Original Manuscript

Recruiting a Probability-Based Online Panel via Postal Mail: Experimental Evidence

Carina Cornesse[1], Barbara Felderer[1,2], Marina Fikel[1],
Ulrich Krieger[1], and Annelies G. Blom[1]

Abstract
Once recruited, probability-based online panels have proven to enable high-quality and high-frequency data collection. In ever faster-paced societies and, recently, in times of pandemic lockdowns, such online survey infrastructures are invaluable to social research. In absence of email sampling frames, one way of recruiting such a panel is via postal mail. However, few studies have examined how to best approach and then transition sample members from the initial postal mail contact to the online panel registration. To fill this gap, we implemented a large-scale experiment in the recruitment of the 2018 sample of the German Internet Panel (GIP) varying panel recruitment designs in four experimental conditions: online-only, concurrent mode, online-first, and paper-first. Our results show that the online-only design delivers higher online panel registration rates than the other recruitment designs. In addition, all experimental conditions led to similarly representative samples on key socio-demographic characteristics.

Keywords
online panels, offline recruitment, probability sampling, mail surveys, push-to-web, concurrent mode, mixed-mode, representativeness

Online panels constitute valuable infrastructures that are increasingly attractive for conducting social science research (Callegaro et al., 2014). The appeal of online panels lies in their speed, flexibility, and cost-effectiveness (Blom et al., 2016). For example, online panels can be adapted flexibly to unforeseen societal changes and shock events, such as the COVID-19 pandemic, and provide high-frequency and high-quality data to decision makers fast when needed (see Blom et al., 2020). In addition, an increasing body of literature shows that, when based on a probability sample of the general population, online panels can yield representative population predictions (for a review, see Cornesse et al., 2020).

---

[1] University of Mannheim, Germany
[2] GESIS—Leibniz Institute for the Social Sciences, Mannheim, Germany

Corresponding Author:
Carina Cornesse, University of Mannheim, Mannheim 68159, Germany.
Email: carina.cornesse@uni-mannheim.de

Despite these advantages, recruiting a high-quality probability-based online panel of the general population is challenging (Baker et al., 2010). Whereas for some special-population online surveys, it may be possible to randomly sample from existing email lists (see, e.g., Couper & Miller, 2008; Hupp, 2015), for the general population, a sampling frame of email addresses is usually unavailable. Thus, online panel recruitment and surveying are typically conducted in two separate, bi-modal stages: A probability sample of the general population is first approached in an offline mode. For this initial offline contact with the sample members, face-to-face interviews, postal mail contacts, or a mix of contact modes which may include telephone calls are typically used (see Callegaro et al., 2014). Contacted sample members then have to be transferred to the online mode for online panel data collection. This can, for example, be done by requesting sample members' email addresses for further contact or by asking them to follow a website link to a registration page.

In Europe, it is common to use interviewers for the initial contact. The Dutch Longitudinal Internet Studies for the Social Sciences (LISS) Panel (Scherpenzeel & Toepoel, 2012), the French Études Longitudinales par Internet Pour les Sciences Sociales (ELIPSS) Panel (Revilla et al., 2016), the German GESIS Panel (Bosnjak et al., 2018), and the first two recruitment rounds of the German Internet Panel (GIP; Blom et al., 2015)[1] all relied on face-to-face and/or telephone interviewers. In the United States, initial contact is more commonly established via postal mail, such as in the American Trends Panel (Pew Research Center, 2019) and Ipsos (2020) Knowledge Panel or via a mix of self-administered and interviewer-administered recruitment modes, such as in the Gallup Panel (GALLUP, 2020) and AmeriSpeak (National Opinion Research Center, 2019).

Regardless of how initial contact is established, many probability-based online panels implement two surveys during their recruitment process: first, an initial recruitment survey during the offline contact (e.g., to ask for an email address); second, a subsequent online registration survey to collect consent to becoming a member of the online panel and additional background information (for examples of such recruitment processes with two surveys, see Blom et al., 2016; for a counter-example that skips the offline recruitment survey, see Høgestøl & Skjervheim, 2014).

Compared to approaches that use interviewers in the online panel recruitment (either for face-to-face or telephone contacts), postal mail contacts tend to be less expensive and avoid interviewer effects (for an overview of potential interviewer effects, see West & Blom, 2017).[2] However, little is known about how to best conduct such a mail-mode recruitment in a way that maximizes panel registration rates and representativeness (for a notable exception, see Rao et al., 2010). The overarching research question in this article, therefore, is:

> In terms of panel registrations and sample representativeness, which postal mail contact design strategy
> is most successful at recruiting sample members into a probability-based online panel?

Literature and Hypotheses

Previous research has shown that probability-based online panels can successfully be recruited via postal mail (see, e.g., Høgestøl & Skjervheim, 2014). However, postal mail recruitments can be conducted in a number of ways that may impact the success of the recruitment in terms of how many sample members can be recruited and to what extent the recruited sample accurately reflects the intended target population.

The most straightforward approach to recruiting a probability-based online panel via mail is to just send sample members a letter that includes a request to register to the panel on the Internet. In the following, this approach will be referred to as an "online-only" design strategy. Online-only design strategies are comparatively inexpensive and particularly attractive for sample members whose preferred survey participation mode is the web anyway (as compared to face-to-face,

telephone, or mail-back surveys; see, e.g., Olson et al., 2012; Shih & Fan, 2007; Smyth et al., 2014). However, online-only approaches are less attractive for people who are reluctant to provide data online (see, e.g., Cornesse & Schauerer, 2021) and they completely exclude people who are unable to participate in online surveys, because they lack internet access or the necessary skills to use the internet (see, e.g., Blom et al., 2017; Herzing & Blom, 2019).

To establish rapport with sample members who prefer survey modes other than the web, offering the possibility of participating in the first part of a two-stage recruitment process (i.e., the initial recruitment survey) via mail-back paper questionnaires may boost response. In addition, the data provided on paper may later be used in calibration weighting or, budget permitting, to provide non-internet users with the necessary equipment to participate in the panel online (see, e.g., Blom et al., 2017, Leenheer & Scherpenzeel, 2013). Moreover, once sample members have experienced being a respondent offline, they may be willing to switch to the online mode to stay in the panel.

Offering sample members the possibility of participating in the initial recruitment survey either online or via mail-back paper questionnaires is a mixed-mode design strategy (see de Leeuw, 2013; de Leeuw & Berzelak, 2016, for definitions). Some evidence from cross-sectional surveys suggests that mixed-mode designs can lead to higher response rates than single-mode designs, in particular, when the single-mode design is online-only (see Couper, 2011; Dillman et al., 2014; de Leeuw, 2018, for an overview). We thus obtain our first hypothesis:

Hypothesis 1: Online panel registration rates will be higher in mixed-mode recruitment designs than in online-only recruitment designs.

Whereas mixed-mode designs may increase response rates as compared to online-only designs, some research suggests that modes should be offered sequentially rather than concurrently (see, e.g., Holmberg et al., 2010; Mauz et al., 2018). In a concurrent mode design, respondents are offered a choice between modes (see de Leeuw & Toepoel, 2018, for a definition). However, a meta-analysis by Medway and Fulton (2012) concludes that, on average, sample members are more likely to respond in a single-mode design than a concurrent mode design. In contrast, sequential mixed-mode designs first offer one mode, and, subsequently, sample members who did not take up this initial mode are offered a different mode (see, e.g., Klausch et al., 2017; Lynn, 2013; for a definition, see also Bethlehem & Biffignandi, 2012, Chapter 7). Several studies on cross-sectional surveys demonstrate the success of sequential mixed-mode designs (see, e.g., Dillman, 2017; Patrick et al., 2018). We thus obtain our second hypothesis:

Hypothesis 2: Online panel registration rates will be higher in sequential mode designs than in concurrent mode designs.

When using a sequential mode design, researchers have to decide which mode to offer first and which mode to offer second (i.e., only to initial nonrespondents). A series of studies conducted at Washington State University (see Dillman, 2017, for an overview) find that inviting sample members to an online survey first and, subsequently, following up nonrespondents with a mail-back paper questionnaire leads to the best cost-benefit ratio and the highest share of online respondents (see Edwards et al., 2014; Dillman et al., 2014; Messer, 2012; Millar & Dillman, 2011; Smyth et al., 2010). This particular sequential strategy is commonly referred to as a "push-to-web" design (see Dillman et al., 2014, Chapter 11). In our study, we will refer to it as an "online-first" design. Our third hypothesis, therefore, is:

Hypothesis 3: Online-first designs will lead to higher online panel registration rates than other mixed-mode designs.

While online-first designs have been found to have many advantages, they have usually not been found to maximize overall response rates. On the contrary, some studies suggest that the highest cross-sectional survey response rates are achieved with sequential designs that prioritize the mail-back mode, that is, a "paper-first" design (see, e.g., Smyth et al., 2010). This may have changed in recent years with the world-wide increase in internet use, but to our knowledge, no recent studies comparing paper-first designs to other mixed-mode strategies are available.

In a paper-first design, sample members are initially invited to participate in a mail-back survey. Subsequently, nonrespondents are followed up with reminder letters that contain a link to an online version of the survey. While paper-first designs may thus boost response in cross-sectional surveys, the same may not be true for a panel context. On the contrary, some studies suggest that switching panel members from an offline survey mode that they were familiar with to the web may lead to increased nonresponse, either directly or in subsequent panel waves (see, e.g., Bianchi et al., 2017; Couper et al., 2007; Jäckle et al., 2015). Since in the GIP, sample members will have to make the switch to the online mode by the time they register to the panel, we obtain our fourth hypothesis:

Hypothesis 4: Paper-first designs will lead to lower online panel registration rates than other mixed-mode designs.

Whereas a number of studies have examined the impact of different mixed-mode and single-mode designs on survey response rates, little is known directly about how these designs affect sample representativeness (for an exception in a special population survey, see McMaster et al., 2017). Indirect evidence from studies into the representativeness of online surveys, however, suggests that people who do not use the internet or who are unwilling to participate in surveys on the internet differ from people who use the internet (see, e.g., Blom et al., 2017; Callegaro et al., 2014; Helsper & Reisdorf, 2017). From this literature, we derive the expectation that different modes may lead to different selection effects, which could potentially cancel each other out. Therefore, our last hypothesis is:

Hypothesis 5: Mixed-mode recruitment designs will lead to higher online panel sample representativeness than online-only designs.

Data and Experimental Design

To test the hypotheses described above, a large-scale experiment was conducted in the 2018 recruitment of the GIP. In an experimental setup that compares the online-only, concurrent mode, online-first, and paper-first designs, we examine response rates and sample representativeness in the GIP recruitment process.

Data. The GIP is a multi-topic, probability-based online panel of the general population aged 16–75 years in Germany.[3] To date, the GIP has seen three recruitment rounds, in 2012, 2014, and 2018. The 2012 and 2014 samples were recruited by means of interviewer-mediated face-to-face interviews (for details, see Blom et al., 2016; Blom et al., 2015; Blom & Herzing, 2017; Blom et al., 2017).

In an effort to reduce costs, eliminate interviewer effects, and explore alternative ways of recruiting probability-based online panels, the 2018 GIP sample was recruited via postal mail with a two-stage probability sample drawn from local community population registers (see Figure 1 for a graphical summary of the sampling and data collection protocol and Appendix B for a detailed description).
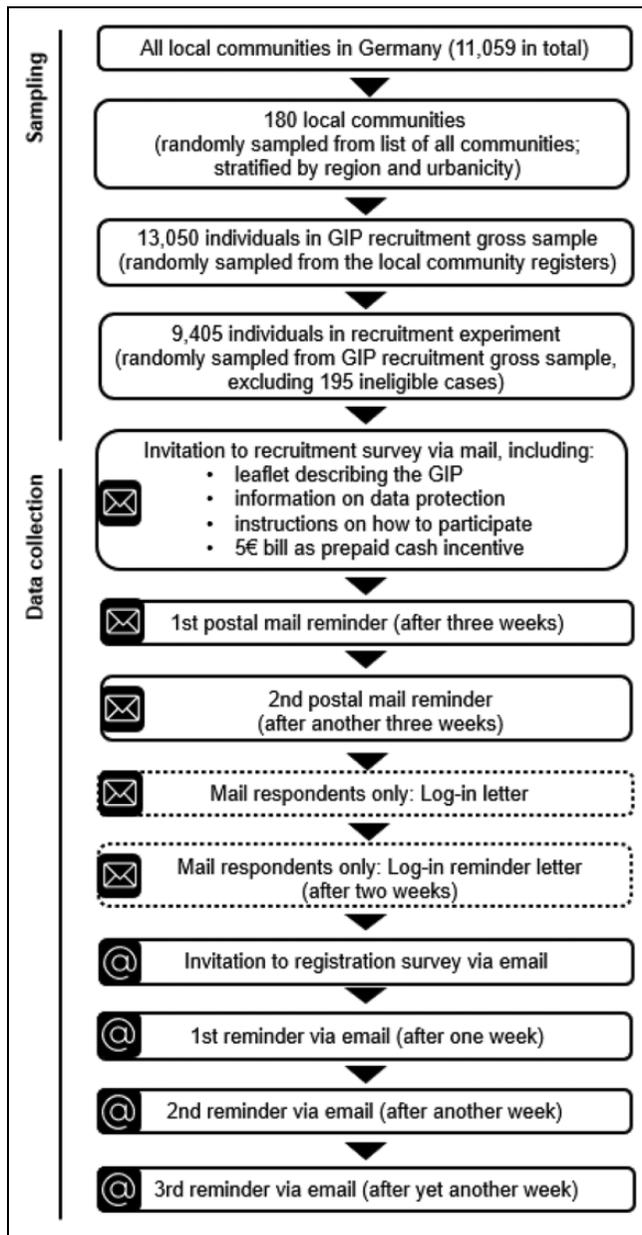
Figure 1. Graphical summary of the sampling and recruitment protocol.

Fieldwork for the panel recruitment started on September 13, 2018, and closed on February 28, 2019, with the large majority of cases recruited during the first 2 months. The recruitment survey contained 13 questions on attitudes toward digitalization and societal change, internet use, and socio-demographics (see the Online Appendix for the full questionnaires). The subsequent registration survey contained approximately 20 questions (depending on how many follow-up questions were triggered by filters). The central purpose of the registration survey was to obtain panel participation consent. In addition, the registration survey contained some questions to complete the
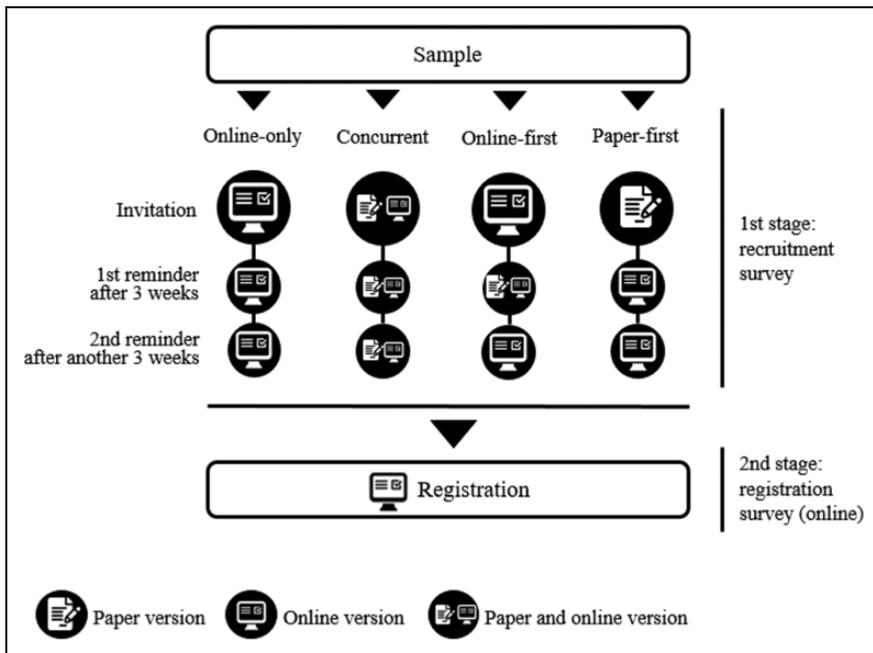
Figure 2. Overview of the recruitment process and experimental setup.

panel members' profiles (e.g., address confirmation, incentive payout preferences, further socio-demographic information; see the Online Appendix). Overall, 37.8% of all gross sample members assigned to the recruitment experiment participated in the recruitment survey and 22.7% of all gross sample members assigned to the recruitment experiment participated in the subsequent registration survey (see Figure A1 in Appendix A for more information).

Experimental groups. Four mail recruitment strategies were tested against each other: online-only, concurrent mode, online-first, and paper-first. The sample of individuals assigned to the recruitment experiment was randomly split into four experimental groups of equal size ($n = 2,400$ in each group; 195 cases were later identified as ineligible, see Appendix B and Table A1 in Appendix A for more information). Depending on their experimental group, sampled individuals were either invited and/or reminded to participate in a paper and/or online version of the recruitment survey (see Figure 2). The content of the invitation letters, reminder letters, and other study material was identical across experimental groups, except for necessary differences needed for the treatment.

Online-only. Sample members in the online-only experimental group received an invitation letter via mail that contained a link and log-in credentials for the online version of the recruitment survey. Sample members who had not logged in to the online version of the recruitment survey after 3 weeks received a first reminder letter with a link and log-in credentials for the online version of the recruitment survey and a second reminder letter with the same information after another 3 weeks.

Concurrent mode. In the concurrent mode, sample members received an invitation letter via mail containing the paper version of the recruitment survey and a prepaid mail-back envelope as well as the link and log-in credentials for the online version of the recruitment survey. In the invitation letter, sample members were informed that they could pick the data collection mode they preferred.

Sample members who had neither sent back the paper questionnaire nor logged in to the online version of the recruitment survey received two reminder letters. Each of the reminder letters included the paper questionnaire and the prepaid mail-back envelope as well as the link and log-in credentials for the online version. In each of the reminder letters, sample members were, again, instructed to pick their preferred mode.

Online-first. In the online-first experimental group, sample members first received an invitation letter via postal mail that contained a link and log-in credentials for the online version of the recruitment survey. The invitation letter did not mention the paper mode offered in follow-up mailings. Sample members who had not logged in to the online version of the recruitment survey after 3 weeks received a reminder letter that, again, included the link and log-in credentials for the online version of the recruitment survey, along with the paper version of the recruitment survey with a prepaid mail-back envelope. Sample members who had neither participated in the online version of the recruitment survey nor sent back the paper questionnaire after another 3 weeks received a second reminder letter that, once again, contained the link and log-in credentials but no paper questionnaire or envelope. However, sample members were informed that they could still send back the paper questionnaire that they had received with the first reminder letter. The decision to only include a paper questionnaire in the first reminder letter and not in the second reminder letter was made to strengthen the emphasis on the web as the dominant mode.

Paper-first. In the paper-first experimental group, sample members received an invitation letter via mail containing the paper version of the recruitment survey and a prepaid mail-back envelope. Sample members who had not sent back the paper questionnaire within 3 weeks received two reminder letters that included the link and log-in credentials for the online version of the recruitment survey. In the reminder letters, sample members were informed that they could—if they preferred—still fill out and send back the paper questionnaire that they had received with the invitation letter. However, they did not receive another paper questionnaire and prepaid mail-back envelope. By not providing the paper questionnaire again with the reminder letters, we aimed to avoid the decision-making burden for the respondents which is suspected to be responsible for the comparatively low response rates of concurrent mode surveys (see Medway & Fulton, 2012).

Subsequently, all recruitment survey respondents—independent of their experimental group—were invited to the online panel registration survey via email. Respondents to the online version of the recruitment survey provided their email addresses at the end of the recruitment survey questionnaire. This email address was then used to send them the invitation to the registration survey. Respondents to the paper version of the recruitment survey received a letter with log-in credentials to the online panel and provided their email address when first logging into the online panel home space. Subsequently, they received the invitation email to the registration survey. This in-between step for the mail respondents was deemed necessary to avoid sending invitation emails to email addresses that the mail respondents may have misspelled or that may have been miscoded when entering their data from the recruitment survey questionnaires. Ensuring that invitation emails were not mistakenly sent to other people (e.g., with a similar email address) was considered particularly crucial because the sample member following the registration survey link would already find some of their personal data (in particular their name and postal address) on their online panel profile.

At the online panel registration survey stage, no paper version of the survey was offered anymore. Nonrespondents to the online panel registration survey received up to three reminders via email. Upon completing the online panel registration survey, respondents become regular members of the GIP and are invited to its bi-monthly online survey waves.

Analytical Strategy

Two metrics are used to assess outcomes of the experiment: response rates (to test Hypotheses 1–4) and sample representativeness (to test Hypothesis 5).

Response rates. For each experimental group, online panel registration rates (RR_{reg}) were computed as an indicator of panel recruitment success. To examine progress in panel registrations across the fieldwork period, the registration rate was calculated for each experimental group cumulatively at each fieldwork day in order to track the overall recruitment success with

$$RR_{\text{reg}} = \frac{R_{\text{reg}}}{N_{\text{rec}_{\text{eligible}}}}, \tag{1}$$

where R_{reg} is the number of completed online panel registration survey questionnaires[4] and $N_{\text{rec}_{\text{eligible}}}$ is the number of eligible sample members in the experimental gross sample. In sum, the registration survey response rate can be considered as the overall final response rate of the panel recruitment process.

Because they might provide valuable background information on whether the different experimental groups perform consistently well in relation to one another or whether some groups perform well at an earlier stage of the recruitment process while others perform better at the final stage, we also compute response rates for the recruitment survey, which takes place prior to the registration survey. To examine progress in participation across the fieldwork period, response rates for each experimental group were calculated cumulatively at each fieldwork day with

$$RR_{\text{rec}} = \frac{R_{\text{rec}_{\text{paper}}} + R_{\text{rec}_{\text{online}}}}{N_{\text{rec}_{\text{eligible}}}}, \tag{2}$$

where $R_{\text{rec}_{\text{paper}}}$ is the number of completed paper questionnaires,[5] $R_{\text{rec}_{\text{online}}}$ is the number of completed online questionnaires,[6] and $N_{\text{rec}_{\text{eligible}}}$ is the number of eligible sample members independent of whether they participated in the recruitment survey or not.[7]

Sample representativeness. For each experimental group, relative biases ($\widehat{RB}$; see Groves, 2006) were estimated for each category of a set of socio-demographic variables collected during the GIP 2018 recruitment process. The relative bias is defined as the difference between each estimate in a survey compared to the respective population total from a benchmark relative to that population total[8] and is given by:

$$\widehat{RB}(\hat{\bar{y}}) = \frac{\hat{\bar{y}}_{\text{svy}} - \bar{y}_{\text{mz}}}{\bar{y}_{\text{mz}}} \times 100, \tag{3}$$

where $\hat{\bar{y}}_{\text{svy}}$ is the estimated proportion of a given category in the survey questionnaire and $\bar{y}_{mz}$ is the respective proportion of the given category in a population benchmark.

The standard error (*SE*) of the relative bias is given by

$$SE\left(\widehat{RB}(\hat{\bar{y}})\right) = \frac{100}{\bar{y}_{\text{mz}}} \times SE\left(\hat{\bar{y}}_{\text{svy}}\right), \tag{4}$$

where the *SE* of the survey estimate $\left(SE\left(\hat{\bar{y}}_{\text{svy}}\right)\right)$ is estimated using a sample variance estimation approach which takes into account the complexity of the sampling design (see Appendix B for more information). Generally, the estimation of the *SE* of the relative bias is straightforward under the assumption that the benchmark values are true population characteristics (i.e., have no sampling variance; see Felderer et al., 2019, for a mathematical derivation).

From the *SE*s of the relative bias estimates, the 95%-confidence intervals were computed. Confidence intervals that do not include zero indicate significant relative biases at the 5%-level.

In addition to exploring the relative biases of all the categories of a set of variables separately, the distribution of the absolute relative biases is investigated by experimental group, focusing on the mean and quartiles of the distribution. The mean across all examined absolute relative biases is commonly referred to as the absolute average relative bias (AARB; Groves, 2006). Similarly, we will refer to the median of the bias distribution as the median absolute relative bias (MARB, see Cornesse & Bosnjak, 2018). Since item nonresponse rates are negligible in all of the included variables (see Table A1 in Appendix A), complete-case analyses were conducted.

The socio-demographic variables included in the bias analyses are gender, age, and education, which were gathered in the recruitment survey, and marital status, household size, and citizenship, which were gathered in the registration survey. The survey data are compared to official German population statistics. These official population statistics are based on the German Mikrozensus (https://www.forschungsdatenzentrum.de/en/household/microcensus) conducted by the German federal statistical agency (Destatis). The Mikrozensus is a mandatory annual survey of 1% of the German resident population.[9]

Results

In this section, we first describe differences in response rates (Hypotheses 1–4) and subsequently provide a detailed analysis of sample representativeness (Hypothesis 5) across the four the experimental conditions. In the subsequent Summary section, we will discuss these results in light of whether they support our research hypotheses or not.
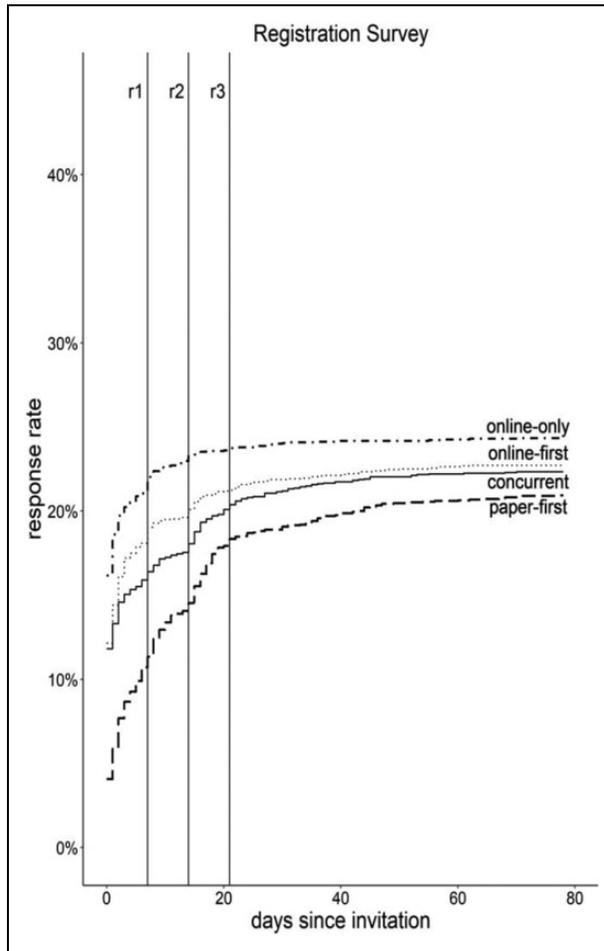
Response Rates

The highest online panel registration rate was achieved by the online-only group (24.5%), followed by online-first (22.8%), concurrent mode (22.4%), and paper-first (21.1%; see Figure 3; for more detailed information, including confidence intervals for all response rate estimates, see Table A2 in Appendix A).

This is despite the fact that the response rate to the recruitment survey, which took place prior to the registration survey, was highest in the concurrent-mode group (43.9%), followed by paper-first (41.1%), online-first (36.6%), and online-only (29.7%; see Figure A2 in Appendix A). This means that, overall, the online-only design, which eliminates paper questionnaires altogether, obtained the lowest recruitment survey response rate but the highest panel registration rate.

Before the first reminder, panel registration rates were higher in the two groups that promoted the online survey versions at the previous recruitment survey stage (online-first: 19.0%, online-only: 22.0%) than in the two groups that put more emphasis on the paper questionnaires at the recruitment survey stage (paper-first: 11.3%; concurrent mode: 16.5%). The concurrent mode group nearly catches up with the online-first group after the reminders though. The paper-first group achieves a steep increase in participation but never reaches the other groups.

The finding that panel registration rates were highest in the online-only group may in part be attributable to the fact that respondents to the paper version of the recruitment survey had an additional step in the recruitment process that required their action (i.e., they received an additional letter with online log-in information after completing the recruitment survey, see Figure 1). The additional letter may have had a negative impact among the respondents to the paper version of the recruitment survey. Indeed, the lowest online panel registration rate was achieved in the paper-first

Figure 3. Registration survey response rate by days since invitation and experimental group. *Note.* Cumulative response rates are based on the gross sample of eligible names and addresses drawn from the population registers. Vertical lines denote sending of reminders (r1 = first reminder, r2 = second reminder, and r3 = third reminder). Graph is truncated at 80 days of data collection.

group, which had the highest overall share of paper respondents in the recruitment survey (share of paper respondents: 83.8%). However, the online panel registration rates in the concurrent mode group and online-first group are nearly identical despite fairly different shares of paper version respondents in the recruitment survey (share of paper respondents in online-first group: 36.0% and share of paper respondents in concurrent mode group: 59.9%).

Sample Representativeness

To gain insights into how well the experimental groups performed in terms of recruiting a sample that adequately represents the general population, relative biases were assessed using a set of socio-demographic characteristics (see Figures 4 and 5 as well as Table A3 in Appendix A). In this section, we focus on the bias results for the online panel registration stage. For some variables (age, gender,
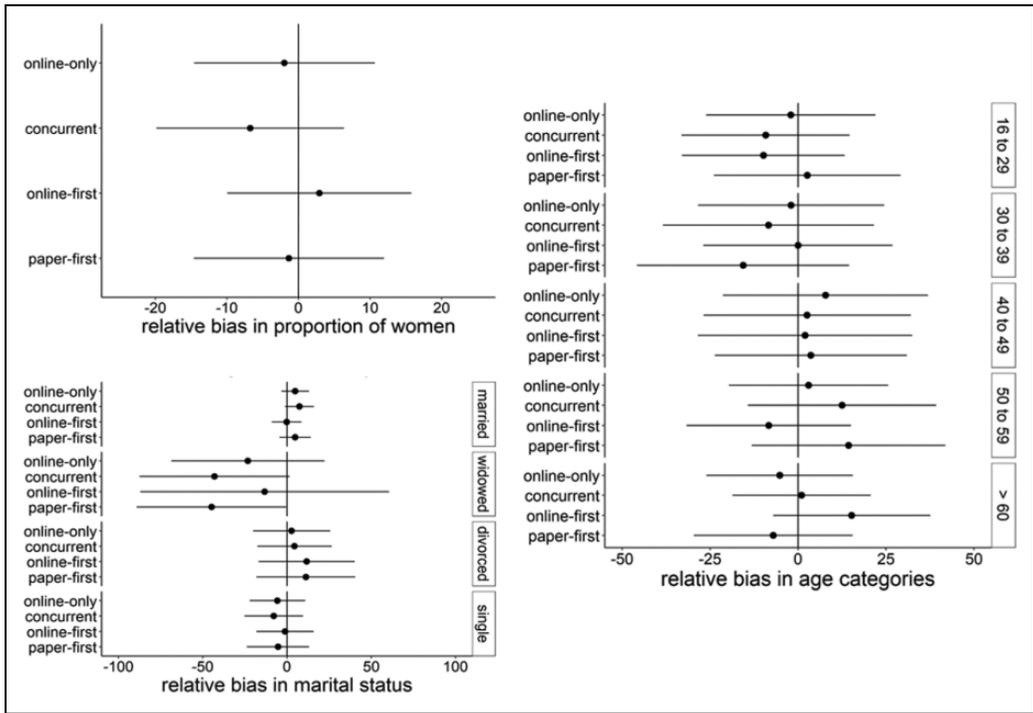
Figure 4. Relative biases in gender, age, and marital status at the registration survey stage across experimental groups. *Note.* The relative bias estimates are denoted by the dots and 95%-confidence intervals by the horizontal lines.

and education), we can also examine sample representativeness at the recruitment survey stage (see Figure A3 in Appendix A).

Overall, we find similar biases across all experimental groups. No significant bias can be detected with regard to gender, age, or marital status (see Figure 4). However, we find biases in terms of education, household size, and citizenship (see Figure 5).

With regard to education, all experimental groups underrepresent persons with low education (relative biases: -60.9% online-only, -61.9% concurrent mode, -61.8% online-first, and 67.0% paper-first) and overrepresent persons with high education (relative biases: 57.1% online-only, 50.3% concurrent mode, 57.1% online-first, and 63.1% paper-first). With regard to household size, all experimental groups underrepresent single households (relative biases: -24.3% online-only, -36.8% concurrent mode, -34.5% online-first, and -22.4% paper-first) and the online-only group overrepresents households with four or more members (relative bias: 24.9%). With regard to citizenship, all experimental groups overrepresent German citizens (relative biases: 7.6% online-only, 13.0% concurrent mode, 9.0% online-first, and 9.5% paper-first).

To aggregate the findings described above, the distribution of the absolute relative biases across experimental groups was examined at the panel registration survey stage (Figure 6; see also Table A4 and Figure A4 in Appendix A for results on the recruitment survey stage). Overall, sample representativeness was very similar across experimental groups in terms of the mean, median, and interquartile range of the bias distribution.

Regarding the mean of the absolute relative bias distribution, all experimental groups performed similarly at the panel registration stage (AARBs: 13.0% for online-only, 16.7% for concurrent mode,
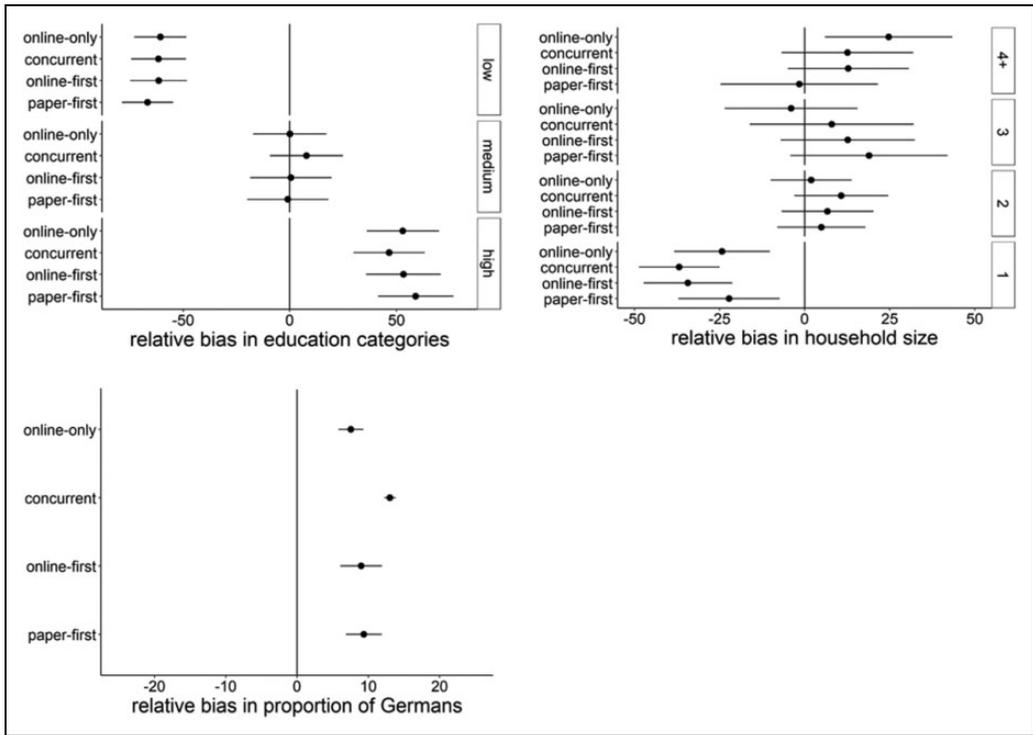
Figure 5. Relative biases in education, household size, and citizenship at the registration survey across experimental groups. *Note.* The relative bias estimates are denoted by the dots, and 95%-confidence intervals are denoted by the horizontal lines.
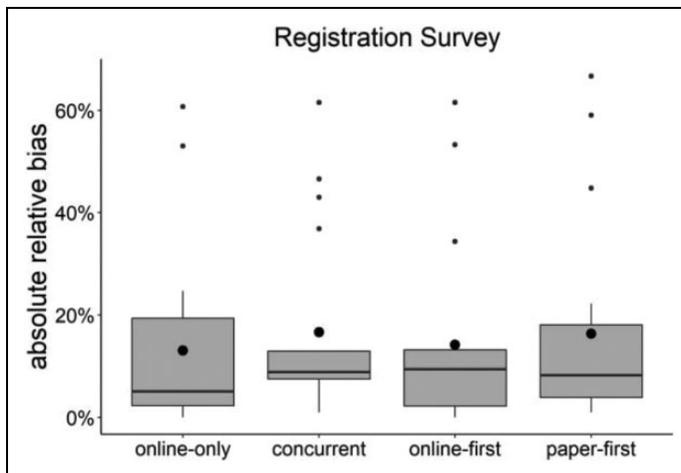
Figure 6. Distribution of absolute relative biases of all categories of the available socio-demographic variables in percent at the registration survey stage across experimental groups. *Note.* The large black dots denote the absolute average relative bias in each of the experimental groups. The horizontal bar in bold print denotes the median absolute relative bias, while the boxes and whiskers represent the interquartile range and the small dots represent outliers of a standard box-and-whiskers plot. The outliers are mostly due to the comparatively large biases on education.

14.2% for online-first, and 16.4% for paper-first). However, the medians were lower than the means (panel registration stage MARBs: 5.1% for online-only, 8.8% for concurrent mode, 9.4% for online-first, and 8.2% for paper-first). Thus, some large biases in a few categories skew the distribution of the absolute relative biases.

Regarding the interquartile range, we find high overlap across experimental groups. The same is true when examining the variation in relative biases. The outliers displayed in Figure 6 are primarily the education biases observed in Figure 5. Most notably, the maximum absolute relative bias in each experimental group (as displayed by the small dots that lie furthest away from the respective boxes) in each case represents the comparatively large underrepresentation of people with low education (see also Table A4 in Appendix A).

Summary

In the following, we discuss the results described above in light of our research hypotheses.

Hypothesis 1: Online panel registration rates will be higher in mixed-mode recruitment designs than in online-only recruitment designs.

The evidence from this study does not support Hypothesis 1. The online-only design led to very similar online panel registration rates compared to the mixed-mode designs (24.5% in the online-only group as compared to between 21.1% and 22.8% in the mixed-mode groups). This suggests that offering a paper version of the recruitment survey may not be necessary and that the cost-efficient online-only design may be sufficient to achieve reasonable online panel registration rates.

Hypothesis 2: Online panel registration rates will be higher in sequential mode designs than in concurrent mode designs.

Based on the evidence from this study, Hypothesis 2 has to be rejected. Online panel registration rates were essentially the same across all mixed-mode designs (between 21.1% and 22.8% in the sequential mode designs compared to 22.4% in the concurrent mode design). Overall, concurrent mode designs performed better than expected (for further evidence of this, see Wolf et al., in press).

Hypothesis 3: Online-first designs will lead to higher online panel registration rates than other mixed-mode designs.

The evidence from this study provides little support for Hypothesis 3. The online-first design led to very similar online panel registration rates compared to the other mixed-mode recruitment designs (22.8% as compared to 21.1% and 22.4% in the other mixed-mode groups). This suggests that online-first designs are a viable option for recruiting a probability-based online panel. However, they do not seem to offer a noteworthy advantage over other designs in terms of registration rates.

Hypothesis 4: Paper-first designs will lead to lower online panel registration rates than other mixed-mode designs.

The evidence from this study provides limited support for Hypothesis 4. The paper-first group had slightly lower registration rates than the other experimental groups (21.1% as compared to between 22.4% and 24.5% in the other experimental groups). These findings suggest that paper-first designs may be a viable option for cross-sectional surveys (see, e.g., Smyth et al.,

2010, and our findings regarding the response rates to the initial recruitment survey), but they cannot be recommended for recruiting a probability-based online panel.

Hypothesis 5: Mixed-mode recruitment designs will lead to higher online panel sample representativeness than online-only designs.

Based on the evidence from this study, Hypothesis 5 has to be rejected. All experimental groups performed similarly with regard to the examined socio-demographic characteristics. This suggests that offering mail-back questionnaires during recruitment does not increase the representativeness of probability-based online panels.

Discussion

The findings from our study provide insights into how to best recruit a probability-based online panel via mail. While mixed-mode approaches, such as online-first (which is often referred to as "push-to-web"), might be beneficial for conducting cross-sectional surveys, especially in terms of response rates, this does not seem to apply to the recruitment of online panels. Both panel registration rates and sample representativeness were just as high in the online-only condition as they were in the mixed-mode conditions. However, our results may be confounded by the potential impact of sending people who responded to the recruitment survey via paper questionnaires an additional log-in letter to ensure we did not mistakenly send registration survey invitation emails to the wrong people. While we deemed this additional step for paper respondents to be necessary for data protection and panel management reasons, future research should assess whether and to what extent this may have a negative impact on panel registration success among recruitment survey mail respondents.

It also remains unclear to what extent our findings are generalizable to other countries than Germany. The fact that online-only recruitment design strategies are successfully applied in Norway and the United States indicates that our findings may be generalizable to some other countries. However, more research is needed, in particular, in countries with different demographic profiles and degrees of digitalization. In addition, cultural differences may explain some of the differences between some of our findings and the cross-sectional survey literature. For example, German study participants may appreciate the possibility to pick their preferred survey mode in a concurrent design (leading to our comparatively high recruitment survey response rates in this group), while people in other countries may find it unnecessarily burdensome to have to choose between modes (leading, e.g., to the comparatively low concurrent mode response rates in the meta-analysis by Medway and Fulton in 2012).

Furthermore, the findings from this study suggest that offering paper questionnaires in the recruitment survey does not have a positive impact on online panel sample representativeness as compared to only offering a link and log-in credentials for online participation. However, it should be noted that this has only been examined with regard to a set of primary socio-demographic characteristics for which widely recognized official statistics were available that could be used as a gold-standard benchmark. It is possible that the mixed-mode recruitment designs increase online panel representativeness with regard to other characteristics, such as political attitudes. Furthermore, the questions asked on topics relating to digitalization and social change that were implemented in the recruitment survey may have kept some non-internet users and late adopters of digital technologies from providing any information, even via postal mail (although Table A5 on internet use by experimental group suggests otherwise). Future research needs to address these limitations and further explore best-practices for recruiting a probability-based online panel via postal mail.

Appendix A
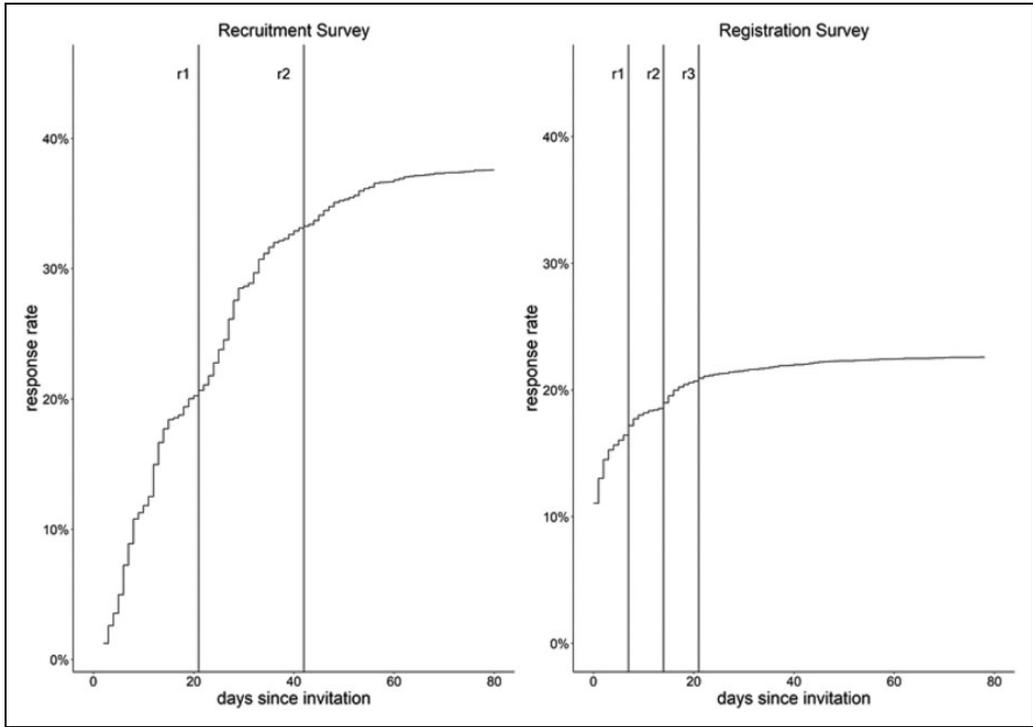
Figure A1. Recruitment and registration survey response rates by days since invitation. *Note.* Cumulative response rates are based on the gross sample of eligible names and addresses drawn from the population registers. Vertical lines denote sending of reminders (r1 = first reminder, r2 = second reminder, and r3 = third reminder). Graphs are truncated at 80 days of data collection.

Table A1. Case Numbers, Completions, and Item Nonresponse in the Recruitment and Registration Stage (Reported in Absolute Numbers).

Characteristic	Online-Only	Concurrent	Online-First	Paper-First	Total
Sample size	2,338	2,359	2,354	2,354	9,405
Recruitment stage responses	695	1,036	862	967	3,560
Registration stage responses	572	528	537	496	2,133
Recruitment					
Gender	0	6	1	6	13
Age	0	6	0	7	13
Education	0	22	12	25	59
Registration					
Gender	0	0	0	1	1
Age	0	0	0	1	1
Education	0	3	4	7	14
Marital status	1	0	1	0	2
Household size	3	2	4	0	9
Nationality	0	1	0	0	1

Table A2. Recruitment Survey Response Rates and Panel Registration Rates Across Experimental Groups in Percentage.[a]

| Rates	Online-Only (%)	Concurrent (%)	Online-First (%)	Paper-First (%)	Mean (%)
Recruitment					
Response rates					
Before first reminder	15.5 [14.9, 16.1]	25.7 [24.7, 26.7]	15.2 [14.6, 15.8]	26.2 [25.2, 27.2]	20.7 [20.3, 21.1]
Before second reminder	25.6 [24.6, 26.6]	38.9 [37.4, 40.5]	31.4 [30.2, 32.7]	37.4 [35.9, 38.9]	33.3 [32.6, 34.0]
Final	29.7 [28.5, 30.9]	43.9 [42.2, 45.7]	36.6 [35.1, 38.1]	41.1 [39.5, 42.7]	37.8 [37.0, 38.6]
Registration					
Registration rates					
Before first reminder	22.0 [21.1, 22.9]	16.5 [15.9, 17.1]	19.0 [18.3, 19.7]	11.3 [10.9, 11.7]	17.2 [16.9, 17.5]
Before second reminder	23.3 [22.4, 24.2]	18.1 [17.4, 18.8]	20.1 [19.3, 20.9]	14.5 [13.9, 15.1]	19.0 [18.6, 19.4]
Before third reminder	23.7 [22.8, 24.6]	20.4 [19.6, 21.2]	21.2 [20.4, 22.0]	18.3 [17.6, 19.0]	20.9 [20.5, 21.3]
Final	24.5 [23.5, 25.5]	22.4 [21.5, 23.3]	22.8 [21.9, 23.7]	21.1 [20.3, 21.9]	22.7 [22.3, 23.1]

[a]The 95% confidence intervals are given in brackets.
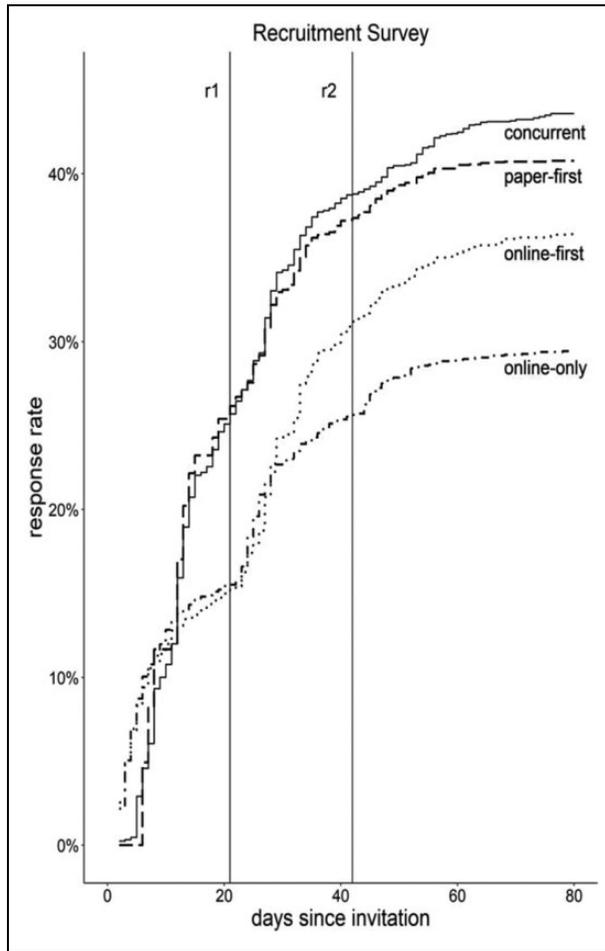
Figure A2. Recruitment survey response rates by days since invitation and experimental group. *Note.* Cumulative response rates are based on the gross sample of eligible names and addresses drawn from the population registers. Vertical lines denote sending of reminders (r1 = first reminder, r2 = second reminder, and r3 = third reminder). Graphs are truncated at 80 days of data collection.

Table A3. Proportions of Sociodemographic Characteristics (in %) Across Experimental Groups Among the Recruitment Survey Participants and the Registered Panel Sample Compared to the Official Statistics From the German Mikrozensus.

| Characteristic	Online-Only (%)	Concurrent (%)	Online-First (%)	Paper-First (%)	MZ (%)
Recruitment survey sample					
Gender					
Male	52.4	51.0	48.1	47.9	50.2
Female	47.6	49.0	51.9	52.1	49.8
Age					
16–29	20.9	17.0	17.5	17.5	21.1
30–39	16.8	13.1	15.3	14.8	16.8
40–49	18.4	18.1	17.4	17.8	17.4
50–59	21.9	24.8	21.5	23.5	21.5
60+	22.0	27.1	28.3	26.3	23.3
Education					
Low	14.4	17.8	16.6	15.9	30.9
Medium	33.1	38.9	35.4	34.0	33.9
High	52.5	43.4	48.0	50.1	35.0
Registered panel sample					
Gender					
Male	51.2	53.6	48.8	50.9	50.2
Female	48.8	46.4	51.2	49.1	49.8
Age					
16–29	20.6	19.1	19.0	21.6	21.1
30–39	16.4	15.3	16.8	14.1	16.8
40–49	18.7	17.8	17.7	18.0	17.4
50–59	22.2	24.2	19.7	24.6	21.5
60+	22.0	23.5	26.8	21.6	23.3
Education					
Low	12.1	11.8	11.8	10.2	30.9
Medium	33.0	35.6	33.2	32.7	33.9
High	54.9	52.6	55.0	57.1	35.0
Marital status					
Married	55.7	57.0	53.0	55.6	53.1
Widowed	2.8	2.1	3.2	2.0	3.7
Divorced	8.8	8.9	9.5	9.5	8.5
Single	32.7	32.0	34.3	32.9	34.7
Household size					
Four+ persons	27.6	24.9	25.0	21.8	22.1
Three persons	17.8	20.0	20.8	22.0	18.5
Two persons	37.8	41.1	39.6	38.9	37.1
Single household	16.9	14.1	14.6	17.3	22.3
Nationality					
German	93.2	97.9	94.4	94.8	86.6
Other	6.8	2.1	5.6	5.2	13.4

Note. MZ = German Mikrozensus, 2018.
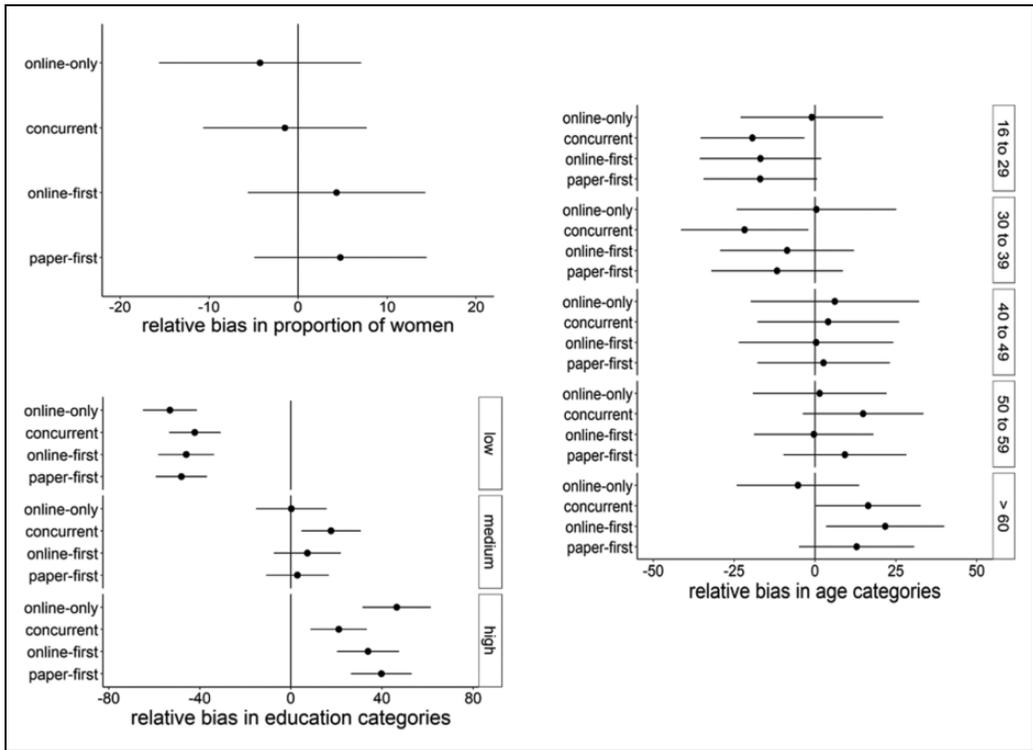
Figure A3. Relative biases in gender, age, and education at the recruitment survey stage across experimental groups. *Note.* The relative bias estimates are denoted by the dots and 95%-confidence intervals by the horizontal lines.

Table A4. Absolute Relative Biases (%) Across Experimental Groups in the Recruitment and Registration Stage.

Group	Minimum (%)	First Quartile (%)	Median (%)	Mean (%)	Third Quartile (%)	Maximum (%)
Recruitment						
Online-only	0.2	1.0	4.3	13.2	6.1	53.2
Concurrent	1.5	14.9	17.7	17.7	21.0	42.2
Online-first	0.3	4.3	8.7	15.5	21.7	46.0
Paper-first	2.7	4.8	11.8	16.5	16.9	48.1
Registration						
Online-only	0.1	2.2	5.1	13.0	19.4	60.7
Concurrent	1.0	7.5	8.8	16.7	12.9	61.6
Online-first	0.0	2.2	9.4	14.2	13.2	61.5
Paper-first	0.9	3.9	8.2	16.4	18.1	66.7
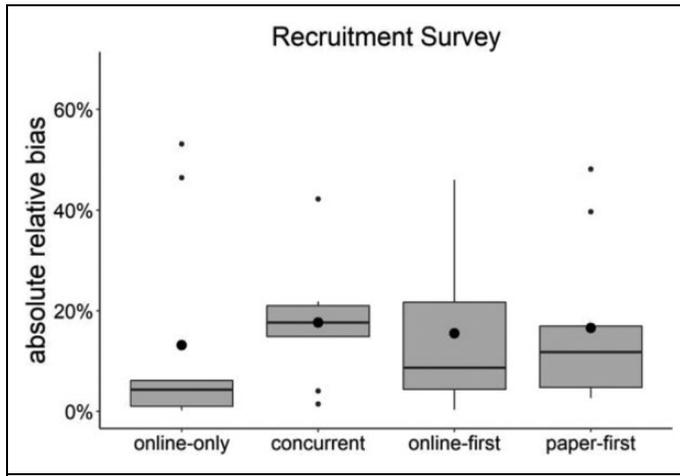
Figure A4. Distribution of absolute relative biases of all categories of the available socio-demographic variables in percent at the recruitment survey stage across experimental groups. *Note.* The large black dots denote the absolute average relative bias in each of the experimental groups. The horizontal bar in bold print denotes the median absolute relative bias, while the boxes and whiskers represent the interquartile range and the small dots represent outliers of a standard box-and-whiskers plot. The outliers are mostly due to the comparatively large biases on education.

Table A5. Internet Use (%) Across Experimental Groups in the Recruitment and Registration Stage.

Internet Use	Online-Only (%)	Concurrent (%)	Online-First (%)	Paper-First(%)
Recruitment				
Daily	84.50	75.40	77.90	74.60
More than once a month	14.10	18.70	16.10	19.40
Once a month or less	1.40	5.90	5.90	6.10
Registration				
Daily	86.20	86.30	86.20	84.80
More than once a month	13.40	13.20	13.10	14.60
Once a month or less	0.40	0.50	0.70	0.60

Appendix B

The Sampling Process Explained

The GIP 2018 sampling scheme was designed in close cooperation with the survey statistics team at GESIS—Leibniz Institute for the Social Sciences. First, a full list of all local communities was acquired from the German Federal Statistical Agency (see Destatis, 2021). The most recent list of communities available at the time of sampling was from 2017. In it, 11,059 communities were listed. Their median population size lay between 1,000 and 1,999 inhabitants, although some communities are considerable larger (e.g., more than 500,000 inhabitants) while some other communities are considerably smaller (e.g., less than 100 inhabitants).

One hundred eighty of these local communities were randomly sampled, stratified by region (operationalized via government districts, which each contain between one and 1,112 local communities) and degree of urbanicity (operationalized via official data on the population size per local community). The 180 sampled local communities were approached with requests to draw samples of

their inhabitants from their population registers. This two-stage sampling process (i.e., first sampling local communities and then individuals within these communities) was necessary, because in Germany, each local community keeps its own, decentralized population register. A centralized population register on the federal level does not exist.

One of the particularities of the German population register system is that the local communities, which hold the decentralized registers, are only encouraged to draw samples for academic research projects but not strictly required to do so by law. One common problem resulting from this rule is that small local communities can be unwilling to draw relatively large samples as compared to their total population size for data protection reasons (e.g., a fear that study participants may be easy to de-anonymize given the small number of people in the community). One solution for this problem is to increase the number of primary sampling units, so that within each community less cases have to be drawn. However, the more surveys pursue this strategy, the higher becomes the burden of requests to sample from their register for each German community, thus decreasing their willingness to comply. To avoid asking small communities to sample a large proportion of their inhabitants and at the same time avoid sampling a large number of local communities, we applied a complex sampling scheme. This sampling scheme was designed to ensure approximately equal inclusion probabilities among all registered German population members within the GIP's envisioned age range (people of ages 16–75) despite the practical constraints described above.

This goal was achieved by assigning all communities to one of the three mutually exclusive domains depending on their population size (i.e., large, small, and medium-sized). Large communities were defined as having at least 499,845 inhabitants. Small communities were defined as having less than 3,800 inhabitants. Medium-sized communities were defined as having between 3,800 and 499,844 inhabitants. The first domain (i.e., large communities) consists of the 15 largest German communities, which we all drew into our sample of local communities. Within these large communities, individuals were sampled via simple random sampling with sample sizes proportional to the community size.

The second domain consists of 7,522 communities with a population size below 3,800 (i.e., small communities). From the local communities in the second domain, 36 communities were sampled with equal probability using a balance sampling strategy (see Tillé, 2006, Chapter 8) accounting for the population size, government district, and degree of urbanicity. From these communities, a fixed number of persons was sampled from the registers.

The third domain consists of all other communities (i.e., 3,522 medium-sized communities). Within the third domain, 129 communities were sampled proportional to population size, stratified by government district and degree of urbanicity. From these communities, a fixed number of persons were sampled from the registers.

As a result of the sampling design being different in the three domains, inclusion probabilities are equal among all population members within each of the three domains but not entirely among all population members across the three domains. To take this fact into account, we used design weights calculated as the inverse of the inclusion probabilities in all of our analyses.

In addition to using the design weights, we account for the clustering and stratification in our sampling design in the variance estimation: We estimate the variances of the Horvitz–Thompson estimator separately for each of the three domains accounting for their specific sampling design described above. Then, we pool these variances to calculate the total variance. This was done by weighting each domain's variance by the domain's sample size and then calculating the weighted average of the three variance estimates.

From the 13,050 individuals drawn into the GIP gross sample, 9,600 individuals were randomly drawn to be part of the experiment that we report on in our article. One hundred ninety-five of those were classified as "not eligible" in accordance with AAPOR (2016, p. 33) Standard Definitions,

because according to information provided by the postal mail service, they no longer lived at the address that they were registered at.

Authors' Note

Declaration of Conflicting Interests

Funding

Supplemental Material

The supplemental material (questionnaires, invitation letters, reminder letters, etc.) is available in the online version of the article.

Notes

1. A notable exception of a probability-based online panel in Europe that uses postal mail to establish initial contact with the sample members is the Norwegian Citizen Panel (Høgestøl & Skjervheim, 2014).
2. It should be noted that self-administered surveys have disadvantages, too, including differential nonresponse (see, e.g., Dillman et al., 2014).
3. German Internet Panel data are available as Scientific Use Files via the GESIS Data Archive for the Social Sciences (https://dbk.gesis.org/dbksearch/gdesc2.asp?no=0109&tab=&ll=10&notabs=1&db=e).
4. Online panel registration surveys were considered completed if they were filled out until the last web survey page, even if they included selective item nonresponse (i.e., if some questions were not answered).
5. Paper questionnaires were considered completed if they were returned, even if they included selective item nonresponse (i.e., if some questions were not answered). However, paper questionnaires were not considered completed if they were returned empty (i.e., without any answers to survey questions).
6. Online questionnaires were considered completed if they were filled out until the last web survey page, even if they included selective item nonresponse (i.e., if some questions were not answered). However, online questionnaires were not considered completed if respondents broke off the survey during the interview (i.e., did not fill out the questionnaire until the last survey page). Overall, 9.6% of sample members who started the online version of the recruitment survey broke it off.
7. It should be noted that in cases where a sample member filled out and returned both the paper and the online version of the recruitment survey (24 cases in total), we prioritized the online version to avoid duplicate cases in the data. In the numerator of Equation 2, we counted these cases once.
8. The population total is assumed to be the true value in the population, that is, without sampling variance.
9. The analyses rely on data from the 2018 German Mikrozensus. The data were provided upon request by the German Federal Statistics Office (Destatis, 2018).

References

The American Association for Public Opinion Research. (2016). *Standard definitions: Final dispositions of case codes and outcome rates for surveys* (9th ed.). https://www.aapor.org/aapor_main/media/publications/standard-definitions20169theditionfinal.pdf

Baker, R., Blumberg, S. J., Brick, M. J., Couper, M. P., Courtright, M., Dennis, J. M., Dillman, D., Frankel, M. R., Garland, P., Groves, R. M., Kennedy, C., Krosnick, J., Lavrakas, P. J., Lee, S., Link, M., Piekarski, L., Thomas, R. K, & Zahs, D. (2010). AAPOR report on online panels. *Public Opinion Quarterly*, *74*(4), 711–781.

Bethlehem, J., & Biffignandi, S. (2012). *Handbook of web surveys*. John Wiley & Sons.

Bianchi, A., Biffignandi, S., & Lynn, P. (2017). Web-face-to-face mixed mode design in a longitudinal survey: Effects on participation rates, sample composition and costs. *Journal of Official Statistics*, *33*(2), 385–408.

Blom, A. G., Bosnjak, M., Cornilleau, A., Cousteaux, A.-S., Das, M., Douhou, S., & Krieger, U. (2016). A comparison of four probability-based online and mixed-mode panels in Europe. *Social Science Computer Review*, *34*(1), 8–25.

Blom, A. G., Cornesse, C., Friedel, S., Krieger, U., Fikel, M., Rettig, T., Wenz, A., Juhl, S., Lehrer, R., Möhring, K., Naumann, E., & Reifenscheid, M. (2020). High-frequency and high-quality survey data collection. *Survey Research Methods*, *14*(2), 171–178.

Blom, A. G., Gathmann, C., & Krieger, U. (2015). Setting up an online panel representative of the general population: The German Internet Panel. *Field Methods*, *27*(4), 391–408.

Blom, A. G., & Herzing, J. M. E. (2017). Face-to-face Rekrutierung für ein probabilistisches Onlinepanel: Einfluss auf die Repräsentativität. In S. Eifler & F. Faulbaum (Eds.), *Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung* [Face-to-face Recruitment for a Probability-Based Online Panel: Impact on Representativeness] (pp. 99–118). VS Verlag.

Blom, A. G., Herzing, J. M. E., Cornesse, C., Sakshaug, J., Krieger, U., & Bossert, D. (2017). Does the recruitment of offline households increase the sample representativeness of probability-based online panels? Evidence from the German Internet Panel. *Social Science Computer Review*, *35*(4), 498–520.

Bosnjak, M., Dannwolf, T., Enderle, T., Schaurer, I., Struminskaya, B., Tanner, A., & Weyandt, K. W. (2018). Establishing an open probability-based mixed-mode panel of the general population in Germany: The GESIS Panel. *Social Science Computer Review*, *36*(1), 103–115.

Callegaro, M., Baker, R. P., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (2014). *Online panel research: A data quality perspective*. John Wiley & Sons.

Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., de Leeuw, E. D., Legleye, S., Pasek, J., Pennay, D., Phillips, B., Sakshaug, J. W., Struminskaya, B, & Wenz, A. (2020). A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research. *Journal of Survey Statistics and Methodology*, *8*(1), 4–36.

Cornesse, C., & Bosnjak, M. (2018). Is there an association between survey characteristics and representativeness? A meta-analysis. *Survey Research Methods*, *12*(1), 1–13.

Cornesse, C., & Schauer, I. (2021). The long-term impact of different offline population inclusion strategies in probability-based online panels: Evidence from the German Internet Panel and the GESIS Panel. *Social Science Computer Review*. https://doi.org/10.1177/0894439320984131

Couper, M. P. (2011). The future of modes of data collection. *Public Opinion Quarterly*, *75*(5), 889–908.

Couper, M. P., Kapteyn, A., Schonlau, M., & Winter, J. (2007). Noncoverage and nonresponse in an Internet survey. *Social Science Research*, *36*(1), 131–148. https://doi.org/10.1016/j.ssresearch.2005.10.002

Couper, M. P., & Miller, P. V. (2008). Web survey methods: Introduction. *Public Opinion Quarterly*, *72*(5), 831–835.

De Leeuw, E. D. (2013). Mixed-mode surveys and the Internet. *Survey Practice*, *3*(6). https://doi.org/10.29115/SP-2010-0030

De Leeuw, E. D. (2018). Mixed-mode: Past, present, and future. *Survey Research Methods*, *12*(2), 75–89.

De Leeuw, E. D., & Berzelak, N. (2016). Survey mode or survey modes. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.) *The Sage handbook of survey methodology* (pp.142–156). SAGE Publications.

De Leeuw, E. D., & Toepoel, V. (2018). Mixed-mode and mixed-device surveys. In D. L. Vannette & J. A. Krosnick (Eds.), *The Palgrave handbook of survey research* (pp. 51–61). Palgrave Macmillan.

Destatis. (2018). *Erhebung über die private Nutzung von Informations- und Kommunikationstechnologien* [Data Collection Regarding the Private Use of Information and Communication Technologies]. Statistisches Bundesamt. https://www.destatis.de/de/methoden/qualitaet/qualitaetsberichte/einkommen-konsum-lebens bedingungen/ikt-private-haushalte-2018.pdf?__blob=publicationfile

Destatis. (2021). *Gemeinden nach Bundesländern und Einwohnergrößenklassen am 31.12.2018* . Statistisches Bundesamt [Local Communities Across Federal States and Population Size Classes on December 31, 2018. Federal Statistical Agency]. https://www.destatis.de/de/themen/laender-regionen/regionales/gemeindever zeichnis/administrativ/archiv/standardtabellen/08_gemeinden_einwohnergroessenklassenvorjahr.html

Dillman, D. A. (2017). The promise and challenge of pushing respondents to the web in mixed-mode surveys. *Survey Methodology*, *43*(1), 3–30.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons.

Edwards, M. L., Dillman, D. A., & Smyth, J. D. (2014). An experimental test of the effects of survey sponsorship on Internet and mail survey response. *Public Opinion Quarterly*, *78*(3), 734–750.

Felderer, B., Kirchner, A., & Kreuter, F. (2019). The effect of survey mode on data quality: Disentangling nonresponse and measurement error bias. *Journal of Official Statistics*, *35*(1), 93–115.

GALLUP. (2020). *How does the Gallup panel work?* https://www.gallup.com/174158/gallup-panel-methodol ogy.aspx

Groves, R. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*(5), 646–675.

Helsper, E. J., & Reisdorf, B. C. (2017). The emergence of a "digital underclass" in Great Britain and Sweden: Changing reasons for digital exclusion. *New Media & Society*, *19*(8), 1253–1270.

Herzing, J. M. E., & Blom, A. G. (2019). The influence of a person's IT literacy on unit nonresponse and attrition in an online panel. *Social Science Computer Review*, *37*(3), 404–424.

Høgestøl, A., & Skjervheim, Ø. (2014). *Norwegian Citizen Panel 2013: First wave methodology report*. NSD. https://nsd.no/data/individ/publikasjoner/nsd2065documentationreport.pdf

Holmberg, A., Lorenc, B., & Werner, P. (2010). Contact strategies to improve participation via the web in a mixed-mode mail and web survey. *Journal of Official Statistics*, *26*(3), 465–480.

Hupp, A. (2015). *Sustainability cultural indicators program (SCIP): Methodology report*. University of Michigan. http://graham.umich.edu/media/files/scip_methodologyreport_2018.pdf

Ipsos. (2020). *Ten differentiators of the Ipsos knowledge panel*. https://www.ipsos.com/sites/default/files/20-04-55

Jäckle, A., Lynn, P., & Burton, J. (2015). Going online with a face-to-face household panel: Effects of a mixed mode design on item and unit non-response. *Survey Research Methods*, *9*(1), 57–70.

Klausch, T., Schouten, B., & Hox, J. J. (2017). Evaluating bias of sequential mixed-mode designs against benchmark surveys. *Sociological Methods & Research*, *46*(3), 456–489.

Leenheer, J., & Scherpenzeel, A. C. (2013). Does it pay off to include non-Internet households in an Internet panel? *International Journal of Internet Science*, *8*(1), 17–29.

Lynn, P. (2013). Alternative sequential mixed-mode designs: Effects on attrition rates, attrition bias, and costs. *Journal of Survey Statistics and Methodology*, *1*(2), 183–205.

Mauz, E., von der Lippe, E., Allen, J., Schilling, R., Müters, S., Hoebel, J., Schmich, P., Wetzstein, M., Kamtsiuris, P, & Lange, C. (2018). Mixing modes in a population-based interview survey: Comparison of a sequential and a concurrent mixed-mode design for public health research. *Archives of Public Health*, *76*(1), 8.

McMaster, H. S., LeardMann, C. A., Speigle, S., & Dillman, D. A. (2017). An experimental comparison of web-push vs. paper-only survey procedures for conducting an in-depth health survey of military spouses. *BMC Medical Research Methodology*, *17*(1), 73.

Medway, R. L., & Fulton, J. (2012). When more gets you less: A meta-analysis of the effect of concurrent web options on mail survey response rates. *Public Opinion Quarterly*, *76*(4), 733–746.

Messer, B. L. (2012). *Pushing households to the web: Experiments of a "Web+Mail" methodology for conducting general public surveys* [PhD dissertation]. Washington State University.

Millar, M. M., & Dillman, D. A. (2011). Improving response to web and mixed-mode surveys. *Public Opinion Quarterly*, *75*(2), 249–269.

National Opinion Research Center. (2019). *Technical overview of the AmeriSpeak Panel. NORC's probability-based household panel*. https://amerispeak.norc.org/documents/research/amerispeak%20technical%20overview%202019%2002%2018.pdf

Olson, K., Smyth, J. D., & Wood, H. M. (2012). Does giving people their preferred survey mode actually increase survey participation rates? An experimental examination. *Public Opinion Quarterly*, *76*(4), 611–635.

Patrick, M. E., Couper, M. P., Laetz, V. B., Schulenberg, J. E., O'Malley, P. M., Johnston, L. D., & Miech, R. A. (2018). A sequential mixed-mode experiment in the US National Monitoring the Future Study. *Journal of Survey Statistics and Methodology*, *6*(1), 72–97.

Pew Research Center. (2019). *Growing and improving Pew Research Center's American Trends Panel*. https://www.pewresearch.org/methods/2019/02/27/growing-and-improving-pew-research-centers-american-trends-panel/

Rao, K., Kaminska, O., & McCutcheon, A. (2010). Recruiting probability samples for a multi-mode research panel with Internet and mail components. *Public Opinion Quarterly*, *74*(1), 68–84.

Revilla, M., Cornilleau, A., Cousteaux, A.-S., Legleye, S., & de Pedraza, P. (2016). What is the gain in a probability-based online panel of providing Internet access to sampling units who previously had no access? *Social Science Computer Review*, *34*(4), 479–496.

Scherpenzeel, A., & Toepoel, V. (2012). Recruiting a probability sample for an online sample: Effects of contact mode, incentives, and information. *Public Opinion Quarterly*, *76*(3), 470–490.

Shih, T.-H., & Fan, X. (2007). Response rates and mode preferences in web-mail mixed-mode surveys: A meta-analysis. *International Journal of Internet Science*, *2*(1), 59–82.

Smyth, J. D., Dillman, D. A., Christian, L. M., & O'Neill, A. C. (2010). Using the Internet to survey small towns and communities: Limitations and possibilities in the early 21st century. *American Behavioral Scientist*, *53*(9), 1423–1448.

Smyth, J. D., Olson, K., & Millar, M. M. (2014). Identifying predictors of survey mode preference. *Social Science Research*, *48*, 135–144.

Tillé, Y. (2006). *Sampling algorithms*. Springer.

West, B. T., & Blom, A. G. (2017). Explaining interviewer effects: A research synthesis. *Journal of Survey Statistics and Methodology*, *5*(2), 175–211.

Wolf, C., Christmann, P., Gummer, T., Schnaudt, C., & Verhoeven, S. (in press). Conducting general social surveys as self-administered mixed-mode surveys. *Public Opinion Quarterly*.

Author Biographies

Carina Cornesse is a postdoctoral researcher at the Collaborative Research Center SFB 884 "Political Economy of Reforms," University of Mannheim, and co- leader of the project "Recruiting probability-based online panels: Cost-efficiency and data quality" at the Mannheim Center for European Social Research (MZES). Her research focuses on innovation in quantitative social science research methods, in particular the recruitment and maintenance of online panels, the quality of probability-based and nonprobability survey samples, as well as mixed-mode and mixed-device data collection.

Barbara Felderer is the head of the team Survey Statistics at the Department of Survey Design and Methodology, GESIS – Leibniz-Institute for the Social Sciences and an associate researcher at the Collaborative Research Center SFB 884 "Political Economy of Reforms," University of Mannheim. Her research focuses on survey nonresponse, especially nonresponse bias analysis and nonresponse adjustment in (mixed-mode) panel surveys.

Marina Fikel is a researcher at the Collaborative Research Center SFB 884 "Political Economy of Reforms", University of Mannheim. Her work focuses survey practice, in particular online panel recruitment and management.

Ulrich Krieger is a post-doc researcher and a survey manager at the Collaborative Research Center SFB 884 "Political Economy of Reforms," University of Mannheim, specializing in longitudinal online surveys. He serves as head of the German Internet Panel, a probability based online panel of the German population. His research focuses on survey participation, the use of respondent incentives and recruiting online panels.

Annelies Blom is a Professor for Data Science at the School Social Sciences, University of Mannheim. She is also the Director of the German Internet Panel (GIP) at the Collaborative Research Center SFB 884 "Political Economy of Reforms". Her research focuses on data science applications in social research, digital exclusion, and ethical questions in this area, as well as response behavior, measurement and nonresponse errors.