

Contact Modes and Participation in App-Based Smartphone Surveys: Evidence From a Large-Scale Experiment

Social Science Computer Review
2022, Vol. 40(5) 1076–1092

© The Author(s) 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0894439321993832
journals.sagepub.com/home/ssc



Mario Lawes¹, Clemens Hetschko^{1,2,3}, Joseph W. Sakshaug^{4,5,6},
and Stephan Griebemer⁴

Abstract

Smartphone apps are increasingly being used for population-based survey research. Recruiting people to sign up for an app-based survey is, however, less straightforward compared to traditional surveys, which risks inflating nonresponse as well as the potential for nonresponse bias. By means of an experiment with over 44,000 recently registered job seekers, we present causal evidence on the effects of using different contact modes (email, postal letter, or preannouncement letter and email) on participation rates in an app-based panel survey. Further, using detailed administrative register data, we investigate whether contact modes differentially affect nonresponse bias. We also examine whether the mode of making contact has a lasting effect on panel participation rates and participation rates in momentary assessments collected using the experience sampling method (ESM). Overall, the preannouncement letter and email invitation strategy maximizes participation compared to stand-alone letters and emails, which do not differ significantly in terms of participation rates. Stand-alone letters and the preannouncement approach perform better than emails when it comes to panel participation and submitted ESM episodes.

Keywords

app-based research, survey participation, contact modes, experience sampling method (ESM), panel stability

¹ Freie Universität Berlin, Germany

² University of Leeds, United Kingdom

³ CESifo, Munich, Germany

⁴ Institute for Employment Research (IAB), Nuremberg, Germany

⁵ Ludwig-Maximilians-Universität Munich, Germany

⁶ University of Mannheim, Germany

Corresponding Author:

Mario Lawes, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany.

Email: mario.lawes@fu-berlin.de

Collecting survey data using smartphone apps has become popular among social scientists. There are a number of advantages over traditional web surveys or face-to-face and telephone surveys, such as convenient high-frequency assessment, reduced burden and increased flexibility for participants, push notifications and reminders, and integrated interventions and incentives (see, e.g., Miller, 2012). Given these advantages and the growing use of smartphones in society (Pew Research Center, 2019), smartphone-based research plays an increasing role in the social sciences.

Some types of data collected through smartphone apps would be impossible to gather by other means without incurring high costs or posing extensive burden to participants. Thus, smartphones enable researchers to conveniently and cost-efficiently unlock new measurement possibilities and new sources of data with great potential for social science research. This becomes most apparent through the collection of passive mobile data, including geolocation, movement patterns, browsing, and communication behavior (e.g., Keusch et al., 2019). In addition, studies benefit from smartphone-based surveying if they require people to actively provide information in real time. One example is the experience sampling method (ESM; Hektner et al., 2007), the “gold standard” of measuring affect (e.g., Kahneman & Krueger, 2006; Organization for Economic Cooperation and Development [OECD], 2013). In the ESM, subjects are asked to report at randomly chosen time points their current activity and emotions. Before the smartphone era, researchers had to equip subjects with pagers. Now, participants only need to install a survey app and wait for the ESM notifications, which renders the data collection less expensive and less burdensome (see, e.g., Bryson & MacKerron, 2017).

Despite the rise of smartphone usage and its potential for research, the act of downloading, registering, and actively engaging with an app for frequent measurement might be too burdensome for some members of the general population (Read, 2019), which in turn might cut participation rates and enlarge nonresponse bias. Up to now, there is a paucity of information on methods for maximizing participation rates in app-based surveys. One crucial design decision for any online survey is the contact mode (e.g., email, postal letter) used to deliver the invitation, communicate the study details, and encourage potential participants to sign up (Menegaki et al., 2016; Millar & Dillman, 2011; Sakshaug et al., 2019). What is currently lacking is knowledge about which contact mode is optimal for maximizing participation and minimizing nonresponse bias as well as ensuring continuous active engagement with the app after it is installed.

By means of a contact mode experiment, we present causal evidence on whether potential participants of an app-based panel survey rather respond to a traditional invitation letter than to an email invitation or to a preannouncement letter followed by an email invitation. We also test whether the mode of making contact affects nonresponse bias, panel participation rates, and response rates to ESM measurements. The experiment was carried out when inviting more than 44,000 recently registered job seekers to take part in the German Job Search Panel (GJSP; Hetschko et al., 2020), a monthly app-based smartphone survey that collects data on well-being and health. A unique advantage of the GJSP is that the sample can be linked to rich administrative records, enabling detailed analysis of nonresponse bias. In addition, we investigate whether contact mode has a lasting impact on panel participation rates and ESM participation patterns after sign-up. Overall, the GJSP allows for analyzing the impact of the contact mode on a number of participation and data quality outcomes that are of interest for survey research based on smartphones and beyond.

Background

Contact modes in app-based surveys. A variety of contact modes, including email, postal letters, or a combination of both, have been used to invite subjects to participate in app-based surveys (e.g., Jäckle et al., 2019; Kreuter et al., 2019; McGeeney & Weisel, 2015; Scherpenzeel, 2017; Wells et al., 2013). Each contact mode has strengths and limitations. Postal letters likely better highlight

the study sponsors and research aims, which in turn increases the perceived legitimacy of the study (Dillman et al., 2014). Postal letters may also enclose a flyer or information pamphlet that renders the invitation more appealing and is handy at later points in time. But mailing letters also increases study costs, especially for large samples, and requires extra effort on the part of participants to download the app using the enclosed QR code or URL as contact mode and survey mode differ. The added costs of mailing letters may, however, be justified if this approach increases participation rates relative to cheaper alternatives. Email contacts are an inexpensive means of communicating the study request and offer convenience for participants to directly follow a link to download the app. That is, if people receive the email on their smartphone, they do not need to switch modes to download the app. Further provided links can lead to websites providing more information than a letter or a flyer. However, email invitations may be perceived as spam or fail to convey the legitimacy of the study. One approach to making an email invitation seem less unsolicited and increase the perceived legitimacy of the study is to send a preannouncement letter (Dykema et al., 2011; Harmon et al., 2005). Such letters are used to describe the study, highlight its sponsorship, and provide advance notice of the impending email invitation to avoid the perception of spam. The preannouncement approach therefore combines the benefits of paper and email into a single design, while also increasing the total exposure to the recruitment message.

Contact mode and sign-up rates. Sign-up rates in app-based surveys have primarily been investigated in nonexperimental studies using relatively cooperative samples of subjects who (1) are already active members of a panel survey, (2) have a preexisting relationship with the survey organization, or (3) have already been prescreened for eligibility and stated willingness to take part in the app study. No study has experimentally assessed the impact of contact mode for recruiting cross-sectional samples of general population members for app-based surveys. The only related evidence comes from web survey experiments, which have yielded mixed results. Sakshaug et al. (2019) found a higher sign-up rate (8%) when paper invitations were used as opposed to email invitations (4%) in a web-only survey of establishments. Kaplowitz et al. (2012) found a slightly higher web survey sign-up rate among university students (22% vs. 19%) and faculty (40% vs. 33%) when an email (vs. paper) invitation was used but no difference among staff (43% vs. 43%). Birnholtz et al. (2004) found no statistically significant difference in sign-up rates between paper and email (40% vs. 32%) invitations in a web survey of engineering researchers. Preceding an email invitation with a preannouncement letter has been shown to be effective in increasing sign-up rates in several web survey experiments (Crawford et al., 2004; Dykema et al., 2011; Harmon et al., 2005; Kaplowitz et al., 2004; Porter & Whitcomb, 2007). While we would expect a similar positive effect of preceding the email invitation with a preannouncement letter in an app-based smartphone survey, to date, no experimental study has tested this hypothesis.

Contact modes and nonresponse bias. Beyond sign-up rates, it is also important to consider the impact of different contact mode strategies on nonresponse bias. People interested in participating in an app survey are likely nonrepresentative, and the mode of contact may vary the degree of selective participation even further. In survey studies, researchers have found a weak relationship between nonresponse rates and nonresponse bias (Groves, 2006; Groves & Peytcheva, 2008). That is, a recruitment strategy that yields a higher participation rate is likely, but not guaranteed, to have lower nonresponse bias over a recruitment strategy that yields a lower participation rate. Only few studies have analyzed nonresponse bias in app-based surveys (Jäckle et al., 2019; McGeeney & Weisel, 2015). However, as mentioned above, these analyses are based on existing panel studies where the invited subjects are drawn from a relatively cooperative, preexisting sample pool compared to a newly drawn population sample. App-based surveys conducted on new population samples are likely to have even lower participation rates than the panel study samples reviewed

above and, thus, a higher risk of nonresponse bias. Insights into how different contact modes affect participation rates and nonresponse bias in app-based surveys based on new population samples are currently lacking.

Panel participation and ESM. Another aspect not addressed in the literature is the potential influence of contact mode on continued adherence in app surveys. Given the large variation in dropout rates among initial app participants (Jäckle et al., 2019; McGeeney & Weisel, 2015), it is unclear whether participants who receive the study invitation through different modes will be differentially motivated to continue using the app. A special focus in this regard is whether or not participants respond to momentary assessments since these require participants to promptly respond to survey questions at random time points throughout the day. We expect that certain contact mode strategies, particularly those that incorporate paper, are likely to do a better job conveying the legitimacy of the study and will, in turn, optimize initial participation and participation in follow-up waves of the panel as well as participation in the ESM measurements.

Research Questions (RQs)

To address the identified gaps in the literature, we address the following RQs:

1. To what extent do sign-up rates to an app-based smartphone survey vary by contact mode, including postal letter, email, and preannouncement letter followed by email?
2. Do different contact modes differentially affect nonresponse bias?
3. Do subjects recruited via different contact modes differ in their (a) panel participation rates and (b) participation rates in ESM measurements?

Data and Methods

The GJSP

The contact mode experiment was embedded in the recruitment phase of the GJSP (Hetschko et al., 2020). The GJSP is a 2-year app-based panel survey collecting monthly data on the well-being and health of recently registered job seekers in Germany. Each month from December 2017 to May 2019, a new sample of subjects was recruited for the GJSP. However, the present experiment ran from November 2018 to May 2019 only. The job search notification process of the German Federal Employment Agency (*Bundesagentur für Arbeit*) was used to identify employed Germans in the age range between 18 and 60 who were about to lose their job. To be eligible for unemployment benefits, people have to register as job seeking with the employment agency 3 months before expecting to terminate an employment. If they learn about the prospect of losing their job at a later point in time, they have 3 days to notify the employment agency.

On a monthly basis, two separate samples of registered job seekers were drawn. The first sample consisted of all Germans who had registered as job seekers between the 15th day of the second to last month and the 14th day of the last month and had been employed in companies that likely conducted mass layoffs since many employees of the same company registered as job seekers at the same time (see Hetschko et al., 2020). Using this novel identification strategy, we were able to contact the whole population of recently registered German job seekers ($N = 38,180$) from companies conducting mass layoffs in Germany during the time period between September 15, 2018, and April 14, 2019. A second monthly random sample was drawn (with equal size to the first) from all job seekers who were not identified as being part of a mass layoff during the same time period as the first sample ($N = 38,178$ in total). Of the 76,358 sampled subjects pooled across both samples and all

seven months of the recruitment period, 44,696 subjects had provided an email address to the German Federal Employment Agency and were eligible for the contact mode experiment. Individuals who did not provide an email address ($N = 31,662$) were contacted via postal mail and are not considered in the analysis (see Figure A1 in the Online Supplement for a flowchart).

Procedure

Experimental conditions. Individuals within each of the seven monthly samples were randomly assigned to be contacted via postal letter ($N = 14,896$), via email ($N = 14,898$), and via preannouncement letter followed by email ($N = 14,902$). All three conditions provided subjects with the same information about the app study and data privacy. Translated versions of all materials can be found in the Online Supplement.

Letter. In this condition, a physical letter was sent to the homes of the potential participants. The letter introduced the basic features of the GJSP and provided information on data privacy. Included with the letter was the personal token needed to register for the study as well as a flyer summarizing key features of the app study. Further, the flyer contained QR codes directing subjects to the Google Play Store or Apple App Store to download the survey app.

Email. In this condition, individuals received an email introducing the basic features of the GJSP as well as a hyperlink enabling subjects to access data privacy information and further details about the survey, including the information provided by the flyer in the letter condition. The personal token to sign up for the study was also part of the email. URL links to the Google Play Store and Apple App Store were also included to simplify downloading the survey app.

Preannouncement letter + follow-up email. In this condition, individuals received a physical preannouncement letter describing the basic features of the GJSP as well as a notification that they will shortly receive an email containing a personal token to sign up for the study. As in the letter condition, the preannouncement letter also included a printout of the data protection policy and the flyer. The email invitation following the preannouncement letter was identical to the one used in the email condition.

Timing of the invitations and sign-up. The German Federal Employment Agency mailed out both letter contacts (i.e., preannouncement and letter invitation) on the same day using the German postal service. As 93% of letters sent via the German postal service are delivered to their recipients on the next working day (Deutsche Post DHL Group, 2018, p. 60), almost all people in the letter and preannouncement groups received the letters in their letterboxes on the same day (i.e., one working day after dispatch). Roughly 2 days after the different letters were delivered, the email containing the personal token needed to sign up for the study was sent simultaneously to those in the two conditions entailing email contact. This approach ensured that people had a high chance of receiving the preannouncement letter before the email, although some people might have received the email first if they did not check their letterbox.

Sign-up was possible either by using the downloaded app on a smartphone or by using a web interface. Actual participation in the panel survey was, however, only possible via the survey app (for details on the survey app see Ludwigs & Erdtmann, 2019). After the registration, subjects filled out an entry survey. For methodological reasons unrelated to the experiment, participants to the entry survey were filtered out of the main study if they reported not being employed anymore ($N = 792$) or if they were still in the probation period of their current position ($N = 127$). Further, one third of all participants were randomly excluded from the GJSP due to a different experiment ($N = 531$). A total of 115 subjects signed up but never submitted the entry survey. The remaining

subjects ($N = 1,035$) were then presented with short questionnaires over the course of 7 days each month (see Hetschko et al., 2020). To motivate continuous participation, subjects were rewarded for their active participation with either a small monthly cash incentive in the form of an Amazon voucher or a cash transfer (if they signed up via the smartphone app) or with a new smartphone (if they signed up via the web interface).

Outcomes and Data Analyses

The following criteria are used to evaluate the different recruitment strategies in accordance with each RQ. All RQs are addressed using two-sided significance tests and a Type I error level of 5%.

Sign-up rates and time to sign-up (RQ1). The sign-up rate is defined as the proportion of subjects who eventually entered their personal token directly into the app or on the study website. We further estimated the time interval between the receipt of the invitation and the actual sign-up for each individual to see whether the invitation mode affects the time to sign up. We operationalized the receipt of the invitation by using the first actually observed sign-up date per experimental group and month of recruitment.

Nonresponse bias (RQ2). For the analysis of nonresponse bias, we use various individual characteristics derived from the administrative records of the German Federal Employment Agency, the so-called Integrated Employment Biographies (IEB; see Antoni et al., 2016, for an account of a 2% subsample of the IEB). The IEB comprise retrospective data from the notification process of the German social insurance system (e.g., the employer must notify the health insurance of the employee when the employment relationship commences or terminates) and from administrative processes at the Federal Employment Agency (e.g., receipt of social benefits in the past). In order to use these data, we have to rely on a reduced sample size ($N = 35,732$) as many contacted job seekers cannot be linked to a corresponding IEB administrative record using the available process data. This is due to the fact that the required data are subject to multiple steps of preparation, editing, and correction until they can be supplied as IEB records to researchers. As these processes take several months, newly registered job seekers are often not assigned to an IEB record yet. All other units are matched directly using exact one-to-one matching on a unique identifier. We have no reason to believe that linked and nonlinked subjects would behave differently in the contact mode experiment, as the participation rates of both groups were comparable within each experimental condition.

In line with the literature on nonresponse bias, this linkage enables us to consider a number of characteristics that tend to be correlated with survey participation, such as age, gender, qualification, whether their employer is based in an Eastern or a Western state of Germany, and income (measured as the daily wage; see, e.g., Eckman & Haas, 2017; West et al., 2018). In addition, we use a wide range of variables that are particularly relevant for labor market research and considered important individual characteristics (Caliendo et al., 2017). Specifically, we include tenure, job type (regular, marginal, other), working time (full-time vs. part-time), and labor market history over the past 10 years (e.g., periods of unemployment benefit receipt/welfare receipt).

Provided that the registration as job seeking took place in 2018, all variables are measured at the time of registering, thus only shortly before people received the study invitation. For 2019, however, the IEB records are not available yet. We therefore assign people who registered as job seekers between January 1, 2019, and April 14, 2019, with IEB data from their last observed employment spell, that is, data from the end of 2018. Thus, we assume temporal stability in the data over a period of at most 3½ months. This is a valid assumption since we require people to have been working in their companies for at least 6 months in order to be eligible to participate in the GJSP. In addition to the IEB variables, we include the recruitment month to account for time trends and a sample

indicator indicating whether or not subjects registered as job seeking from companies conducting mass layoffs.

We calculate the average absolute nonresponse bias (Groves & Peytcheva, 2008; Sakshaug et al., 2019) as an aggregate-level measure of nonresponse bias to examine the extent to which the three experimental conditions differentially affect selection to sign up. The proportions of persons who signed up for the study under the different conditions ($Y_{k,condition}$) are compared to the full sample of job seekers ($Y_{k,sample}$) across all categories ($k = 1, \dots, K$) of the administrative variables. The average absolute nonresponse bias is calculated as the sum of all absolute differences between those who signed up and the full sample ($|Y_{k,condition} - Y_{k,sample}|$) for all categories, divided by the total number of categories.

$$\text{Average absolute bias}_{condition} = \frac{\sum_{k=1}^K |Y_{k,condition} - Y_{k,sample}|}{K}.$$

Binary variables have one category only (e.g., female) whereas metric variables need to be categorized to obtain proportions that can be compared between conditions and sample. If the distribution of a metric variable is highly skewed (e.g., periods of welfare benefit receipt over the past 10 years), we collapse them into a binary variable (any indication/no indication throughout). Other metric variables (e.g., age) are categorized into quartiles of the sampled job seekers.

Apart from average absolute bias, we estimate the likelihood of signing up depending on the aforementioned sociodemographic and other characteristics to reveal group-specific patterns of selective participation. To this end, we employ logistic regression and allow for interaction effects of the contact mode with a number of individual characteristics. This enables us to test whether a certain mode of contact amplifies or weakens selective participation in the smartphone survey, for instance, based on education or gender. To ease the interpretation of the interaction effects, metric variables whose distributions are not highly skewed (age, wage, tenure, and periods of employment in the last 10 years) are included in their continuous form and not as categorical variables reflecting different quartiles. Qualification is only included as a binary variable dividing the sample into roughly equal halves (academic degree or not).

Panel participation (RQ3a). As a measure of panel participation, we compute for each contact mode the percentage of GJSP participants fulfilling the inclusion criteria who answered at least one item in each of the first four waves of the GJSP after the entry survey. We focus on the first four waves as this is the period with the largest expected drop in participation. Logistic regression is used to predict whether subjects were still actively participating in each of the first 4 months of the survey. We fit two separate models: (a) an unconditional model with contact mode as the only predictor to test whether the different experimental conditions yield different panel participation rates and (b) a model including covariates collected during the entry survey and the first wave of the GJSP (which directly follows the entry survey). Specifically, the second model includes age, gender, education (academic degree or not), Big Five personality traits (openness to experience, conscientiousness, extroversion, agreeableness, neuroticism), marital status, and household income as covariates. The model with covariates allows to examine the effects of contact mode after controlling for potential observed differences in the composition of the participants that are caused by the different contact modes and the study exclusion criteria. We use the survey data instead of the administrative records for these analyses for two reasons: First, using the administrative records would reduce the sample size due to the aforementioned linkage issues. Second, the survey data allow for including more demographic and psychological variables. Personality traits, in particular, have been repeatedly shown to be related to panel attrition in Internet surveys (Cheng et al., 2018; Lugtig, 2014).

Participation in the ESM (RQ3b). In the GJSP, affective well-being is measured using the ESM. During the last day of the monthly survey blocks, subjects receive six push notifications with short questionnaires at random time points throughout the day between 9 a.m. and 9 p.m. In these short questionnaires, subjects are asked to indicate their current activity, location, and the people they are with. For each ESM episode, the subjects are further asked to answer a number of questions concerning their emotions and moods (for items, see Steyer et al., n.d., 1997). As an overall measure of data quality, we compute the average number of submitted ESM episodes for each of the first four survey waves for each contact mode. To test whether subjects are differentially likely to participate in the ESM measurements at all, we employ logistic regression to predict for each of the first four waves of the GJSP whether subjects submitted at least one ESM episode within a given month. Again, we run these analyses in two ways. First, using an unconditional model with the contact modes as predictors, and second, adding the same survey variables used in the panel participation analysis as covariates. This approach enables us to test whether differences in ESM participation between the experimental conditions are driven by the different composition of people responding to the different modes or by the different features of the contact modes themselves.

Results

Descriptive Statistics

Table A1 in the Online Supplement presents descriptive statistics for the entire sample of subjects eligible for the contact mode experiment. The three experimental groups are practically the same size and do not differ substantially in the individual characteristics measured in the IEB administrative data. Thus, we conclude that the randomization was successful.

Sign-Up Rates and Time to Sign Up

The sign-up rates for the three different recruitment strategies are depicted in Figure 1A. In the letter condition, 733 (4.92%) individuals signed up for the study, compared to 786 (5.28%) individuals contacted via email, and 1,081 (7.25%) individuals in the preannouncement condition. Logistic regression reveals that the sign-up rate in the preannouncement condition is significantly higher compared to the letter ($b = 0.41, p < .001$) and email conditions ($b = 0.34, p < .001$). The email and letter conditions do not differ with regard to the observed sign-up rates ($b = 0.07, p = .16$).

Figure 1B depicts the median number of days until subjects signed up after they received the study invitations. For 35 subjects, the exact sign-up date cannot be reconstructed due to technical issues. Thus, the following analyses are based on the remaining sample ($N = 2,565$). Individuals contacted via letter signed up on average 5.1 days ($SD = 8.9$, maximum: 113.8) after the first subject in the letter condition signed up for the study, our proxy date of invitation receipt. Subjects in the email condition signed up on average 2.8 days ($SD = 5.6$, maximum: 90.1) after receiving the email invitation. In the preannouncement condition, subjects signed up on average 3.7 days ($SD = 9.4$, maximum: 211.4) after receiving the email invitation. The distribution of days until subjects signed up was heavily skewed, that is, many subjects signed up in the first couple of days, and only few subjects signed up several days later. Applying the Mann–Whitney U test indicates that the subjects in the letter condition needed significantly longer to sign up compared to subjects in the email ($z = -10.3, p < .001$) and preannouncement ($z = -7.95, p < .001$) conditions. The difference regarding the time to sign up between the preannouncement and email condition is also statistically significant ($z = -3.1, p = .002$).

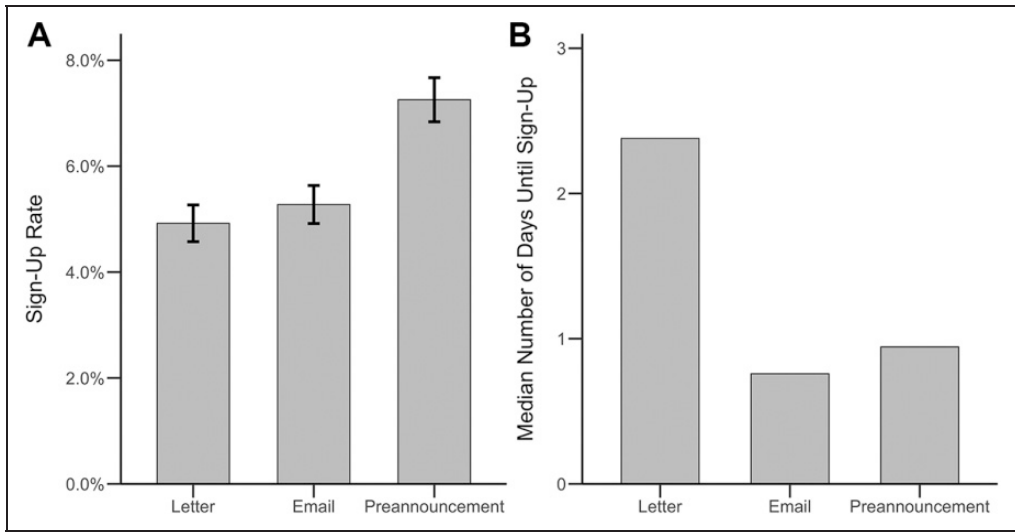


Figure 1. Panel A: Sign-up rates for the three contact modes. Error bars indicate the 95% confidence interval of the sign-up percentages. Panel B: Average (median) number of days subjects needed to sign up for the study in each experimental group.

Nonresponse Bias

Table A2 in the Online Supplement presents the proportions of categories for all IEB administrative variables for subjects who signed up in each experimental condition relative to the entire eligible sample. These values are used in the calculation of the average absolute nonresponse bias for each experimental condition.¹ As Figure 2 shows, the average absolute bias is relatively low across conditions and no greater than 4% in any one condition. There is some variation across conditions, with the letter condition (2.82%) and the preannouncement condition (2.79%) yielding somewhat lower average absolute biases than the email condition (3.63%).

To test whether the contact mode conditions produce more or less selective sign-up with regard to the IEB administrative characteristics, we estimate interaction effects of the conditions

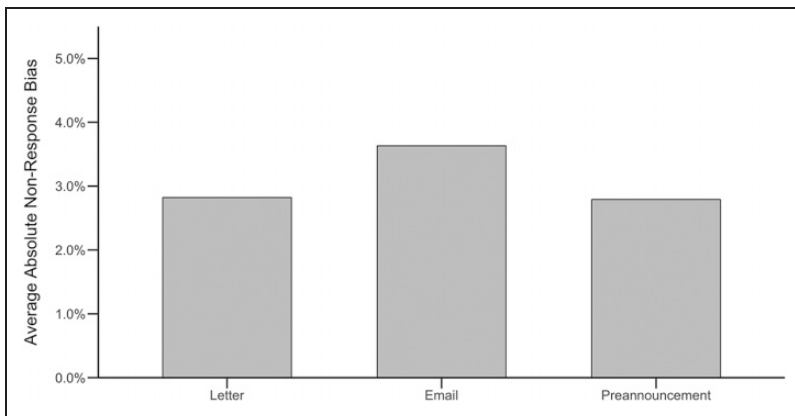


Figure 2. Average absolute nonresponse bias.

with these characteristics. The corresponding logistic regression results are presented in Table 1. Columns 1 (Logit Estimates) and 2 (Marginal Effects) reveal general patterns of selective participation, that is, irrespective of the contact mode. It turns out that age reduces the probability to sign up, whereas being female, having higher levels of earnings (daily wage), and having an academic degree render it more likely that people sign up. In addition, the estimation shows that the significant main effect of the preannouncement condition remains after controlling for numerous characteristics. Columns 3 and 4 display the same sets of results for the interaction model. The positive effect of being female on the willingness to sign up is neutralized by the preannouncement condition. Apparently, men respond much stronger to the preannouncement condition than to the stand-alone email and letter conditions and are just as likely as women to participate under this condition. In addition, we find that the email condition works increasingly better as people's daily earnings get higher. For most of the variables, however, we do not find differential sign-up rates dependent on the mode of contact.

Panel Participation

Figure 3A depicts the percentages of all GJSP participants who fulfilled all the inclusion criteria and submitted at least one item within each of the first four waves (M1–M4) of the experiment ($N = 1,035$). The participation rates in M1, for example, correspond to the percentage of GJSP participants fulfilling all inclusion criteria who answered at least one item in the first wave of the survey. After an initial drop following M1, participation rates stabilize after M2 and decrease more slowly. Particularly in the letter condition, subjects have stable participation rates after M2. In M4, the participation rates in the preannouncement (68.3%) and letter (68.8%) conditions are roughly 10 percentage points higher than those in the email (59.2%) condition. Participation rates in the letter and preannouncement conditions are not statistically different across the first four waves of the GJSP. However, in M2 ($b = 0.47, p = .04$), M3 ($b = 0.52, p = .001$), and M4 ($b = 0.40, p = .009$), participation rates are significantly larger in the preannouncement condition compared to the email condition. In M1, there are no significant differences between these conditions in terms of the participation rate. In M3 ($b = 0.42, p = .015$) and M4 ($b = 0.42, p = .015$), participation rates in the letter condition are significantly higher than in the email condition. In M1 and M2, there are no significant differences in participation between these conditions. When the survey covariates are included in the logistic regression model to control for differences in sample compositions, these differences remain statistically significant. The only exception is that the participation rates in M4 for subjects recruited via preannouncement and email are not statistically different from each other anymore. This result, however, is driven by the reduced sample due to item nonresponse on the covariates and not by the covariates themselves (see Table A3 in the Online Supplement). These results indicate that the observed differences in terms of sample composition in the different experimental groups (due to selective participation) cannot account for the differences in participation rates of the three recruitment conditions. Detailed results of these analyses are depicted in Table A3 in the Online Supplement.

Participation in the ESM

Figure 3B depicts the average number of submitted ESM episodes over the first four waves of the app survey for all GJSP subjects who answered at least one item in the respective wave. Individuals in the letter and preannouncement groups have a similar average number of roughly 3.2 submitted ESM episodes per day, whereas subjects recruited via email submit about 0.4 fewer episodes per day, on average. The unconditional logistic regression analyses reveal that across the first four waves, there are no significant differences between the letter and preannouncement conditions with

Table 1. Determinants of Sign-Up and Interactions With Contact Mode.

Variable	Main Effects		Main Effects and Interactions		
	Logit Estimates	Marginal Effects	Logit Estimates	Marginal Effects	
Email (ref. letter)	0.071 (0.060)	0.004	-0.305 (0.622)		-0.016
Preannouncement (ref. letter)	0.401*** (0.056)	0.021***	0.790 (0.500)		0.040
Age (in years)	-0.013*** (0.003)	-0.001***	-0.009 (0.005)		-0.000
× Email			-0.003 (0.007)		-0.000
× Preannouncement			-0.008 (0.006)		-0.000
Daily wage (in €10)	0.017*** (0.005)	0.001***	0.005 (0.008)		0.000
× Email			0.036** (0.012)		0.002**
× Preannouncement			0.014 (0.012)		0.001
Tenure (years)	0.012 (0.007)	0.001	0.012 (0.012)		0.001
× Email			-0.002 (0.017)		-0.000
× Preannouncement			0.002 (0.016)		0.000
Female	0.209*** (0.048)	0.011***	0.335*** (0.090)		0.017***
× Email			-0.041 (0.125)		-0.002
× Preannouncement			-0.260* (0.118)		-0.013*
East German origin	0.081 (0.060)	0.004	0.088 (0.112)		0.004
× Email			-0.023 (0.158)		-0.001
× Preannouncement			0.017 (0.147)		0.001
Regularly employed, when registering as job seeker (ref. other)	0.087 (0.203)	0.004	0.171 (0.395)		0.009
× Email			0.416 (0.607)		0.021
× Preannouncement			-0.394 (0.483)		-0.020
Marginally employed, when registering as job seeker (ref. other)	0.067 (0.224)	0.003	0.089 (0.434)		0.005
× Email			0.645 (0.653)		0.033
× Preannouncement			-0.394 (0.539)		-0.020
Mass layoff sample	-0.061 (0.047)	-0.003	-0.011 (0.087)		-0.001
× Email			-0.149 (0.122)		-0.008
× Preannouncement			-0.010 (0.114)		-0.000
Academic degree	0.813*** (0.051)	0.042***	0.840*** (0.094)		0.043***
× Email			-0.163 (0.132)		-0.008
× Preannouncement			0.022 (0.125)		0.001
Overall employment periods, past 10 years (in years)	0.005 (0.011)	0.000	0.025 (0.020)		0.001
× Email			-0.054 (0.028)		-0.003
× Preannouncement			-0.011 (0.027)		-0.001
Any unemployment benefit receipt, past 10 years (ref. none)	0.023 (0.052)	0.001	-0.003 (0.097)		-0.000
× Email			-0.004 (0.135)		-0.000
× Preannouncement			0.082 (0.128)		0.004
Any welfare benefit receipt, past 10 years (ref.: none)	0.098 (0.059)	0.005	-0.005 (0.113)		-0.000
× Email			0.162 (0.154)		0.008
× Preannouncement			0.135 (0.146)		0.007
Number of persons	34,241	34,241	34,241		34,241

Note. All estimates are based on a model specification that considers binary variables for the recruitment waves besides the variables listed in the table. Continuous variables (e.g., age) are mean-centered. The standard errors are depicted in parenthesis. Marginal effects are calculated at the grand mean level of all continuous variables and the grand share of dummy coded variables.

* $p < .05$. ** $p < .01$. *** $p < .001$.

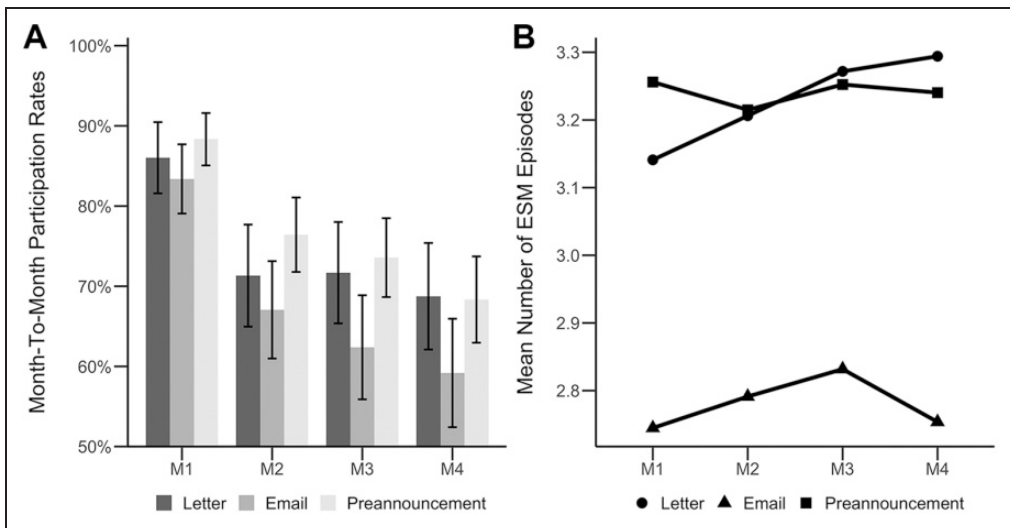


Figure 3. Panel A: Percentage of GJSP participants fulfilling all inclusion criteria and answering at least one item in the respective wave (M1–M4). Error bands correspond to the 95% confidence intervals. Panel B: Average (mean) number of submitted experience sampling method (ESM) episodes of subjects who answered at least one item in the respective GJSP waves (M1–M4).

regard to participation in the ESM among all participants who answered at least one item in the respective GJSP waves. ESM participation rates are, however, significantly greater in M4 for subjects in the letter (86.6%) and preannouncement (87.4%) conditions compared to those recruited via the email (75.9%) condition (letter vs. email [$b = 0.72, p = .007$], preannouncement vs. email [$b = 0.80, p = .001$]). Further, ESM participation rates are greater in M1 for subjects recruited in the preannouncement (85.2%) condition compared to subjects recruited in the email (77.6%) condition ($b = 0.50, p = .013$). The remaining contrasts are not statistically significant. These reported significant differences remain after controlling for sample selection through the inclusion of the survey variables. This result indicates that observed differences in terms of sample composition in the different experimental groups cannot explain the differences in regard to ESM participation. Still, these analyses reveal that subjects with an academic degree are more likely to participate in the ESM module. Detailed results are depicted in Table A4 in the Online Supplement.

Discussion

The present study used experimental data to examine whether the contact mode affects sign-up rates, nonresponse bias, and subsequent survey participation in a large app-based smartphone survey. Subjects were randomly assigned to be invited via email, postal letter, or preannouncement letter followed by email. The preannouncement approach yielded a roughly 1.4 times higher sign-up rate compared to the email and letter invitations (7.25% vs. ~5% for email and letter). Further, the average absolute nonresponse bias was rather low across all recruitment methods (< 4%) and lowest for the letter and preannouncement conditions relative to the email condition. Moreover, the three contact methods seem to be differentially appealing to certain individuals. For example, unlike emails and postal letters, the preannouncement approach does not yield different participation rates for men and women.

The contact mode also affected subsequent panel participation after the sign-up. Specifically, subjects in the email condition were found to have lower participation rates throughout the first four

waves of the GJSP compared to individuals in the letter and preannouncement conditions. These differences remained even after controlling for observed differences (i.e., demographic and psychological variables) in the sample composition of each recruitment mode.

Moreover, the experiment provided evidence that the contact mode has an effect on the data quality of momentary assessments using the ESM. Throughout the panel, subjects in the email condition were less likely to participate in the ESM module compared to the other two conditions. These statistical differences persisted after accounting for differential sample composition.

Implications

Our study extends the literature on app-based surveys in four important directions. First, it provides strong causal evidence that sending out a preannouncement letter in advance of an email invitation is the preferred method to increase sign-up rates for app surveys relative to stand-alone letter or email invitations. This finding is consistent with the literature on the use of preannouncement or advance letters to boost response rates to web surveys (Crawford et al., 2004; Dykema et al., 2011; Harmon et al., 2005; Kaplowitz et al., 2004; Porter & Whitcomb, 2007).

Second, for the first time, participation rates in an app-based population-based study were investigated using a sample of newly contacted subjects making the findings especially relevant for researchers who cannot recruit subjects from an ongoing study. Further, in the present study, we had the unique opportunity to contact the whole target population (i.e., job seekers from companies conducting mass layoffs), thus eliminating all confounding influences during the identification of potential subjects. Both of these novelties increase the value of the presented findings and account for the observed lower initial sign-up rates compared to other app-based surveys that draw their subjects from ongoing panel studies (e.g., Jäckle et al., 2019; Kreuter et al., 2019; McGeeney & Weisel, 2015; Scherpenzeel, 2017; Wells et al., 2013).

Third, the availability of rich administrative records allowed detailed nonresponse bias analyses. Despite the comparably low sign-up rates, the nonresponse bias was rather small, on average. This finding is in line with existing studies on the relationship between nonresponse rates and nonresponse bias (Groves, 2006; Groves & Peytcheva, 2008) and underlines that app surveys can be an effective tool to survey a representative sample of the target population.

Fourth, this study presents the first evidence that the recruitment mode used for app-based surveys has a long-lasting effect on panel participation as well as participation in the ESM measurements. The present study suggests that properties of the contact modes themselves are likely to be the main drivers of panel participation and ESM participation rather than selective participation. A plausible post hoc explanation for these findings is that subjects in the letter condition consider the study invitation more thoroughly and only sign up when they are truly motivated to continuously participate in the survey. Contrarily, the email invitation might make the registration process so easy that even slightly interested individuals decide to spontaneously sign up for the study without grasping its full scope, which in turn leads to poorer long-term participation rates, especially in more demanding questionnaires like the ESM.

Recommendations for Practice

Based on the present experiment, we derive the following recommendations for recruiting participants for app-based smartphone surveys:

1. *Do not use an email-only strategy to recruit subjects for panel surveys.* While the email invitation yielded similar sign-up rates compared to the letter invitation, it turned out to be disadvantageous in the longer term as it tended to bring in participants with lower rates of

panel participation and ESM participation. Thus, unless the survey is a one-off event, stand-alone emails should not be used to invite people to app-based surveys if letters or ideally preannouncement letters are feasible. Inviting subjects using a stand-alone letter yields comparable data quality to the preannouncement condition but fewer overall participants. Thus, the cost per respondent is lower in the preannouncement condition compared to the letter condition (assuming emails cost nothing).

2. *Combine different contact modalities.* We recommend using at least one nondigital channel (e.g., postal letter) to initiate first contact with potential subjects to increase the perceived trustworthiness of the survey. The first contact need not be a prenotification but could be the initial invitation containing all of the sign-up instructions. It remains a direction for future research whether the nondigital contact mode needs to be a letter, a postcard, or phone call. To facilitate the sign-up process, we suggest at least one email contact after the nondigital channel has been used. This email should include a direct download link to the survey app alongside the registration credentials.

Our results indicate that subjects who receive their sign-up credentials via email (i.e., email-only and preannouncement conditions) sign up faster compared to subjects recruited via letter. This finding suggests that an email contact makes the sign-up process particularly convenient, as subjects can directly click on the provided link to download the survey app. Contrarily, subjects recruited via letter have to set the letter aside in order to download the survey app on their smartphone. This delay, however, could act as a benefit of using a nondigital contact preceding an email contact, as it gives subjects more time to contemplate whether they actually want to commit to the study.

Suggestions for Future Research

A task for future studies is to better understand the mechanisms that make a contact strategy effective in app-based surveys. Specifically, it is important to investigate whether the effect of the preannouncement condition is simply driven by the increased exposure to the recruitment message (*exposure effect*) or the different modalities used (*multimodality effect*). To do so, one might use an email preannouncement instead of the paper letter as a further treatment.

To make the planning of app-based surveys easier and more effective, it is further important to extend the current findings to other populations, content domains, and survey structures. For well-being research, special focus should be put toward understanding selective participation in ESM questionnaires. Existing ESM studies indicate that individual differences in dispositional features (e.g., education, gender, psychological functioning) predict ESM nonresponse (Courvoisier et al., 2012; Messiah et al., 2011; Silvia et al., 2013). These studies however are based on relatively short time periods (i.e., usually 1 week). It remains a direction for future research to investigate which individual differences explain participation in ESM modules in panel surveys.

Authors' Note

We thank Benjamin Küfner for valuable research assistance. Furthermore, we are grateful for comments by seminar participants at ZEW Mannheim.

Data Availability

The paradata of the GJSP used in the analyses as well as the supplementary materials are available online at <https://osf.io/6ghsf/>. The administrative records cannot be made publicly available, but the files can be

requested from the Research Data Center (FDZ) of the Institute for Employment Research (IAB). The following link provides further information about IAB data access: https://fdz.iab.de/en/FDZ_Data_Access.aspx

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study has been supported by Deutsche Forschungsgemeinschaft (Grant ID: EI 379/11-1, SCHO 1270/5-1, and STE 1424/4-1).

Supplemental Material

Supplemental material for this article is available online.

Note

1. Here and in the forthcoming logistic regression models of sign-up, we do not include all available Integrated Employment Biographies variables as were used when assessing the balance of the experimental conditions (see Supplemental Table A1), as otherwise some variables would overlap. This is the reason why, for instance, part-time job, episodes job seeking/participation in policy schemes in the past 10 years are not considered.

References

- Antoni, M., Ganzer, A., & vom Berge, P. (2016). Sample of integrated labour market biographies (SIAB) 1975-2014. *FDZ-Datenreport, 4*. http://doku.iab.de/fdz/reporte/2016/DR_04-16_EN.pdf
- Birnholtz, J. P., Horn, D. B., Finholt, T. A., & Bae, S. J. (2004). The effects of cash, electronic, and paper gift certificates as respondent incentives for a web-based survey of technologically sophisticated respondents. *Social Science Computer Review, 22*(3), 355–362. <https://doi.org/10.1177/0894439304263147>
- Bryson, A., & MacKerron, G. (2017). Are you happy while you work? *Economic Journal, 127*(599), 106–125. <https://doi.org/10.1111/eoj.12269>
- Caliendo, M., Mahlstedt, R., & Mitnik, O. A. (2017). Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies. *Labour Economics, 46*(January 2016), 14–25. <https://doi.org/10.1016/j.labeco.2017.02.001>
- Cheng, A., Zamarro, G., & Orriens, B. (2018). Personality as a predictor of unit nonresponse in an internet panel. *Sociological Methods and Research*. <https://doi.org/10.1177/0049124117747305>
- Courvoisier, D. S., Eid, M., & Lischetzke, T. (2012). Compliance to a cell phone-based ecological momentary assessment study: The effect of time and personality characteristics. *Psychological Assessment, 24*(3), 713–720. <https://doi.org/10.1037/a0026733>
- Crawford, S. D., McCabe, S. E., Saltz, B., Boyd, C. J., Freisthler, B., & Paschall, M. J. (2004). *Gaining respondent cooperation in college web-based alcohol surveys: Findings from experiments at two universities* [Paper Presentation]. Annual Meeting of the American Association for Public Opinion Research, Phoenix, Arizona, USA.
- Deutsche Post DHL Group. (2018). *Annual report 2018: Acting agile*. <https://www.dpdhl.com/content/dam/dpdhl/en/media-center/investors/documents/annual-reports/DPDHL-2018-Annual-Report.pdf>
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley.
- Dykema, J., Stevenson, J., Day, B., Sellers, S. L., & Bonham, V. L. (2011). Effects of incentives and prenotification on response rates and costs in a national web survey of physicians. *Evaluation and the Health Professions, 34*(4), 434–447. <https://doi.org/10.1177/0163278711406113>

- Eckman, S., & Haas, G. C. (2017). Does granting linkage consent in the beginning of the questionnaire affect data quality? *Journal of Survey Statistics and Methodology*, 5(4), 535–551. <https://doi.org/10.1093/jssam/smx016>
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646–675. <https://doi.org/10.1093/poq/nfl033>
- Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2), 167–189. <https://doi.org/10.1093/poq/nfn011>
- Harmon, M. A., Westin, E. C., & Levin, K. Y. (2005). *Does type of pre-notification affect web survey response rates?* [Paper Presentation]. Annual Conference of the American Association for Public Opinion Research, Miami Beach, Florida, USA.
- Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Sage Publications.
- Hetschko, C., Eid, M., Lawes, M., Schöb, R., & Stephan, G. (2020). *The German job search panel*. OSF Preprints. <https://doi.org/10.31219/osf.io/7jazr>
- Jäckle, A., Burton, J., Couper, M. P., & Lessof, C. (2019). Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: Coverage and participation rates and biases. *Survey Research Methods*, 13(1), 23–44. <https://doi.org/10.18148/srm/2019.v13i1.7297>
- Kahneman, D., & Krueger, A. B. (2006). Developments in measurements of subjective well-being. *The Journal of Economic Perspectives*, 20(1), 3–24. <https://doi.org/10.1257/089533006776526030>
- Kaplowitz, M. D., Hadlock, T. D., & Levine, R. (2004). A comparison of web and mail survey response rates. *Public Opinion Quarterly*, 68(1), 94–101. <https://doi.org/10.1093/poq/nfh006>
- Kaplowitz, M. D., Lupi, F., Couper, M. P., & Thorp, L. (2012). The effect of invitation design on web survey response rates. *Social Science Computer Review*, 30(3), 339–349. <https://doi.org/10.1177/0894439311419084>
- Keusch, F., Struminskaya, B., Antoun, C., Couper, M. P., & Kreuter, F. (2019). Willingness to participate in passive mobile data collection. *Public Opinion Quarterly*, 83(S1), 210–235. <https://doi.org/10.1093/poq/nfz007>
- Kreuter, F., Haas, G. C., Keusch, F., Bähr, S., & Trappmann, M. (2019). Collecting survey and smartphone sensor data with an app: Opportunities and challenges around privacy and informed consent. *Social Science Computer Review*, 1–17. <https://doi.org/10.1177/0894439318816389>
- Ludwigs, K., & Erdtmann, S. (2019). The Happiness Analyzer – Developing a New Technique for Measuring Subjective Well-Being. *International Journal of Community Well-Being*, 1(2), 101–114. <https://doi.org/10.1007/s42413-018-0011-3>
- Lugtig, P. (2014). Panel attrition: Separating tayers, fast attriters, gradual attriters, and lurkers. *Sociological Methods and Research*, 43(4), 699–723. <https://doi.org/10.1177/0049124113520305>
- McGeeney, K., & Weisel, R. (2015). App vs. web for surveys of smartphone users: Experimenting with mobile apps for signal-contingent experience sampling method surveys. *Pew Research Center Report*. <https://www.pewresearch.org/methods/2015/04/01/app-vs-web-for-surveys-of-smartphone-users/>
- Menegaki, A. N., Olsen, S. B., & Tzagarakis, K. P. (2016). Towards a common standard—A reporting checklist for web-based stated preference valuation surveys and a critique for mode surveys. *Journal of Choice Modelling*, 18, 18–50. <https://doi.org/10.1016/J.JOCM.2016.04.005>
- Messiah, A., Grondin, O., & Encrenaz, G. (2011). Factors associated with missing data in an experience sampling investigation of substance use determinants. *Drug and Alcohol Dependence*, 114(2–3), 153–158. <https://doi.org/10.1016/j.drugalcdep.2010.09.016>
- Millar, M. M., & Dillman, D. A. (2011). Improving response to web and mixed-mode surveys. *Public Opinion Quarterly*, 75(2), 249–269. <https://doi.org/10.1093/poq/nfr003>
- Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, 7(3), 221–237. <https://doi.org/10.1177/1745691612441215>
- Organization for Economic Cooperation and Development [OECD]. (2013). *OECD guidelines on measuring subjective well-being*. <https://doi.org/10.1787/9789264191655-en>

- Pew Research Center. (2019). *Mobile fact sheet*. <https://www.pewresearch.org/internet/fact-sheet/mobile/>
- Porter, S. R., & Whitcomb, M. E. (2007). Mixed-mode contacts in web surveys: Paper is not necessarily better. *Public Opinion Quarterly*, 71(4), 635–648. <https://doi.org/10.1093/poq/nfm038>
- Read, B. (2019). Respondent burden in a mobile app: Evidence from a shopping receipt scanning study. *Survey Research Methods*, 13(1), 45–71. <https://doi.org/10.18148/srm/2019.v13i1.7379>
- Sakshaug, J. W., Vicari, B., & Couper, M. P. (2019). Paper, e-mail, or both? Effects of contact mode on participation in a web survey of establishments. *Social Science Computer Review*, 37(6), 750–765. <https://doi.org/10.1177/0894439318805160>
- Scherpenzeel, A. (2017). Mixing online panel data collection with innovative methods. In S. Eifler & F. Faulbaum (Eds.), *Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung* [Methodological problems of mixed-mode approaches in survey research] (pp. 27–49). Springer.
- Silvia, P. J., Kwapil, T. R., Eddington, K. M., & Brown, L. H. (2013). Missed beeps and missing data: Dispositional and situational predictors of nonresponse in experience sampling research. *Social Science Computer Review*, 31(4), 471–481. <https://doi.org/10.1177/0894439313479902>
- Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (n.d.). *The multidimensional mood state questionnaire (MDMQ) (English version of MDBF)*. Retrieved January 16, 2020, from <https://www.metheval.uni-jena.de/mdbf.php>
- Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Der mehrdimensionale Befindlichkeitsfragebogen* [The multidimensional mood state questionnaire] (MDBF). Hogrefe Verlag.
- Wells, T., Bailey, J. T., & Link, M. W. (2013). Filling the void: Gaining a better understanding of tablet-based surveys. *Survey Practice*, 6(1), 1–9. <https://doi.org/10.29115/SP-2013-0002>
- West, B. T., Conrad, F. G., Kreuter, F., & Mittereder, F. (2018). Nonresponse and measurement error variance among interviewers in standardized and conversational interviewing. *Journal of Survey Statistics and Methodology*, 6(3), 335–359. <https://doi.org/10.1093/JSSAM/SMX029>

Author Biographies

Mario Lawes is a doctoral candidate at the Freie Universität Berlin and a pre-doctoral fellow of the International Max Planck Research School on the Life Course (LIFE). Email: mario.lawes@fu-berlin.de

Clemens Hetschko is an associate professor in economics at the University of Leeds, United Kingdom, a CESifo Affiliate, and a visiting researcher at Freie Universität Berlin. Email: c.hetschko@leeds.ac.uk

Joseph W. Sakshaug, PhD, is a researcher at the Institute for Employment Research (IAB, Nuremberg, Germany) and a professor at the Ludwig-Maximilians-Universität Munich and the University of Mannheim. Email: joe.sakshaug@iab.de

Stephan Griebemer is a researcher at the "Data Distribution" unit of the Institute for Employment Research (IAB, Nuremberg, Germany). Email: stephan.griessemer@iab.de