-2012	€20 cash, cond.
pairfam	2008-2009	€10 cash, cond.[d]
PASS	2007	€1.50 lottery coupon, cond. for household
	2008	€5 lottery coupon, cond. for household
	2009	€5 lottery coupon, cond. for household of refreshment sample and half of panel sample, €10 cash, prepaid for household of other half of panel sample
	2010	€10 cash, prepaid for each person for panel sample w/ participation in 2009, €10 cash, cond. for each person for refreshment sample and panel sample w/o participation in 2009
	2011	€10 cash, prepaid for each person for panel sample, €10 cash, cond. for each person for refreshment sample
PIAAC[e]	2011-2012	€50 cash, cond.
SHARE[f]	2004	Low-value gift, prepaid
	2010	€10 cash, cond.
SOEP[g]	1984-2007	Lottery coupon, cond.
	2008-2012	Lottery coupon, prepaid
	2009-2012	€5 cash cond. for household and €10 cash cond. for each person in new refreshment samples (samples I, J, and K)

[a]For face-to-face recruitment interviews of GIP, further incentives for online participation.
[b]Only cohort 6 of NEPS is considered (Allmendinger et al. 2011; Kleinert et al. 2011).
[c]Incentives were increased because of slow progress in the field: €10 Nov 10-May 3; €50 May 4-July 30.
[d]Only incentives for "anchor" persons are shown.
[e]Only the German part of PIAAC is considered (Zabal et al. 2014).
[f]Only the German part of SHARE is considered (de Luca and Lipps 2005; Lynn et al. 2013).
[g]See Schröder et al. (2013).

on participation, whereas the control group (N = 908) received no incentive. At the end of the field period after only four weeks, only 190 interviews were realized; thus, the response rate was low overall.
Another survey experiment was conducted in the pretest study of pairfam (Castiglioni, Pforr, and Krieger 2008). The pretest was conducted in 2005 and

was repeated in half-yearly intervals for two more waves. A random sample of 1,664 persons in three birth cohorts was drawn from the resident registers of four German cities. The sample was split into three treatment groups. The first experimental group (N = 576) received a prepaid €10 voucher incentive, the second group (N = 562) received a €10 voucher, conditional on participation, and the control group (N = 526) received no incentive. The assignment was held constant across all three panel waves. The interviewers were blinded to the treatment condition of each sample unit.[7]

In the PASS survey, an experiment was implemented in the third wave, fielded in 2009. The sample of households that had participated in at least one of the earlier waves was randomly split into two treatment groups. The first experimental group of 5,349 households was given the same incentives as in the previous wave (see table 3): a lottery ticket with a social sponsor worth €5 per person, conditional on participation. For the second experimental group of 5,362 households, the incentives were increased to a prepaid €10 cash incentive per household. For both groups, the incentives were paid at the household level.

In the German field test of PIAAC in 2010, an incentive experiment was conducted (Zabal et al. 2014). The field-test design was similar to the main study design described in table 1. The target population was the same, but the sampling area was restricted to five federal states. The gross sample was split into three treatment groups. The first experimental group (N = 1,384) was assigned a conditional incentive of €50 in cash, and the second experimental group received (N = 1,391) €25 in cash, conditional on participation. The third experimental group (N = 674) received a commemorative silver coin worth €10 with the emblem of the 2006 World Cup soccer tournament.

The German division of SHARE conducted an incentive experiment in the refreshment sample in the fourth wave in 2010 (Börsch-Supan, Krieger, and Schröder 2013). The experiment consisted of four treatment groups. The first experimental group of 750 persons received a prepaid €40 cash incentive, and the second group of 750 persons received a prepaid €20 cash incentive. The third group of 1,375 persons received a prepaid €10 cash incentive, and the control group of 1,025 persons received no incentive. The analysis sample had to be restricted to the 2,241 cases (57.5 percent of 3,900 total) that had been entirely worked by the survey agency.[8] A case was defined as processed if it resulted in an interview, received a hard or soft refusal, or was visited eight

[7] The conditional voucher incentive was sent by the field management team after the completed interview, without interference from the interviewer.

[8] The gross sample consisted of 3,900 addresses in 156 sample points. Experimental conditions were randomly allocated in sample points. The households were assigned to interviewers in the fourth calendar week of January 2010, and the first interview was conducted in the seventh calendar week, in mid-February 2010 (Malter 2013). Fieldwork in the refreshment sample was aborted in August 2012 because of interviewer inactivity and slow progress. The agency was advised to focus on the panel sample. Overall, 1,900 cases out of 3,900 (48.7 percent) were contacted at least once (Börsch-Supan, Krieger, and Schröder 2013). The severity of selection bias in the contact process is unclear.

times. Because of these restrictions, nonstandard response rates are reported here, defined as the ratio of households with at least one complete interview to the number of processed addresses.
In the SOEP study, two experiments regarding incentive effects were conducted. The first experiment was implemented in the "Innovation Sample," drawn in 2009 (Richter and Schupp 2012; Schröder et al. 2013). The outcomes were measured for the 2009 wave and the subsequent 2010 wave. The experiment consisted of four treatment groups with approximately 1,240 households per treatment, all of which received incentives conditional on participation. The first experimental group ("moderate cash") was promised €5 in cash for the household head and €10 in cash for each individual respondent. The second group ("low cash") received €5 in cash for the household head and €5 cash for each individual. The third group could choose between a lottery ticket and the "low cash" incentive. Finally, the control group received the standard SOEP incentive in the form of a lottery ticket for a charity worth €5 for each participating household member. The experimental variation was removed in the following wave in 2010, in which all of the households were promised the "low cash" incentive.
The second SOEP experiment was implemented with the 1,604 households of the ongoing panel samples A-H, based on the gross sample of 2011 (Schröder et al. 2013). All of the households were sent the usual charity lottery ticket (value €5) before participation. Additionally, the experimental groups received €5 cash for the household head and €10 cash for each household member, conditional on the participation of each individual respondent in the household.

MEASUREMENT AND ANALYSIS DESIGN

To test our hypotheses, we examine experimental variations in the offered incentives in cross-sectional studies and in first waves, new refreshment samples, and subsequent waves of panel studies. For all of the experiments described above, except the SHARE experiment, we analyze the original contact record data and the respective realized samples. The information for SHARE is taken from Börsch-Supan, Krieger, and Schröder (2013) and is reported here for comparison with the other experiments.[9]
For the cross-sectional studies and for the first waves and refreshment samples in panel studies, we consider AAPOR RR1 response rates. In addition, for pairfam, PASS, and SOEP, we examine analogously defined retention rates in later panel waves. More specifically, we examine the differences in response and retention rates across experimental groups with the respective x^2 statistics. For multi-arm experiments, we report differences in comparison with one reference group.

[9] The SHARE organization did not provide any further information and referred to the informa¬tion published by Börsch-Supan, Krieger, and Schröder (2013) and Malter and Börsch-Supan (2013). With the available information, it is impossible to compute a standard AAPOR response or participation rate for the refreshment sample in the fourth wave of SHARE.

Considering the relative nonresponse bias across incentive groups, we compare the distributions of variables across the experimental conditions that are available for all of the surveys. We do not use external data as a reference for the degree of nonresponse bias because the reference populations differ widely across studies. For surveys for which the sampling units and, therefore, the basis of response rates are households, we examine household size, municipality size, and the proportion of households in the eastern states of Germany. If the sampling units are persons, we additionally examine age, gender, and education. The selection criteria for these variables are eligibility and measurement comparability across all of the surveys.[10]

In the pairfam and PIAAC pretests, because of the restricted samples, the East-West comparisons are not applicable. Additionally, for pairfam, the distributions in municipality size are not comparable because this study was conducted only in large cities. For SHARE, Börsch-Supan, Krieger, and Schröder (2013) do not report any information regarding the variables examined here. Therefore, SHARE is excluded from the nonresponse bias analyses (see footnote 9).

For differences in age and household size, we compare the means of the continuous variables with the respective t-statistics. For the categorical variables gender and proportion of households in East Germany, we compare the respective proportions across experimental conditions. For education, we compare across experimental conditions the proportion of respondents with an academic degree that allows for access to tertiary education. Municipality size is derived from the population in the BIK region containing the municipality in which the respondent unit resides (Statistisches Bundesamt 2014). This information is coded as a categorical variable with the following groups: under 50k, 50k-under 100k, 100k-under 500k, and 500k or greater. We compare the proportions of these categorical variables across experimental conditions and report the respective x^2 statistics.

Because of the heterogeneity in experimental designs, not all of the hypotheses can be tested with all experiments. Table 4 shows the stated hypotheses and the relevant experiments.

The experimental designs in GIP and PASS are problematic given our hypotheses because the monetary value variation is confounded by the conditionality or the cash payment variation. However, with the GIP experiment, we can examine the experimental conditions that lead to a higher response rate and to a smaller effect on nonresponse bias while

[10] The wording of the questions and answers is laid out in the online appendix. Regarding filter¬ing and otherwise intentional respondent selection, the response and retention rates in tables 5 and 6 and the differences in tables 7 and 8 are based on the sample of all eligible cases for the respec¬tive waves. For tables 7 and 8, units with item nonresponse for those indicators that are not taken from frame information are disregarded. The resulting sample sizes for the analyses of incentive effects are shown in tables 5 and 6. Sampling error is reflected in the reported t and x^2 statistics. The response and retention rates in tables 5 and 6, and the differences and t and x^2 statistics in tables 7 and 8, are computed assuming simple random samples.

Table 4. Relationships of Hypotheses to Experiments

Hypothesis	Experiments
(1) Incentives increase response rates	ALLBUS, NEPS, pairfam, SHARE
(2) Effect increases with monetary value	ALLBUS, PIAAC, SHARE, SOEP (2009)
(3) Effect stronger for cash vs. other forms	SOEP (2009)
(4) Effect stronger for prepaid vs. conditional incentives	GIP
(5) Increase in incentives in later waves increases response rates	PASS, SOEP (2011)
(6) Continuous incentives continuously increase retention rates	pairfam
(7) Decreasing incentives later does not decrease retention rates	SOEP (2009)
(8) Incentives do not affect nonresponse bias regarding socio-demographic variables in cross-sectional samples	ALLBUS, GIP, NEPS, pairfam, PASS, PIAAC, SOEP
(9) Incentives do not affect nonresponse bias regarding socio-demographic variables in subsequent panel waves	pairfam, SOEP (2009)

holding costs constant. Furthermore, with the PASS experiment, we can examine whether a relative increase in incentive value in a later panel wave increases response rates. In addition, the experiments in NEPS, pairfam, and PIAAC were conducted in pilot studies. However, the numbers of observations in all of the experimental groups are sufficient for testing our hypotheses.

Results

The experimental variations and the respective response rates for the crosssectional studies, first waves, and refreshment samples in panel studies are shown in table 5.

Supporting hypothesis 1, we find that offering incentives, compared with not offering incentives, significantly increases the response rates in the first ALLBUS experiment and the NEPS and SHARE experiments. In the first wave of pairfam, offering €10 vouchers does not increase response rates significantly. Considering hypothesis 2, we find an increase in response rates with increasing monetary value in the PIAAC and SHARE experiments. The differences in response rates in the second ALLBUS and the “moderate-cash” and “low-cash” groups in the first SOEP experiment are not significant. This result can be interpreted as weak support for hypothesis 2. In agreement with hypothesis 3, the results of the first SOEP experiment

show that the response rate in the “low-cash” group is higher than in the control group. This difference is significant at the 10 percent level. Confirming hypothesis 4, the GIP experiment shows that prepaid incentives lead to significantly higher response rates than conditional incentives. Considering that the monetary value of the prepaid incentive is lower than the value of the conditional incentive, this is strong evidence that prepaid incentives increase response rates.

Regarding incentive effects on panel studies, our experiments yield several results. Considering hypothesis 5, the change from a conditional lottery ticket incentive worth €5 to a prepaid €10 cash incentive in a later wave in the PASS experiment significantly increases the response rate in that wave. However, an increase in incentive value in a later wave in the second SOEP experiment does not increase the response rate significantly. Note that the PASS experiment was implemented in the third wave, whereas the second SOEP experiment was introduced in the 28th wave. Considering this difference, our results show that a change in incentives that is expected to increase response rates in cross-sectional surveys also increases response rates in later panel waves. This result can be interpreted as providing some support for hypothesis 5.

Table 6 shows the incentive effects on retention rates in subsequent waves in pairfam and in the first SOEP experiment. Regarding hypothesis 6, the pairfam experiment shows that compared with the control group without incentives, offering a conditional €10 voucher across three waves increases the retention rate, conditional on participation in the previous waves 2 and 3. Offering a prepaid €10 voucher significantly increases the retention rate only in the third wave. Considering the incentive effects on the first wave and the effect instability concerning incentive form, these results provide weak support for hypothesis 6.

Confirming hypothesis 7, the first SOEP experiment shows that decreasing the “moderate cash” incentive in the first wave to the “low cash” incentive in the second wave does not lead to a decreased retention rate in the second wave, compared with consecutively offering the “low cash” incentive.

In addition to the effects on response and retention rates, the experiments examined whether incentives affect sample composition. Table 7 shows the differences in the socio-demographic variables on the respondent level in the cross-sectional studies and in the first waves and refreshment samples in the panel studies.

The comparisons of variable distributions across experimental conditions show mixed results. The differences in mean age and the proportions of female respondents do not differ significantly across incentive conditions in any of the seven comparisons. The differences in proportions of respondents with access to tertiary education are significantly different only in one comparison group in pairfam (conditional €10 voucher group: 52.6 percent

Table 5. Effects of Incentives on Response Rates in Cross-Sectional Surveys and First Waves and Refreshment Samples in Panel Surveys

Study	Year	Incentive	N	Response rate (%)	x^2
ALLBUS[a]	2010	€10 cash, cond.	2,592	25.8	11.50**
		No incentive	3,888	22.2	
	2010	€20 cash, cond.	972	32.2	0.00
		€10 cash, cond.	972	32.2	
GIP[b]	2012	€5 cash, prepaid	1,464	44.9	26.40**
		€10 cash, cond.	2,436	36.2	
NEPS	2009	€10 cash, cond.	861	13.4	11.97**
		No incentive	908	8.3	
pairfam	2005/06	€10 voucher, cond.	526	41.7	0.01
		€10 voucher, prepaid	559	41.0	0.01
		No incentive	576	41.4	
PASS	2009	€10 cash, prepaid	5,362	72.0[c]	64.67**
		€5 lottery ticket, cond.	5,349	64.7[c]	
PIAAC	2010	€50 cash, cond.	1,384	40.6	43.44**
		€25 cash, cond.	1,391	34.9	17.76**
		€10 commemorative coin, cond.	674	25.5	
SHARE	2010	€40 cash, prepaid	456	54.2[d]	94.15**
		€20 cash, prepaid	436	40.8[d]	19.75**
		€10 cash, prepaid	801	38.3[d]	14.43**
		No incentive	548	27.4[d]	

Continued

Table 5. Continued

Study	Year	Incentive	N	Response rate (%)	x^2
SOEP	2009	€5 cash for HH, €10 cash for respondent, cond.	1,241	32.9	3.53*
		€5 cash for HH, €5 cash for respondent, cond (A)	1,240	33.2	4.13+
		€5 lottery ticket for respondent, cond. (B)	1,243	29.3	
		Free choice between A and B, cond.	1,240	30.8	0.67
	2011	€5 lottery ticket, prepaid; €5 cash for HH, €10			
		Cash for respondent, cond.	803	89.5[c]	0.25
		€5 lottery ticket for HH, prepaid	801	88.8[c]	

Note. – AAPOR RR1 rates reported.
[a]Only the main fielding period.
[b]Households without names on address frame were excluded from the experiment.
[c]Response rates conditional on participation in at least one previous wave.
[d]Response rate not according to AAPOR standard.
**p < .01; *p < .05; +p < .1

Table 6. Effects of Incentives on Retention Rates in Subsequent Waves in Panel Surveys

				Retention rate[b]			
Study	Year	Incentive	N[a]	W2 (%)	x^2	W3 (%)	x^2
pairfam	2005-2006	€10 voucher, cond.	526	79.6	3.70+	87.8	5.35*
		€10 voucher, prepaid	559	71.9	0.00	89.6	7.59**
		No incentive	576	71.7		78.2	
SOEP	2009	€5 cash for HH. €10 cash for respondent, cond.	1.241	76.0	4.53*		
		€5 cash for HH. €5 cash for respondent, cond. (A)	1.240	71.4	0.52		
		€5 lottery ticket for respondent, cond. (B)	1.243	68.9			
		Free choice between A and B. cond.	1.240	71.1	0.39		

Note. – AAPOR RR1 rates reported.
[a]Sample sizes in respective first waves.
[b]Retention rates are conditional on participation in the previous wave.
**p < .01; *p < .05; +p < .1

Table 7. Effects of Incentives on Differences in Respondent Unit Characteristics in Cross-Sectional Face-to-Face Surveys

			Differences in					
Study	Year	Incentive	Mean age (t)	% female (x^2)	% high educ (x^2)	Mean HH-size (t)	Municip. size (x^2)	% east (x^2)
ALLBUS[a]	2010	€10 cash, cond.	1.82^{+}	0.04	1.36	0.03	1.13	0.38
		No incentive						
	2010	€20 cash, cond.	1.11	2.45	0.17	1.13	3.71	0.16
		€10 cash, cond.						
GIP[b]	2012	€5 cash, prepaid				-1.09	1.78	0.00
		€10 cash, cond.						
NEPS	2009	€10 cash, cond.	0.37	0.15	0.30	-1.27	5.38	1.57
		No incentive						
pairfam[c]	2005/06	€10 voucher, cond.	-0.92	1.35	5.98*	-1.70^{+}		
		€10 voucher, prepaid	0.22	2.05	0.93	-0.69		
		No incentive						
PASS	2009	€10 cash, prepaid				0.70	70.09**	22.68**
		€5 lottery ticket, cond.						
PIAAC[d]	2010	€50 cash, cond.	0.13	1.35	0.42	0.62	2.97	
		€25 cash, cond.	-0.51	1.79	2.02	0.26	3.29	
		€10 commemorative coin, cond.						
SOEP	2009	€5 cash for HH, €10 cash for respondent, cond.				-1.58	2.56	3.37^{+}
		€5 cash for HH, €5 cash for respondent, cond. (A)				-1.37	1.14	2.89^{+}
		€5 lottery ticket for respondent, cond. (B)						
		Free choice between A and B, cond.				-2.27*	1.45	1.69
	2011	€5 lottery ticket, prepaid; €5 cash for HH, €10 cash for respondent, cond.				-0.77	2.95	1.69
		€5 lottery ticket for HH, prepaid						

[a]Only the main fielding period.
[b]Households without names on address frame were excluded from the experiment.
[c]Differences in municipality size and % east not applicable.
[d]Differences in % east not applicable.
$^{**}p < .01$; $^{*}p < .05$; $^{+}p < .1$

versus control group: 38.5 percent).[11] Considering the household-level variables, we find that the differences in mean household size across 13 experimental groups are significant only for one comparison group in the first SOEP experiment ("choice" group: 2.4 versus "lottery group": 2.2). The distributions across the four municipality size categories differ significantly only in the PASS experiment. Here, prepaid cash incentives work better than the conditional lottery incentives in metropolitan areas (500k or greater, (34.0 versus 30.5 percent) and vice versa in rural areas (under 50k, 20.2 versus 28.1 percent). The proportions of households in East Germany differ significantly in the PASS experiment (lottery: 27.1 percent, cash: 32.2 percent) and in two comparison groups in the first SOEP experiment (21.2 percent for "moderate cash," 20.8 percent for "low cash," and 15.9 percent for "lottery"). Overall, we find that, in some studies, specific groups of respondents are more responsive to incentives than other respondents, but there are no significant differences between the variables examined in multiple studies. These results lend support for hypothesis 8.

Regarding hypothesis 9, table 8 shows the differences in the socio-demographic variables on the respondent level in subsequent waves for pairfam and the first SOEP experiment. Mean age, mean household size, the distribution across the four municipality size categories, and the proportion of respondents in East Germany do not differ significantly across the examined experimental contrasts. However, in pairfam, the proportions of female respondents differ significantly across both contrasts with respect to the control group in both subsequent waves (W2: conditional 54.9 percent, prepaid 56.1 percent, control 49.3 percent; W3: conditional 58.2 percent, prepaid 59.2 percent, control 46.1 percent). Furthermore, the proportions of respondents with access to tertiary education differ significantly for one contrast in both subsequent pairfam waves (W2: conditional 47.1 percent versus control 66.0 percent; W3: conditional 44.6 percent versus control 65.8 percent). We interpret these results as weak evidence against hypothesis 9 because the significant differences in gender and education were found in both subsequent waves in pairfam. Because this experiment is designed as a pretest, conducted only in four cities with a restricted population, the transferability to general population surveys is limited.

Discussion and Conclusion

The results partly confirm our hypotheses regarding the effects of incentives on response rates and retention rates in large-scale, face-to-face surveys in Germany. Incentives increase response rates. We find weak support for an

[11] This result seems to contradict the literature, which shows that socio-economically deprived respondents are more susceptible to incentives.

Table 8. Effects of Incentives on Differences in Respondent Unit Characteristics in Subsequent Waves in Panel Surveys

			Differences in					
Study	Wave	Incentive	Mean age (t)	% female (x^2)	% high educ (x^2)	Mean HH-size (t)	Municip. size (x^2)	% east (x^2)
pairfam“	2	€10 voucher, cond.	0.81	3.15^{+}	7.59*	0.60		
		€10 voucher, prepaid	-0.25	4.23*	1.55	-0.26		
		No incentive						
	3	€10 voucher, cond.	0.98	3.94*	8.13*	0.34		
		€10 voucher, prepaid	1.07	4.45*	2.14	0.63		
		No incentive						
SOEP(2009)	2	€5 cash for HH, €10 cash for respondent, cond.				-0.55	0.98	0.37
		€5 cash for HH, €5 cash for respondent, cond. (A)				-0.74	0.71	
		€5 lottery ticket for respondent, cond. (B)						0.21
		Free choice between A and B, cond.				-1.35	0.96	0.54

[a] Differences in municipality size and % east not applicable.

$^{*}p < .05$; $^{+}p < .1$

increase in incentive effects with monetary value. Cash incentives have a stronger positive effect on response rates than lottery tickets. Assuming an RR2 response rate of 50 percent, prepaid incentives increase response rates more cost-efficiently than conditional incentives. However, earlier research has indicated that prepaid incentives can cause distrust with some respondents (Börsch-Supan, Krieger, and Schröder 2013). Moreover, we examine the incentive effects in panel studies from a longitudinal perspective. A later rise in incentive value increases response rates in the wave of implementation. We find weak support for a constant increase in retention rates when incentives are consistently offered across multiple waves. Decreasing incentives in a later wave does not decrease retention rates.

In addition to the effects of incentives on response rates, the experiments provide information concerning the effects on nonresponse bias. For crosssectional samples, our results indicate that incentives do not differentially affect nonresponse bias regarding the socio-demographic variables considered here. Regarding the influence on nonresponse bias in subsequent waves in panel surveys, we find mixed results. From a conservative perspective, we found in some studies that specific respondent groups are more responsive to incentives than other respondents. Further research is necessary to assess whether incentives improve or aggravate nonresponse bias beyond key sociodemographic variables.

As our study focuses on specific incentive effects on nonresponse, attrition, and nonresponse bias in Germany, we had to put several otherwise interesting aspects in the rear. First, the scope of our study did not permit an investigation of specific incentive effects on mail, telephone, or web surveys. Second, some experiments permit only limited inferences because multiple variations are confounded. Third, our focus on differences between monetary incentives and lottery tickets ignored how symbolic in-kind incentives affect response rates and nonresponse bias. Fourth, we might have overlooked incentive effects on nonresponse bias concerning other variables in addition to those examined here. In addition, our method of analysis does not allow us to infer whether incentives reduce or increase nonresponse bias. Fifth, we ignore incentive effects on measurement error and item nonresponse, which can lead to biased sample distributions. Sixth, our results provide no evidence of whether incentives are cost efficient because alternative methods for increasing response rates and for affecting nonresponse bias are not examined in this paper. Future research should address these issues in greater detail.

The results show that from a cross-sectional perspective, incentive effects in general and the effects of cash and prepayment in particular also apply in Germany. From a panel perspective, the effects of incentives in later panel waves on response rates and the stability of retention rates after decreased incentives can also be transferred to Germany. Therefore, findings in the international literature regarding incentive effects are at least partly generalizable to Germany. However, with our data, the greater incentive susceptibility

of socio-economically deprived respondents can be neither confirmed nor rejected for Germany.

Supplementary Data

Supplementary data are freely available online at http://poq.oxfordjournals.org/.

References

AAPOR. 2011. Standard Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th ed. Available at http://www.aapor.org/AAPORKentico/AAPOR_Main/media/MainSiteFiles/StandardDefinitions2011_1.pdf.

Allmendinger, Jutta, Corinna Kleinert, Manfred Antoni, Bernhard Christoph, Katrin Drasch, Florian Janik, Kathrin Leuze, Britta Matthes, Reinhard Pollak, and Michael Ruland. 2011. “Adult Education and Lifelong Learning.” Zeitschrift für Erziehungswissenschaft Sonderheft 14:283-99.

Atrostic, Barbara Kathrin, Nancy Bates, Geraldine Burt, and Adriana Silberstein. 2001. “Nonresponse in US Government Household Surveys: Consistent Measures, Recent Trends, and New Insights.” Journal of Official Statistics 17:209-26.

Becker, Rolf, and Guido Mehlkop. 2011. “Effects of Prepaid Monetary Incentives on Mail Survey Response Rates and on Self-Reporting about Delinquency: Empirical Findings.” Bulletin of Sociological Methodology 111:5-25.

Berger, Fred, Urs Grob, Helmut Fend, and Wolfgang Lauterbach. 2009. “Möglichkeiten zur Optimierung der Rücklaufquote in postalischen Befragungen: Bericht über die Vorstudie zum Forschungsprojekt LifE.” Zeitschrift für Soziologie der Erziehung und Sozialisation 25:99-107.

Blohm, Michael, and Achim Koch. 2013. “Respondent Incentives in a National Face-to-Face Survey: Effects on Outcome Rates, Sample Composition, and Fieldwork Efforts.” Methoden, Daten, Analysen (MDA) 7:89-122.

Blom, Annelies G., Christina Gathmann, and Ulrich Krieger. Forthcoming. “Setting Up an Online Panel Representative of the General Population.” Field Methods 27(4).

Blom, Annelies G., Annette Jäckle, and Peter Lynn. 2010. “The Use of Contact Data in Understanding Cross-National Differences in Unit Nonresponse.” In Survey Methods in Multinational, Multiregional, and Multicultural Contexts, edited by Janet A. Harkness, Michael Braun, Brad Edwards, Timothy P. Johnson, Lars E. Lyberg, Peter Ph. Mohler, BethEllen Pennell, and Tom W. Smith, ch. 18. Hoboken, NJ: Wiley.

Booker, Cara L., Seeromanie Harding, and Michaela Benzeval. 2011. “A Systematic Review of the Effect of Retention Methods in Population-Based Cohort Studies.” Public Health 11:1-12.

Börsch-Supan, Axel, and Hendrik Jürges, eds. 2005. The Survey of Health, Aging, and Retirement in Europe-Methodology. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).

Börsch-Supan, Axel, Ulrich Krieger, and Mathis Schröder. 2013. Respondent Incentives, Interviewer Training, and Survey Participation. SHARE Working Paper, München.

Bosnjak, Michael, and Tracy L. Tuten. 2003. “Prepaid and Promised Incentives in Web Surveys: An Experiment.” Social Science Computer Review 21:208-17.

Brick, J. Michael, and Douglas Williams. 2013. “Explaining Rising Nonresponse Rates in CrossSectional Surveys.” ANNALS of the American Academy of Political and Social Science 645:36-59.

Cantor, David, Barbara C. O'Hare, and Kathleen S. O'Connor. 2008. “The Use of Monetary Incentives to Reduce Nonresponse in Random Digit Dial Telephone Surveys.” In Advances

in Telephone Survey Methodology, edited by James M. Lepkowski, Clyde Tucker, J. Michael Brick, Edith D. de Leeuw, Lilli Japec, Paul J. Lavrakas, Michael W. Link, and Roberta L. Sangster, ch. 22. Hoboken, NJ: Wiley.

Castiglioni, Laura, Klaus Pforr, and Ulrich Krieger. 2008. “The Effect of Incentives on Response Rates and Panel Attrition: Results from a Controlled Experiment.” Survey Research Methods 2:151-58.

Church, Allan H. 1993. “Estimating the Effect of Incentives on Mail Survey Response Rates. A Meta-Analysis.” Public Opinion Quarterly 57:62-79.

Couper, Mick P., and Edith D. de Leeuw. 2003. “Nonresponse in Cross-Cultural and CrossNational Surveys.” In Cross-Cultural Survey Methods, edited by Janet A. Harkness, Fons J. R. van de Vijver, and Peter Ph. Mohler, ch. 11. Hoboken, NJ: John Wiley.

de Luca, Giuseppe, and Oliver Lipps. 2005. “Fieldwork and Survey Management in SHARE.” In The Survey of Health, Aging, and Retirement in Europe-Methodology, edited by Axel Börsch-Supan, and Hendrik Jürges, ch. 7. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).

de Leeuw, Edith D., and Wim de Heer. 2002. “Trends in Household Survey Nonresponse: A Longitudinal and International Comparison.” In Survey Nonresponse, edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, ch. 3. New York: John Wiley.

Diekmann, Andreas. 2011. Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek: Rowohlt-Taschenbuch-Verlag.

Fienberg, Howard. 2013. Threat to Respondent Incentives in Federal Surveys May Be Averted This Year in House of Representatives, July 10. Available at http://www.marketingresearch.org/article/ threat-respondent-incentives-federal-surveys-may-be-averted-year-house-representatives.

Gelman, Andrew, Matt Stevens, and Valerie Chan. 2003. “Regression Modeling and MetaAnalysis for Decision Making: A Cost-Benefit Analysis of Incentives in Telephone Surveys.” Journal of Business & Economic Statistics 21:213-25.

Goldenberg, Karen L., David McGrath, and Lucilla Tan. 2007. “The Effects of Incentives on the Consumer Expenditure Interview Survey.” Proceedings of the Survey Research Methods Section of the American Statistical Association, pp. 5985-5999. Available at http://www.amstat.org/sections/srms/Proceedings/y2009/Files/400053.pdf.

Göritz, Anja S. 2006. “Incentives in Web Studies: Methodological Issues and a Review.” International Journal of Internet Science 1:58-70.

———. 2010. “Using Lotteries, Loyalty Points, and Other Incentives to Increase Participant Response and Completion.” In Advanced Methods for Conducting Online Behavioral Research, edited by Samuel D. Gosling and John A. Johnson, ch. 14. Washington, DC: American Psychological Association.

Göritz, Anja S., Hans-Georg Wolff, and Daniel G. Goldstein. 2008. “Individual Payments as a Longer-Term Incentive in Online Panels.” Behavior Research Methods 40:1144-1149.

Groves, Robert M., and Mick P. Couper. 1998. Nonresponse in Household Interview Surveys. New York: John Wiley.

Groves, Robert M., Floyd J. Fowler Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. Survey Methodology. 2nd ed. Hoboken, NJ: John Wiley.

Häder, Sabine, and Siegfried Gabler. 2003. “Sampling and Estimation.” In Cross-Cultural Survey Methods, edited by Janet A. Harkness, Fons J. R. van de Vijver, and Peter Ph. Mohler, ch. 8. Hoboken, NJ: John Wiley.

Haisken-De New, John P., and Joachim R. Frick. 2005. DTC. Desktop Companion to the German Socio-Economic Panel Study (GSOEP). DIW. Berlin.

Hanefeld, Ute. 1987. Das sozio-ökonomische Panel. Grundlagen und Konzeption. Frankfurt am Main and New York: Campus.

Hox, Joop, and Edith D. de Leeuw. 2002. “The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison.” In Survey Nonresponse,

edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, ch. 7. New York: Wiley.

Huinink, Johannes, Josef Brüderl, Bernhard Nauck, Sabine Walper, Laura Castiglioni, and Michael Feldhaus. 2011. “Panel Analysis of Intimate Relationships and Family Dynamics (pairfam): Conceptual Framework and Design.” Zeitschrift für Familienforschung 23:77-101.

Infas, ed. 2009. Nationales Bildungspanel (NEPS) Etappe 8: “Adult Education and Lifelong Learning” Erste Welle der Erwachsenenetappe. Methodenbericht Machbarkeitsstudie Dezember 2009. Bonn.

Jäckle, Annette, and Peter Lynn. 2008. “Respondent Incentives in a Multi-Mode Panel Survey: Cumulative Effects on Nonresponse and Bias.” Survey Methodology 34:105-17.

Johnson, Timothy P., Diane O'Rourke, Jane Burris, and Linda Owens. 2002. “Culture and Survey Nonresponse.” In Survey Nonresponse, edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, ch. 4. New York: Wiley.

Keil, Silke I., and Jan W. van Deth. 2012. Deutschlands Metamorphosen. Ergebnisse des European Social Survey 2002 bis 2008. Baden-Baden: Nomos.

Kleinert, Corinna, Britta Matthes, Manfred Antoni, Katrin Drasch, Michael Ruland, and Annette Trahms. 2011. “ALWA-New Life Course Data for Germany.” Schmollers Jahrbuch 131:625-34.

Klevmarken, Anders, Patrick Hesselius, and Bengt Swensson. 2005. “The SHARE Sampling Procedures and Calibrated Design Weights.” In The Survey of Health, Aging, and Retirement in Europe-Methodology, edited by Axel Börsch-Supan and Hendrik Jürges, ch. 5. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).

Koch, Achim, and Martina Wasmer. 2004. “Der ALLBUS als Instrument zur Untersuchung sozialen Wandels. Eine Zwischenbilanz nach 20 Jahren.” In Sozialer und politischer Wandel in Deutschland. Analysen mit ALLBUS-Daten aus zwei Jahrzehnten, edited by Rüdiger Schmitt-Beck, Martina Wasmer, and Achim Koch, ch. 2. Wiesbaden: VS Verlag für Sozialwissenschaften.

Laurie, Heather. 2007. The Effect of Increasing Financial Incentives in a Panel Survey: An Experiment on the British Household Panel Survey, Wave 14. ISER Working Paper Series. University of Essex, Colchester. Available at https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2007-05.pdf.

Laurie, Heather, and Peter Lynn. 2009. “The Use of Respondent Incentives on Longitudinal Surveys.” In Methodology of Longitudinal Surveys, edited by Peter Lynn, ch. 12. Chichester: John Wiley.

Lynn, Peter, Giuseppe de Luca, Matthias Ganninger, and Sabine Häder. 2013. “Sample Design in SHARE Wave Four.” In SHARE Wave 4 Innovations & Methodology, edited by Frederic Malter and Axel Börsch-Supan, ch. 8. Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Mack, Stephen, Vicki Huggins, Donald Keathley, and Mahdi Sundukchi. 1998. “Do Monetary Incentives Improve Response Rates in the Survey of Income and Program Participation?” Proceedings of the Survey Research Methods Section of the American Statistical Association, pp. 529-34. Available at https://www.amstat.org/sections/SRMS/Proceedings/ papers/1998_089.pdf.

Malter, Frederic. 2013. “Fieldwork Management and Monitoring in SHARE Wave Four.” In SHARE Wave 4 Innovations & Methodology, edited by Frederic Malter and Axel Börsch-Supan, ch. 9. Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Malter, Frederic, and Axel Börsch-Supan, eds. 2013. SHARE Wave 4 Innovations & Methodology. Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Marketing Research Association. 2015 “Respondent Incentives in Federally Conducted or Funded Survey Research.” Position Paper, January 12, 2015. Available at http://www.marketingresearch.org/legal-article/respondent-incentives-federally-conducted-or-funded-survey-research.

Martin, Elizabeth, Denise A. Abreu, and Franklin Winters. 2001. “Money and Motive: Effects of Incentives on Panel Attrition in the Survey of Income and Program Participation.” Journal of Official Statistics 17:267-84.

Mutti, Seema, Ryan David Kennedy, Mary E. Thompson, and Geoffrey T. Fong. 2014. “Prepaid Monetary Incentives – Predictors of Taking the Money and Completing the Survey: Results from the International Tobacco Control (ITC) Four-Country Survey.” Sociological Methods & Research 43:338-55.

Nicoletti, Cheti, and Nicholas Buck. 2004. “Explaining Interviewee Contact and Co-Operation in the British and German Household Panels.” ISER Working Paper Series. University of Essex, Colchester. Available at https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2004-06.pdf.

Philipson, Tomas. 1997. “Data Markets and the Production of Surveys.” Review of Economic Studies 64:47.

Pierson, Steve. 2013. “Of Note: Bill Banning Incentives on Federal Surveys Introduced in House.” ASA SMRS Newsletter 36:7.

Rammstedt, Beatrice, ed. 2013. Grundlegende Kompetenzen Erwachsener im internationalen Vergleich. Ergebnisse von PIAAC 2012. Münster: Waxmann.

Richter, David, and Jürgen Schupp. 2012. SOEP Innovation Panel (SOEP-IS): Description, Structure, and Documentation. SOEP Papers on Multidisciplinary Panel Data Research. Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

Rubin, Donald B. 1976. “Inference and Missing Data.” Biometrika 63:581-92.

Scherpenzeel, Annette, and Vera Toepoel. 2012. “Recruiting a Probability Sample for an Online Panel. Effects of Contact Mode, Incentives, and Information.” Public Opinion Quarterly 76:470-90.

Schnell, Rainer. 1997. Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske + Budrich.

———. 2012. Survey-Interviews. Methoden standardisierter Befragungen. Wiesbaden: VS Verlag für Sozialwissenschaften.

Schoeni, Robert F., Frank P. Stafford, Katherine McGonagle, and Patricia Andreski. 2013. “Response Rates in National Panel Surveys.” ANNALS of the American Academy of Political and Social Science 645:60-87.

Schröder, Mathis, Denise Saßenroth, John Körtner, Martin Kroh, and Jürgen Schupp. 2013. Experimental Evidence of the Effect of Monetary Incentives on Cross-Sectional and Longitudinal Response: Experiences from the Socio-Economic Panel (SOEP). SOEP Papers on Multidisciplinary Panel Data Research. 603. Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

Schupp, Jürgen. 2012. “Die verborgenen Kosten monetärer Anreize. Lohnt sich Motivierung durch Incentivierung?” DIW-Wochenbericht 79(6):20. Available at http://www.diw.de/documents/publikationen/73/diw_01.c.392570.de/12-6-4.pdf.

Siegel, Nico A., Simon Huber, and Anne Bohlender. 2012. “Summary Report SOEP Fieldwork in 2011.” In SOEP Wave Report 2011, edited by Sandra Gerstorf and Jürgen Schupp, ch. 11. Berlin: DIW Berlin.

Simmons, Eleanor, and Amanda Wilmot. 2004. “Incentive Payment on Social Surveys: A Literature Review.” Social Survey Methodology Bulletin 53:1-11.

Singer, Eleanor. 2002. “The Use of Incentives to Reduce Nonresponse in Household Surveys.” In Survey Nonresponse, edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, ch. 11. New York: John Wiley.

Singer, Eleanor, John van Hoewyk, Nancy Gebler, Raghunathan Trivellore, and Katherine McGonagle. 1999. “The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys.” Journal of Official Statistics 15:217-30.

Singer, Eleanor, and Cong Ye. 2013. “The Use and Effects of Incentives in Surveys.” ANNALS of the American Academy of Political and Social Science 645:112-41.

Smith, Tom W. 1978. “In Search of House Effects: A Comparison of Responses to Various Questions by Different Survey Organizations.” Public Opinion Quarterly 42:443.

Stadtmüller, Sven. 2009. “Rücklauf gut, alles gut? Zu erwünschten Effekten monetärer Anreize bei postalischen Befragungen.” Methoden, Daten, Analysen (MDA) 3:167-85.

Statistisches Bundesamt. 2014. “Demographic and Regional Standards.” Available at https://www.destatis.de/EN/Methods/DemographicStandards/DemograpicStandards.html.

Stoop, Ineke, Jaak Billiet, Achim Koch, and Rory Fitzgerald. 2010. Improving Survey Response: Lessons Learned from the European Social Survey. Oxford: Wiley-Blackwell.

Trappmann, Mark, Stefanie Gundert, Claudia Wenzig, and Daniel Gebhardt. 2010. “PASS-A Household Panel Survey for Research on Unemployment and Poverty.” Schmollers Jahrbuch 130:609-22.

Wagner, Gert G., Joachim R. Frick, and Jürgen Schupp. 2007. “The German Socio-Economic Panel Study (SOEP): Scope, Evolution, and Enhancements.” Schmollers Jahrbuch 127:139-69.

Wasmer, Martina, Evi Scholz, Michael Blohm, Jessica Walter, and Regina Jutz. 2012. Konzeption und Durchführung der “Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften” (ALLBUS) 2010. GESIS Technical Report. Gesis-Leibniz-Institut für Sozialwissenschaften, Mannheim.

Zabal, Anouk, Silke Martin, Natascha Massing, Daniela Ackermann, Susanne Helmschrott, Ingo Barkow, and Beatrice Rammstedt. 2014. PIAAC Germany 2012. Technical Report. Münster: Waxmann.

Zagorsky, Jay L., and Patricia Rhoton. 2008. “The Effects of Promised Monetary Incentives on Attrition in a Long-Term Panel Survey.” Public Opinion Quarterly 72:502-13.