Survey Research Methods (2025)
Vol. 19, No. 4, pp. 431-452
https://doi.org/10.18148/srm/2025.v19i4.8553
European Survey Research Association

Effects of Replacing Telephone with Web, Mail, and Mixed-Mode Data Collection in an Establishment Follow-Up Survey

Benjamin Küfner[1] · Joseph W. Sakshaug[1,2,3] · Stefan Zins[1] · Claudia Globisch[1]
[1]Institute for Employment Research (IAB)
[2]LMU-Munich
[3]University of Mannheim

Due to rising data collection costs, there is growing pressure to move away from traditional interviewer-administered mode designs in favor of fully self-administered mode designs in ongoing panel surveys, including large-scale establishment panels. However, the consequences of moving to a fully self-administered mode design on follow-up and cumulative participation in ongoing establishment panel surveys are largely unknown. To address this research gap, we report the results of a follow-up mode design experiment conducted in the second wave of the 2020 IAB-Job Vacancy Survey, an ongoing panel study in Germany. The experiment builds on a previously reported mode design experiment conducted in the first wave survey, where establishments were randomized to four self-administered mode designs (concurrent mail-web mixed-mode, sequential web-to-mail mixed-mode, single-mode web, and single-mode mail). In the second wave (i.e., follow-up) survey, reported here, respondents from the first wave were further randomly allocated to 1) a continuation of the same self-administered mode design from the first wave, or 2) a single-mode telephone design. The results show that the continuation of self-administration leads to higher response rates (both follow-up and cumulative) for the single-mode mail and concurrent mail-web mixed mode designs and comparable response rates for the single-mode web and sequential web-to-mail mixed-mode designs, compared to the telephone follow-up design. Using extensive administrative data, we do not find evidence that forgoing telephone follow-ups adversely affects nonresponse bias or subgroup participation compared to continuing with self-administration in the follow-up wave. Potential cost savings (of up to 67%) were evidenced when replacing the telephone mode with a self-administered follow-up mode design. In-depth qualitative interviews revealed that establishments prefer a constant mode design across waves due to the familiarity and routine of the response process.

Keywords: nonresponse; data collection; self-administered modes; CATI

Establishment panel surveys are a vital tool for measuring the economic activity of businesses and organizations over time. By repeatedly collecting data from the same units, they allow for the measurement of dynamic changes and trends within and across business sectors. However, the quality of panel survey data are threatened by attrition, decreasing response rates, and rising data collection costs (e.g., König et al. 2021; Küfner et al. 2022a; U.S. Bureau of Labor Statistics 2023a), which puts pressure on survey agencies to develop more efficient data collection strategies, including greater use of self-administered modes (e.g., web). Most establishment surveys around the

Supplementary Information The online version of this article (https://doi.org/10.18148/srm/2025.v19i4.8553) contains supplementary material, which is available to authorized users.

Corresponding author: Joseph W. Sakshaug, Institute for Employment Research (IAB), Nuremberg, Germany (Email: joe.sakshaug@iab.de)

world use self-administered mode designs as opposed to more traditional interviewer-administered modes, such as computer-assisted telephone interviewing (CATI) or computer-assisted personal interviewing (CAPI) (e.g., Nicholls II 2000; Haraldsen 2013; Scholtus 2014; U.S. Bureau of Labor Statistics 2023a). However, interviewer-administered modes are still used in many voluntary establishment panel surveys, as interviewers are well-suited to motivate respondents to continue participating in follow-up interviews (Haraldsen 2013). For example, the voluntary Job Openings and Labor Turnover Survey in the U.S., which is designed as a panel study, uses CATI for approximately the first five waves of data collection to establish a reporting routine before establishments can decide whether they want to participate via web, e-mail, or continue with CATI (U.S. Bureau of Labor Statistics 2023b). The voluntary Establishment Panel run by the Institute for Employment Research (IAB) in Germany makes extensive use of both CAPI and CATI modes for nonresponse follow-ups after an initial invitation to web, resulting in more than half of all responses collected by interviewers (Bächmann et al. 2023). The voluntary IAB-Job Vacancy Survey (IAB-JVS) in Germany is another ongoing panel study, which annually recruits a new cohort using a concurrent mail-web mixed-mode design and conducts three CATI follow-up interviews in the first, second, and third quarters of the following year (Bossler et al. 2020).

In the case of the IAB-JVS, declining response rates and rising data collection costs recently prompted the survey team to experiment with four alternative self-administered mode designs (concurrent mail-web mixed-mode, sequential web-to-mail mixed-mode, single-mode web, single-mode mail) in the first panel wave and replacing the CATI mode with self-administration in the follow-up waves, with an eye towards making the panel fully self-administered in the future. The results of the experiments conducted in the first panel wave of the 2020 survey, reported in (Küfner 2024), showed that neither response rates nor nonresponse bias substantially differed between the four self-administered mode designs. Further, the results indicated a large potential for cost-savings (up to 50% per respondent) when utilizing a sequential web-to-mail or a single-mode web design compared to the concurrent mail-web mixed-mode design that is currently used in the IAB-JVS. The motivation to build on this prior work and experiment with replacing the CATI mode with self-administration in the second panel wave is based on addressing the current decline in CATI response rates (Küfner et al. 2022b), reducing costs, and minimizing the risk of longitudinal measurement mode effects (Cernat et al. 2021). Additionally, maintaining self-administration across all panel waves has the potential to streamline survey organizational processes and allow more flexibility in the development of the survey design. However, forgoing telephone follow-ups may also come with potential risks. For instance, the absence of interviewers could actually have a negative impact on response rates in the follow-up waves. In addition, the risk of nonresponse bias may increase if self-administered follow-ups are less effective at motivating certain types of establishments to continue participating in the panel compared to interviewer-administered follow-ups. With the COVID-19 pandemic reshaping business practices and fostering an era of remote work and reliance on online tools, there's a growing need to identify the most effective modes (or mode designs) for maximizing establishment panel survey participation in this new business environment.

The empirical literature offers limited guidance in evaluating the potential benefits and risks associated with replacing interviewer-administered follow-ups with self-administered ones in establishment panel surveys. Furthermore, the literature mainly focuses on response rates, leaving the effects on nonresponse bias and costs unquantified (for an exception, see Gleiser et al. 2022). To address this research gap, we present results from the aforementioned follow-up mode design experiment conducted in the second wave of the 2020 IAB-JVS panel survey, building on the first wave experiment. Specifically, we compare the effects of using CATI follow-ups versus following up with the same self-administered mode design used in the first wave. In addition to evaluating both follow-up and cumulative response rates over both waves, we also examine nonresponse bias and predictors of follow-up and cumulative survey participation by exploiting a large-scale administrative database containing detailed information on establishment and aggregate employee characteristics. Further, we present results from a cost analysis to quantify the cost implications of the different mode design combinations. Lastly, post-survey qualitative interviews were conducted with establishments to acquire insights into their perceptions of CATI and self-administered modes and their preferences for continuing with the same self-administered mode design or switching to an interviewer-led mode design from one panel wave to the next.

The remainder of this article is structured as follows. Section 1 provides an overview of self- and interviewer-administered modes in establishment surveys, summarizes the literature on mode design changes between panel waves, and presents the research questions to be addressed. The experimental design, the data sources used, and the analytic strategy are described in Sect. 2. Sect. 3 presents the results of the mode design experiments and reports insights from both short structured and in-depth qualitative interviews with establishments. Lastly, Sect. 4 provides a summary of the key findings and implications for survey practice.

1 Background

1.1 Self-Administered and Interviewer-Administered Mode Designs in Establishment Surveys

Mail and web modes (incl. electronic data interchange) are the primary modes of data collection used in establishment surveys with other self-administered modes (e.g., fax, touchtone data entry) and interviewer-administered modes (e.g., CAPI, CATI) used to a lesser extent (e.g., Buiten et al. 2018; Robertson 2012; Haraldsen 2013; Scholtus 2014). All of these modes can be deployed as part of a single-mode design, in which all establishments are offered only one mode of data collection, or a mixed-mode design, in which multiple modes are deployed to collect data from establishments. A further distinction can be made between concurrent mixed-mode designs, in which multiple modes are offered simultaneously and the establishment can choose which one they use, and sequential mixed-mode designs, in which data collection starts with one mode in the first phase of fieldwork and additional modes are deployed to remaining nonrespondents in the subsequent fieldwork phases. Mixed-mode designs can comprise a mix of either self-administered or interviewer-administered modes, or a combination of both mode types. For example, a common sequential mixed-mode design is to use a self-administered starting mode, such as web, and follow-up remaining nonrespondents with an interviewer-administered mode, such as CATI (De Leeuw 2005; De Leeuw et al. 2018). In addition to mixing modes within one wave of a survey, one can also mix modes or mode designs across multiple waves of a panel survey. For instance, a survey might use a face-to-face design in the first wave and introduce a web design in subsequent waves. The present study focuses on the impact of mixing modes between waves by examining the effect of switching from a self-administered mode design in the first wave to a CATI mode design in the follow-up wave, compared to using a constant self-administered mode design in both waves.

The popularity of self-administered modes is driven by at least four factors. First, establishment surveys collect extensive data on quantities and facts that respondents may not be able to recall on the spot in the presence of an interviewer. To provide these data, establishments often must check their records, balance sheets, or internal documentation systems. With self-administered modes, respondents can take their time to look for this information and verify the correct response. Second, self-administered questionnaires are easier to share among colleagues within an establishment who could contribute to answering the questions (e.g., Haraldsen 2013; Scholtus 2014; Haraldsen et al. 2023; Küfner 2024). Third, self-administered modes, especially web and mail, facilitate the use of matrix and grid tables (Moore 2016), which are often used in establishment surveys. Administering a two-dimensional matrix in interviewer-administered modes, especially CATI, could be very demanding for the interviewer and respondent (Haraldsen 2013), especially when the items involve disaggregated numbers and arithmetic sums. And fourth, self-administered modes are associated with lower survey costs compared to interviewer-administered modes (e.g., Rosen 1997; Haraldsen 2013; Scholtus 2014; Moore 2016). Further, web questionnaires offer additional features and capabilities, including automated filtering, prompts on item nonresponse, and preloading of available data, which may improve data quality and facilitate the response process (e.g., Couper 2008; Moore and Wojan 2016).

Although interviewer-administered modes are less common in establishment surveys, they possess certain advantages. Most dominantly, interviewer-administered modes enable a two-way communication that allows respondents to ask questions to clarify vague or ambiguous questions. Moreover, interviewers can motivate establishments to take part in the survey and thus influence the decision to participate, which could ultimately lead to higher response rates (e.g., Haraldsen 2013; Scholtus et al. 2014; Haraldsen 2023). As a consequence, CATI and CAPI modes are primarily employed for specific types of surveys and contact screening. This is evident in the European Company Survey (Ipsos 2020), where direct conversation is used to identify the most knowledgeable respondent within a company and assess the eligibility of the company. In establishment panel surveys, establishments face a high burden of participation due to repeated survey requests. The likelihood of panel participation in such cases could benefit from a trust-based relationship between the establishment and the interviewer, who actively encourages and motivates the establishment to continue participating in subsequent survey waves. Additionally, interviewers are extensively used for nonresponse follow-ups. For example, the Australian Bureau of Statistics (2023) regularly deploys interviewers, who play a key role in persuading reluctant establishments to participate. The main disadvantages of interviewer-administered modes are that they are more expensive than self-administered survey modes, less timely, and less suitable for high frequency data collection (e.g., Rosen 1997; Haraldsen 2013; Scholtus 2014; Moore 2016).

1.2 Transitioning from Interviewer-Administration to Self-Administration in Establishment Panel Surveys

The AAPOR Task Force Report by (Olson 2021) offers a comprehensive overview of surveys shifting from CATI to self-administered data collection. Although the report ac-

knowledges the dearth of evidence specific to establishment surveys (both cross-sectional and panel), it outlines two important motivations relevant to the implementation of self-administered modes in establishment surveys: lower survey costs and increased data quality. As previously outlined, CATI is known to entail higher costs than self-administered modes. Therefore, it is common to transition from CATI to self-administered modes in order to reduce expenses and increase cost efficiencies. With respect to data quality, the aim is to maintain (or increase) response rates that have been steadily declining in CATI surveys over time (Olson 2021; Küfner et al. 2022b) without exacerbating nonresponse bias. Specific to establishment surveys, meta-analyses are lacking which examine whether the general trend of decreasing response rates (e.g., König et al. 2021; Küfner et al. 2022a; U.S. Bureau of Labor Statistics 2023a) is more pronounced in self-administered or interviewer-administered modes.

For panel surveys, changing the mode design between waves comes with certain risks. As it is key for panel surveys to minimize panel attrition over the waves, changing the mode design could backfire if panel respondents prefer the previous mode design over the new one (see for example (Jäckle et al. 2015). Thus, altering the mode design between waves of a panel study could lead to an increased risk of nonresponse or attrition and associated selection biases in subsequent waves of data collection.

1.3 Empirical Evidence on Replacing Interviewer-Administered with Self-Administered Modes in Establishment Surveys

Limited research exists comparing interviewer and self-administered modes of participation in establishment surveys, both in a cross-sectional (DesRoches et al. 2007; Moore 2016; Zuckerbraun 2013; Gleiser et al. 2022) and a panel setting (Ellguth et al. 2014). For instance, (Moore 2016) run an experimental pilot in the cross-sectional U.S. Rural Establishment Innovation Survey in 2014 comparing five sequential mixed-mode designs with varying order of mail, web, and CATI modes (two mail-first, one web-first, and two CATI-first), additionally varying incentives (no incentive, 2$, 4$), e-mail reminders (yes/no), the use of priority mail (none, once, twice), and the use of refusal aversion questionnaires (yes/no). The two mail-first, the web-first, and the CATI-first designs using incentives had similar response rates (about 30%), but the CATI-first design without incentives and priority mailing had a significantly lower response rate of 19%. Surprisingly, both CATI-first designs had significantly lower shares of CATI responses (up to 17%) compared to the web (up to 58%) and mail (up to 35%) follow-up responses. This implies that even when offering CATI as the first mode of data collection, most respondents seem to prefer mail and web as follow-up modes.

Within the 2018 cross-sectional refreshment sample of the voluntary German IAB-Establishment Panel, (Gleiser et al. 2022) conducted an experimental comparison between a face-to-face mode design and a sequential web-to-face-to-face mixed-mode design. Both designs included a "mop-up" phase with self-administered paper questionnaires offered to all remaining nonrespondents. The results show similar response rates (22% and 21%, respectively) and (aggregate) levels of nonresponse bias between both mode designs, and about 14% cost savings per respondent in the web-first design, which yielded fewer face-to-face contact attempts (Gleiser et al. 2022).

(Ellguth et al. 2014) investigated the possibility of switching the mode design of the IAB-Establishment Panel from face-to-face paper & pencil interviews with the option of paper self-completion to computer-assisted personal interviews with the option of self-completion of web questionnaires by asking panellists about their ability and willingness to participate online. While 95% of panel respondents indicated that they have internet access and are able to participate online, only 38% expressed a willingness to do so. The main barrier cited was security concerns about data transfer.

Additional research has found limited effects of using telephone calls to prompt establishments that broke off a web survey to complete the web questionnaire compared to not using them within a cross-sectional establishment survey (Zuckerbraun 2013), and a lower likelihood of establishments without employees participating via web than CATI within the single wave of a panel survey (DesRoches et al. 2007).

1.4 Research Gaps and Research Questions

While several studies examine introducing self-administered modes or mixed-mode designs in follow-up waves of household panel surveys (e.g., Jäckle et al. 2015; Bianchi et al. 2017; Allum et al. 2018), such studies are limited in the establishment survey literature (e.g., Gleiser et al. 2022; Zuckerbraun 2013; Ellguth et al. 2014; DesRoches et al. 2007). The reviewed literature points to a lack of experimental evidence on the impact of changing or continuing the same mode design in the follow-up waves of establishment panel surveys, as well as a limited understanding of the effects of various mode design sequences over multiple panel waves, particularly when starting with a self-administered mode design in the first wave. Moreover, there is an absence of evidence on the effects of various mode design combinations in panel surveys with respect to nonresponse bias, subgroup participation, and costs. Qualitative data on

establishments' perceptions of continuing or changing the mode design from self-administration to interviewer-administration in follow-up waves and its impact on their response processes is also missing from the literature.

A better understanding of these issues would inform survey organizations and broaden their consideration of alternative mode designs in establishment panel surveys. This includes providing reference estimates for impacts on response rates, nonresponse bias, and costs. Additionally, subgroup analyses would help identify those establishments that are more (or less) likely to participate in follow-up waves of a panel study under different mode designs. Finally, collecting qualitative insights from establishments on their mode design preferences and response processes in panel studies could help to inform survey designs that mitigate or remedy the disadvantages of the different mode designs.

Against this background, the present study pursues the following five research questions:

- RQ1: To what extent do follow-up and cumulative response rates differ between a continuation of a self-administered mode design and a switch to a CATI mode design in the follow-up wave of an establishment panel survey?
- RQ2: Do follow-up and cumulative nonresponse biases differ between a continuation of a self-administered mode design and a switch to a CATI mode design in the follow-up wave of an establishment panel survey?
- RQ3: Are some establishment subgroups more (or less) likely to participate in the follow-up wave conducted under a continuation of a self-administered mode design or a switch to a CATI mode design in the follow-up wave of an establishment panel survey? Do subgroup patterns of cumulative participation in both waves of the panel vary between the different mode design sequences?
- RQ4: To what extent does a continuation of a self-administered mode design affect both follow-up and cumulative survey costs compared to a switch to a CATI mode design in the follow-up wave of an establishment panel survey?
- RQ5: What do establishments perceive to be the advantages and disadvantages of web, mail, and CATI mode designs, and switching from one mode design to another between waves of a panel survey?

Although establishment surveys primarily collect factual data, which should be less prone to measurement mode effects, changing the mode design in an ongoing panel survey comes with the risk of confounding true change and measurement mode effects, which could compromise estimates of change and trends over time (Smith 2019; Cernat et al. 2021). While we acknowledge the importance of investigating such risks in panel surveys, this article focuses solely on participation and selection effects and we leave the topic of measurement mode effects to future research.

2 Data and Methods

2.1 Data

2.1.1 IAB Job Vacancy Survey

The IAB Job Vacancy Survey (IAB-JVS) is a voluntary establishment survey covering all establishments with at least one employee contributing to social security in Germany (Bossler et al. 2020). The IAB-JVS collects extensive data on job vacancies, recruiting processes, and job flows, such as hires and layoffs. It serves as the official data source for the number of job vacancies in Germany and these results are reported to Eurostat on behalf of Germany. It is designed as a four-wave panel survey with the first wave starting in the fourth quarter of each year followed by three quarterly follow-ups conducted in the following year. Since 2002, the IAB-JVS has used a concurrent mail-web mixed-mode design in which paper invitation letters are sent to a cross-sectional sample of establishments in the first panel wave along with a copy of the paper questionnaire and an optional web link to complete the survey online. Since 2006, the three follow-up panel waves have been conducted via single-mode CATI. The follow-up panel waves are designed to be very short (3–5 min, on average, depending on the quarter), collecting updated information about the quarter-specific number and structure of employees, vacancies, and the average number of hours per employee on working time accounts. The sample is drawn from the population of all establishments in Germany with at least one employee who is subject to social insurance contributions. Establishments that have indicated they do not wish to be contacted for surveys are excluded from the sampling frame. The IAB-JVS uses a stratified sample design with region, establishment size, and industry as stratification variables. All data used in this study are available from the Research Data Centre (RDC) of the Federal Employment Agency in Germany. Restrictions apply to the availability of these data, which are not publicly available. For more information on data access, see https://fdz.iab.de/en.aspx.

2.1.2 Experimental Design

The second panel wave experiment analyzed in this article builds on a preceding large-scale mode design experiment conducted in the first panel wave (i.e., the fourth quarter)

of the 2020 IAB-JVS (Q4/2020), in which four self-administered mode designs were compared: concurrent mail-web mixed-mode, sequential web-to-mail mixed-mode, single-mode web, and single-mode mail (Küfner 2024). The rationale for the mode experiment in the first wave was to examine the potential for cost savings with the single-mode web and sequential web-to-mail mixed-mode designs (i.e., the push-to-web designs), while also evaluating their impact on various aspects of survey participation compared to concurrent mail-web mixed-mode and single-mode mail designs. Because of the importance of the survey for official statistics and substantive research, the decision to adopt a new mode design—despite its known advantages (see Sects. 1.1 and 1.2)—was only to be made on the basis of empirical evidence from these experiments. The sample size in the first wave experiments was 132,433 establishments and the entire field period lasted from the 26th of September 2020 to the 6th of January 2021.

The first wave experiment showed no statistically significant difference between the experimental groups in terms of response rates and aggregate nonresponse bias. Larger establishments were more likely to participate via web than mail in the single-mode designs. In addition, the web-to-mail and single-mode web designs indicated the highest potential for cost-savings. See (Küfner 2024) for complete details of this experiment and outcomes.

Building on this previous experiment, a follow-up survey experiment was conducted in the second panel wave, i.e., the first quarter of 2021 (Q1/2021), which is illustrated in Fig. 1. Specifically, respondents within each of the four mode designs of the first panel wave (fourth quarter of 2020; Q4/2020) were randomly subdivided into two experimental mode design groups for the follow-up panel wave: 1) a mode design continuation group that continued the same self-administered mode design that was randomly assigned to establishments in the first wave, and 2) a mode design switch group that switched the mode design to single-mode CATI (i.e., the standard IAB-JVS follow-up mode design). This random split was carried out by applying systematic random sampling after sorting the sampling frame by establishment size and industry. Similarly to the first wave experiment, the rationale behind the follow-up wave experiment was to investigate potential cost savings and the implications of replacing interviewers with self-administration on continued participation in the survey.

A total of 15,890 establishments were fielded in the follow-up wave (Q1/2021) experiment. These establishments are respondents from the first wave sample, where a sample of 132,433 establishments was fielded. The field period of the follow-up wave started on the 9th of January 2021 and ended 31th of March 2021. All establishments received an advance letter several days before the follow-up wave, thanking them for participating in the first wave and informing them about the upcoming follow-up survey. As in the first panel wave, all self-administered mode designs in the follow-up panel wave included postal invitation letters and one post-due-date reminder letter[1]. Following completion of the first wave questionnaire, participants were asked to provide details of a designated contact person and their corresponding telephone number. This designated contact person served as the recipient for all mailed correspondence in the follow-up wave, with a parallel arrangement for directing telephone contacts in the CATI mode design. If the telephone number used in the CATI mode design turned out to be invalid, alternative telephone numbers were obtained through extensive Internet research or from data available at the German Federal Employment Agency (Bundesagentur für Arbeit). We note that due to administrative reasons, all mode design groups were fielded in two tranches, resulting in a 37 day longer field period for observations in the first tranche than for those in the second tranche. In the Online Supplement B.1 Sample Tranches, we provide more details on the reasons for using separate tranches and dates of invitations and reminders of both tranches as well as additional sensitivity checks, which show that our results are robust to the impact of the tranches.

2.1.3 Establishment History Panel

The Establishment History Panel (BHP) is an administrative data source, which we use to analyze nonresponse bias (RQ2) and predictors of survey participation (RQ3). The BHP contains detailed establishment and employee characteristics for all establishments in Germany with at least one employee (Ganzer et al. 2022). These characteristics are expected to be correlated with important variables collected in the IAB-JVS, including the number of job vacancies and the share of hired employees on fixed-term contracts. These correlations make the BHP well-suited for analyzing nonresponse bias in the IAB-JVS. By using a unique establishment identifier the IAB-JVS response indicator can be directly linked to the 2020 BHP, resulting in a dataset with detailed information for both respondents and nonrespondents. Exceptions are for establishments that ceased to exist between the reference date for the BHP generation (30th of June) and the IAB-JVS sampling (31th of December of the previous year). These non-linked cases account for 4% of the IAB-JVS sample. These establishments are excluded from the analysis of nonresponse bias (RQ2) and predictors of survey participation (RQ3), but are included

[1] This paper encompasses all experimental groups in which a post-due-date reminder was used in the first wave. Additional experimental groups in which a pre-due-date reminder was tested are not considered in this paper.

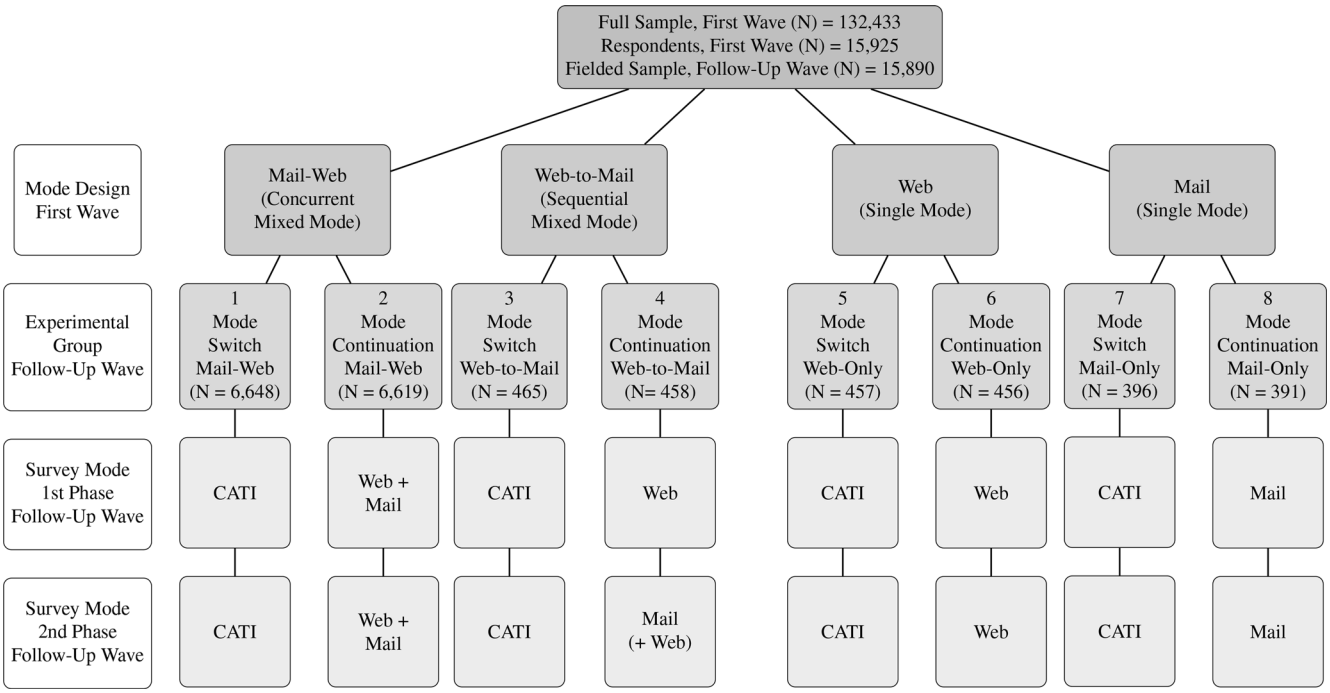


Fig. 1

Experimental Design

in the response rate (RQ1) and cost analyses (RQ4). Online Supplement A contains the full list of BHP variables used in the analysis and descriptive statistics of each for the fielded sample and each experimental group.

2.2 Empirical Strategy

2.2.1 Response Rates

For each wave of data collection, we define a completed interview (or response) for an establishment that answers two key survey items: the number of job vacancies and the number of employees. Both items are necessary to meet the reporting requirements of Eurostat. In addition, all self-administered interview data has to be actively submitted, either by mailing the paper questionnaire back to the survey institute or by clicking the submit button in the web survey. For CATI interviews, the same two items have to be answered and all questions have to be read out loud to be considered a complete interview.

Response rates are computed using the AAPOR Response Rate 1 definition (American Association for Public Opinion Research 2023). This response rate is seen as the minimum response rate as it is simply the ratio of the number of respondents to the fielded sample. Response rates are reported separately for the follow-up wave and cumulatively for both panel waves (Q4/2020 and Q1/2021). As shown in Eq. 1, the follow-up response rate relates the respondents of the follow-up wave to the fielded sample of this wave.

$$\text{Follow-Up Response Rate}_{\text{fup,exp}} = \frac{R_{\text{fup,exp}}}{N_{\text{fup,exp}}} \tag{1}$$

where $R_{fup,exp}$ denotes the number of respondents in the follow-up wave in the experimental group *exp* and $N_{fup,exp}$ is the number of establishments fielded in the follow-up wave *fup* in the respective experimental group *exp*.

To enable inference to the population, we use nonresponse weights to estimate the follow-up response rate. These nonresponse weights account for nonresponse in the first panel wave by estimating response propensities separately for each experimental group using the standard IAB-JVS auxiliary variables, namely, establishment size, region, industry, the wage of employees, and age of employees (see Brenzel et al. 2016). The inverse of the response propensities is then multiplied by the inverse of the inclusion probability, i.e., design weights. To account for the systematic random split of respondents into the mode design switch and continuation groups, we multiply the weights by two and add an adjustment factor to scale the weights to the population total.

Cumulative response rates are calculated by relating respondents who participated in both waves to the fielded sample of the first wave.

$$\text{Cum. Response Rate}_{\text{cum,exp}} = \frac{R_{\text{cum,exp}}}{N_{fir,\text{exp}}} \tag{2}$$

where $R_{cum,exp}$ denotes the number of respondents that participated in the first and follow-up wave in the experimental group *exp* and $N_{fir,exp}$ is the number of establishments fielded in the first panel wave *fir* in the respective experimental group *exp*.

For the estimation of cumulative response rates, we only apply design weights as the reference group is the full sample of the first panel wave. Hence, we don't consider nonresponse in the first wave. Similar to the nonresponse weights, we multiply the design weights by two to compensate for the systematic random split within the first wave experimental groups and scale the weights afterward to the population total. Stratification is accounted for in all standard error estimations.

2.2.2 Nonresponse Bias

To evaluate the effects of the mode designs and mode design combinations on nonresponse bias (RQ2), we compute follow-up and cumulative nonresponse bias estimates for various establishment and employee characteristics from the BHP. Nonresponse bias is defined as the difference between the estimate of interest for respondents and the corresponding estimate of interest for the fielded sample, which are proportions for the categorized BHP variables. Hence, nonresponse bias in the follow-up panel wave is defined as:

$$\begin{aligned}\widehat{\text{Follow-Up NR Bias}}_{\text{fup,exp},k} =& \hat{Y}_{\text{fup,exp},k,r,p} - \\ & \hat{Y}_{\text{fup,exp},k,n,p}\end{aligned} \tag{3}$$

where $\hat{Y}_{fup,exp,k,r,p}$ denotes the estimator for the k^{th} statistic of interest in the follow-up wave for the respective experimental group *exp* based on the respondents (r) using design and nonresponse weights (p) and estimator $\hat{Y}_{fup,exp,k,n,p}$ is analogously defined for the fielded sample with samples sizes n as shown in Fig. 1 (see Eqs. 1 and 2 in the Online Supplement C for definitions of $\hat{Y}_{fup,exp,k,r,p}$ and $\hat{Y}_{fup,exp,k,n,p}$, respectively).

Analogous to the response rate analysis, we apply different sets of weights for the nonresponse bias analysis. To isolate the effect of mode design on nonresponse bias in the follow-up panel wave, we use nonresponse weights to account for the selectivity of response in the first panel wave.

In order to estimate the cumulative nonresponse bias of the different mode design combinations and enable the decomposition of the impacts of the first and follow-up panel waves, we estimate the nonresponse bias of the first and follow-up panel waves separately and apply only design weights. The cumulative nonresponse bias is then calculated as the sum of the design-weighted nonresponse biases of the first and follow-up panel waves for each statistic of interest. This approach enables us to identify whether the mode designs of the follow-up wave offset or reinforce the nonresponse bias of the first panel wave. The cumulative nonresponse bias of the first and follow-up panel waves for the k^{th} statistic of interest is then estimated by:

$$\begin{aligned}\widehat{\text{Cum. NR Bias}}_{\text{cum,exp},k} =& (\hat{Y}_{fir,\text{exp},k,r,d} - \hat{Y}_{fir,\text{exp},k,n,d}) + \\ & (\hat{Y}_{\text{fup,exp},k,r,d} - \hat{Y}_{\text{fup,exp},k,n,d})\end{aligned} \tag{4}$$

where $\hat{Y}_{w,exp,k,r,d}$ denotes the estimator for the k^{th} statistic of interest in the first ($w = fir$) or follow-up panel ($w = fup$) wave for the respective experimental group *exp* based on the respondents (r) using design weights (d) and estimator $\hat{Y}_{w,exp,k,n,d}$ is analogously defined for the fielded sample of the first ($w = fir$) or follow-up panel ($w = fup$) wave (see Eqs. 3 and 4 for definitions of $\hat{Y}_{w,exp,k,r,d}$ and $\hat{Y}_{w,exp,k,n,d}$ in the Online Supplement C).

For summary comparisons, we also compute absolute nonresponse biases and average absolute nonresponse biases for the follow-up panel wave. Absolute nonresponse bias is defined as the absolute value of the nonresponse bias, as shown in the following equation:

$$\begin{aligned}\widehat{\text{Abs. Follow-Up NR Bias}}_{\text{fup,exp},k} = \\ \left|\widehat{\text{Follow-Up NR Bias}}_{\text{fup,exp},k}\right|\end{aligned} \tag{5}$$

Average absolute nonresponse bias is defined as the average nonresponse bias across all statistics of interest, as shown in the equation below:

$$\begin{aligned}\widehat{\text{Avg. Abs. Follow-Up NR Bias}}_{\text{fup,exp}} = \\ \frac{\sum_{k=1}^{K} \widehat{\text{Abs. Follow-Up NR Bias}}_{\text{fup,exp},k}}{K}\end{aligned} \tag{6}$$

with K being the total number of statistics of interest considered.

Analogously, we compute absolute and average absolute cumulative nonresponse bias estimates across both panel waves using the following formula:

$$\widehat{\text{Abs. Cum. NR Bias}}_{\text{cum,exp},k} = \left|\widehat{\text{Cum. NR Bias}}_{\text{cum,exp},k}\right| \tag{7}$$

$$\widehat{\text{Avg. Abs. Cum. NR Bias}}_{\text{cum,exp}} = \frac{\sum_{k=1}^{K} \widehat{\text{Abs. Cum. NR Bias}}_{\text{cum,exp},k}}{K} \quad (8)$$

with K being the total number of statistics of interest considered.

To obtain an aggregate measure of nonresponse bias across a range of individual variables, we compute the average absolute nonresponse bias separately for three groups of BHP administrative variables: establishment characteristics, employee characteristics, and all BHP variables (see Table A.2 of the Online Supplement for a comprehensive overview).

The concurrent mail-web mixed-mode groups have a disproportionately large sample size compared to other groups; hence, we apply a repeated downsampling approach to ensure comparable sample sizes to the other mode design groups (around 7500 in the first wave and 450 in the follow-up wave) are used for this analysis, thereby avoiding any potential sample size effects on the bias results. This approach also allows for the computation of confidence intervals for the average absolute bias estimates in this mode design group. For the other experimental groups, we estimate confidence intervals using standard errors estimated by a bootstrap method based on 500 replicates and normal-approximated quantiles.

2.2.3 Predictors of Survey Participation

To better understand the factors that influence survey participation in the follow-up panel wave, we fitted logistic regression models based on all establishments fielded in the follow-up wave separately for each mode design used in the first wave. This analysis sheds light on which establishments are more (or less) likely to participate in the follow-up wave under the mode design continuation or switch scenarios. Using establishment characteristics and their interactions with the experimental indicator (also included as main effects) as explanatory variables, we modelled survey participation (1 = response in the follow-up wave; 0 = nonresponse in the follow-up wave) and estimated predicted probabilities of response for different establishment characteristics with the covariates fixed at their global means. For evaluating differences in predicted probabilities between the mode design continuation and switch designs, the results of a Wald test are displayed (see Mize 2019). The following formula shows the logistic regression model used to predict the participation probabilities:

$$Pr(R_{\text{fup,exp},i} = 1) = \frac{1}{1 + e^{-(\alpha + \boldsymbol{\beta} \boldsymbol{x}_{\text{fup,exp},i}^{\top} + \boldsymbol{\zeta}[\boldsymbol{x}_{\text{fup,exp},i}\ \boldsymbol{m}_{\text{fup,exp},i}]^{\top})}} \quad (9)$$

where $Pr(.)$ is the probability function of $R_{fup,exp,\ i}$, the response indicator for the i^{th} establishment ($R_{fup,exp,i} = 1$ = response in the follow-up wave, $R_{fup,exp,i} = 0$ = nonresponse in the follow-up wave) in the respective experimental group exp, $\boldsymbol{x}_{fup,exp,i}$ is a vector of independent variables (Industry Sector, Number of Employees, and Year of Fundation plus and the experimental mode design indicator), $\boldsymbol{\beta}$ the corresponding vector of coefficients, $[\boldsymbol{x}_{fup,exp,i}\ \boldsymbol{m}_{fup,exp,i}]$ is a vector of establishment characteristics interacting with the experimental mode design indicator, $\boldsymbol{\zeta}$ is the corresponding vector of coefficients, and α the intercept.

We estimate further logistic regression models to identify predictors of cumulative survey participation separately for each mode design combination across the first and follow-up waves. As in the regressions analyzing the follow-up wave, establishment characteristics are used as explanatory variables (without interaction terms). The larger sample size enables us to use a more detailed industry aggregation. For these regressions, the dependent variable, survey participation, is defined as 1 if an establishment participated in both waves and 0 if it did not participate in the first or the follow-up panel waves. Using these survey participation models, we estimate response probabilities for various establishment characteristics with the covariates fixed at their global means. The following formula is applied to predict the probabilities of cumulative survey participation:

$$Pr(R_{\text{cum,exp},i} = 1) = \frac{1}{1 + e^{-(\alpha + \boldsymbol{\beta} \boldsymbol{x}_{cum,exp,i}^{\top})}} \quad (10)$$

where $Pr(.)$ is the probability function of $R_{cum,exp,i}$, the response indicator for the i^{th} establishment ($R_k = 1$ = response in the first and follow-up wave, $R_k = 0$ = nonresponse in the first or follow-up wave) in the exp^{th} experimental group, $\boldsymbol{x}_{cum,exp,i}$ is a vector of independent variables (the same establishment characteristics as for model in Eq. 9, but with more categories for Industry Sector), $\boldsymbol{\beta}$ the corresponding vector of coefficients, and α the intercept.

Analyses of the follow-up panel wave account for inclusion probabilities and nonresponse in the first panel wave. The cumulative regressions are estimated using design weights. All computations consider the stratification variables and were conducted using Stata 17 (StataCorp 2021).

2.2.4 Survey Costs

The fourth research question focuses on the survey costs incurred during the follow-up panel wave and separately for both panel waves. Unfortunately, we are unable to disclose the actual costs due to contractual obligations. Instead, we report relative costs by comparing the costs of different mode designs relative to each other. A fixed price per CATI respondent was agreed upon with the survey institute, which includes multiple call attempts, unsuccessful contacts, and supervision among other aspects. For the self-administered mode designs, the price includes postage, printing, envelopes, and the handling of the survey by the institute, including data entry, programming the web questionnaire, etc. Due to constraints of the contract with the survey institute for the first panel wave, it is not feasible to assign costs separately for each mode design. To overcome this limitation, we utilize the costs for web and mail responses in the follow-up wave and adjust them based on the length of the questionnaire. In doing so, we assume that the costs related to web and mail responses are closely tied to the questionnaire's length, including factors such as more data entry, more programming requirements for a longer questionnaire, and additional data processing. Since this approach ignores fixed costs that are unrelated to the size of the questionnaire, this analysis can be considered more as an upper bound of the costs for the first panel wave. Additionally, we account for costs of printing and postage for invitations and reminders that originate from a separate contract and can be clearly assigned to the mode designs. To compare costs per respondent, we divide the total costs by the number of respondents and then relate these costs to the fixed price per CATI respondent for the follow-up wave and to the current IAB-JVS standard design (first wave: mail-web; follow-up wave: CATI) for both waves.

2.2.5 Qualitative Interviews

In order to gather further insights into the perceptions of web, mail, and CATI modes and to assess the effects of switching mode designs between the starting wave and the follow-up wave on survey participation processes, a total of 46 short structured interviews and 8 in-depth semi-structured interviews were conducted with respondents of both panel waves of the experiment. The following analysis focuses specifically on perceptions around the CATI mode itself and the current IAB-JVS mode design of switching respondents from self-administration to CATI in the follow-up waves. A more comprehensive qualitative analysis of web and mail modes can be found in (Küfner 2024).

As part of the routine pretest for the IAB-JVS, short structured interviews were conducted with respondents from the first wave. These interviews aimed to gather data on establishments' mode preferences and their perceived advantages and disadvantages of web, mail, and CATI modes. Ensuring a representative sample, participating establishments were balanced across experimental groups, establishment sizes, industries, and regions (see Online Supplement Table S71). The interviews, conducted via telephone between February and May 2022, targeted human resources representatives and managers responsible for responding to the IAB-JVS. All interviews were executed by trained interviewers from the IAB, with expertise in short structured interviews.

3 Results

3.1 Response Rates

For all the following analyses, we first compare the effects of the mode design continuation and switch groups on survey participation in the follow-up wave for each of the mode design groups of the first panel wave. This is followed by comparing the effects of the mode design combinations over both panel waves on cumulative survey participation. Online Supplement B provides a tabular version of these response rates, including absolute numbers of responses and unweighted results.

3.1.1 Follow-Up Participation

Fig. 2 presents the nonresponse-adjusted response rates for each experimental group in the follow-up panel wave (Q1/2021). The results show that for each self-administered mode design implemented in the first wave, continuing with the same mode design in the follow-up wave leads to similar, or higher, response rates compared to switching to a CATI follow-up design. Specifically, the concurrent mail-web and single-mode mail designs exhibit significantly higher response rates in the follow-up wave under a mode design continuation than a mode design switch to CATI, with rates of 72% versus 59% and 77% versus 52%, respectively. For the sequential web-to-mail and single-mode web designs, the difference in follow-up response rates between continuing with these mode designs or switching to a CATI mode design is not statistically significant (52% versus 60% and 58% versus 65%, respectively). Additional sensitivity checks (see Online Supplement B) show this pattern is consistent for both tranches and whether or not the establishment provides a contact name in the first panel wave. Overall, these findings suggest that continuing with the same self-administered mode design in the follow-up

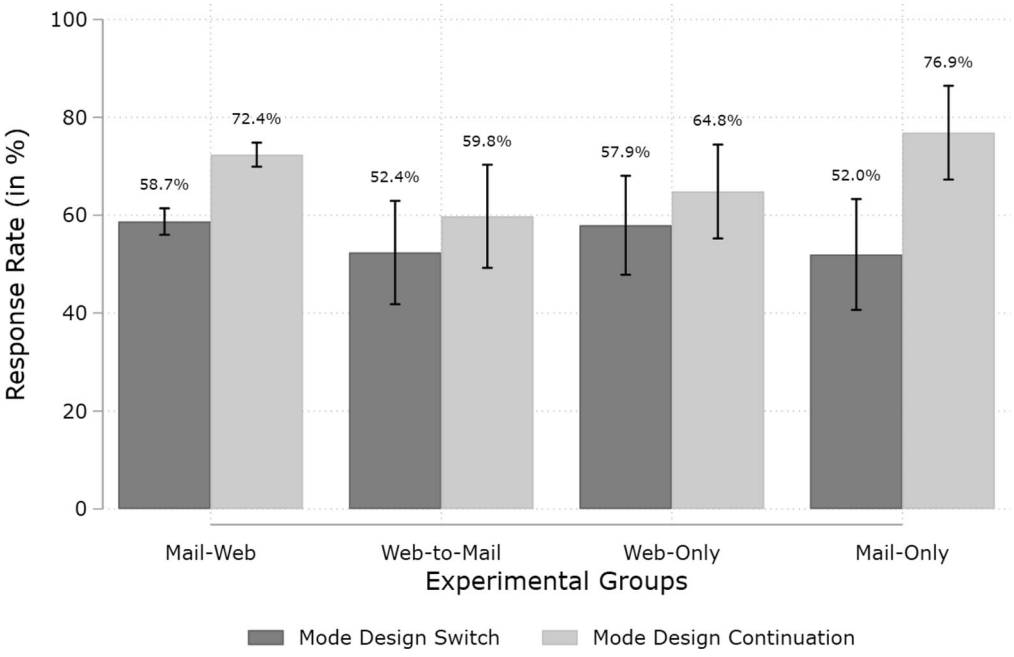


Fig. 2

Follow-Up Response Rates (Nonresponse-Adjusted) and 95% Confidence Interval by Experimental Group, IAB-JVS 2021/Q1

wave leads to similar, or higher, response rates compared to changing to a CATI mode design.

3.1.2 Cumulative Participation

Fig. 3 shows the design-weighted cumulative response rates over both survey waves (Q4/2020 and Q1/2021). Examination of continued participation over the first two waves indicates that using the same self-administered mode design in both waves results in significantly higher cumulative response rates for the concurrent mail-web mixed-mode (11% vs. 9%) and mail-only designs (11% vs. 7%) compared to using these mode designs in the first wave and switching to a CATI design in the follow-up wave. The corresponding differences between a mode design continuation and switch in the web-to-mail (9% vs. 8%) and single-mode web groups (9% vs. 9%) are not statistically significant, although the results marginally favor the mode design continuation. Across all four self-administered mode designs in the first wave with a mode design continuation in the follow-up wave, the cumulative response rates are similar and exhibit no significant differences.

3.2 Nonresponse Bias

In the following, the results regarding the average absolute nonresponse bias overall BHP administrative variables for the follow-up wave and the cumulative average absolute nonresponse bias across both the first and follow-up panel waves are discussed.

For supplementary information, including tables and figures displaying the average absolute nonresponse bias for subgroups of establishment and employee characteristics, please refer to Online Supplement C. Additionally, Online Supplement C.3 and C.4 provide tables of estimated nonresponse biases for individual variables.

3.2.1 Follow-Up Participation

Fig. 4 depicts the average absolute nonresponse bias overall BHP administrative variables for the follow-up wave. The pattern of average absolute nonresponse bias in the follow-up wave mostly reflects the response rate patterns observed earlier. Switching to a CATI design in the follow-up wave exhibits slightly larger nonresponse biases in this wave, on average, compared to continuing the same self-administered mode design from the first wave. This is true for all of the self-administered mode designs considered. However, none of these differences is statistically significant. Overall, the

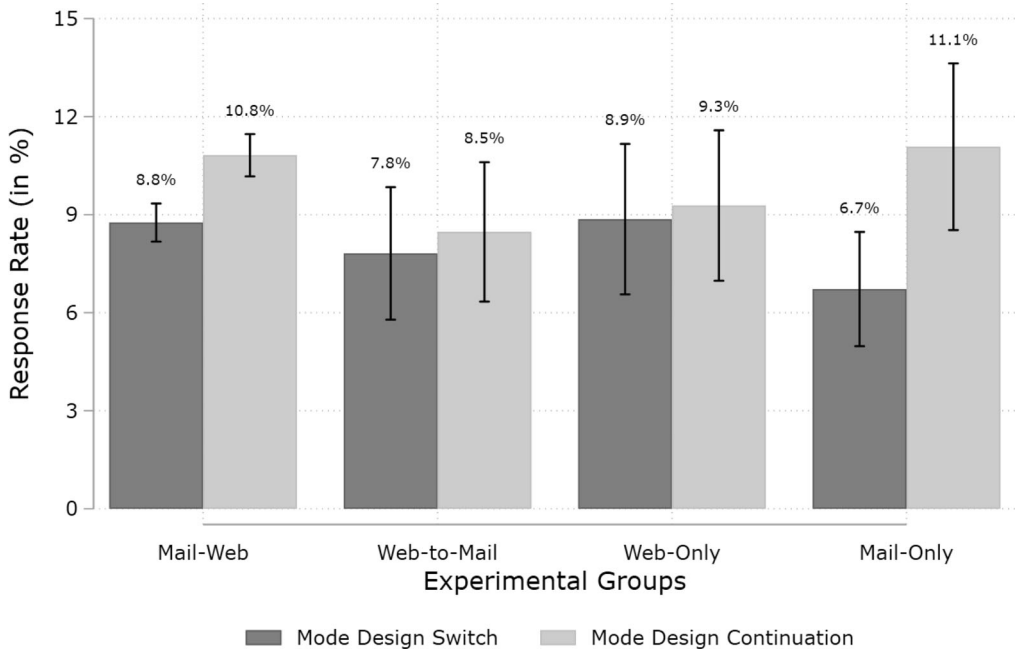


Fig. 3

Cumulative Response Rates (Design-Weighted) by Experimental Group, IAB-JVS 2020/Q4–2021/Q1

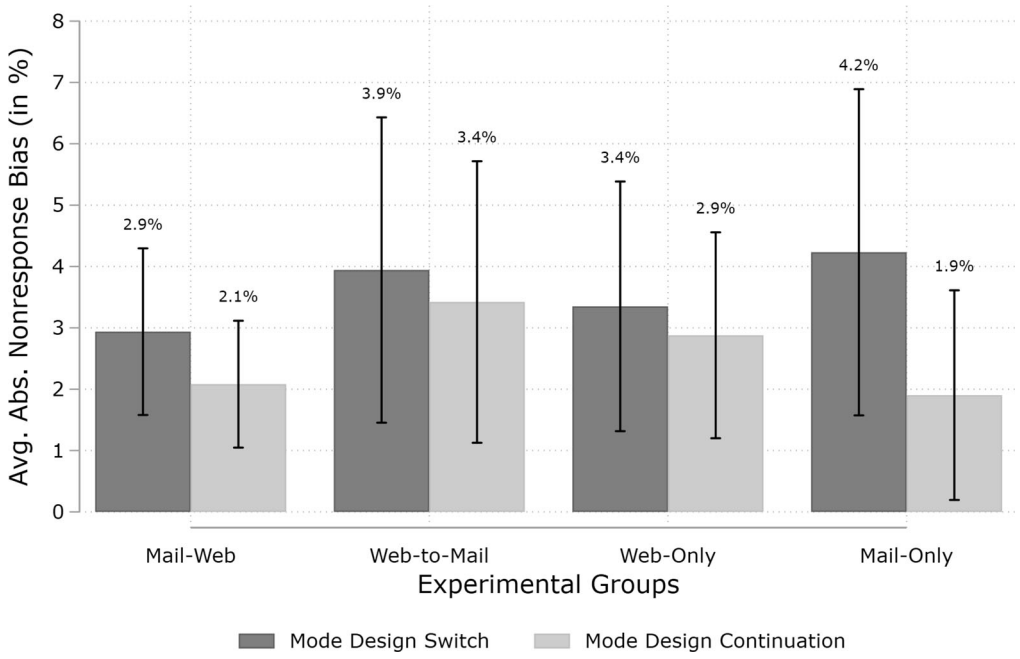


Fig. 4

Average Absolute Nonresponse Bias for the Follow-Up Wave (Nonresponse-Adjusted) by Experimental Group, All Variables BHP 2020

average nonresponse bias estimates are not very large and are always below 4% for every mode design.

Large and statistically significant estimates of nonresponse bias are observed for some individual variables in the follow-up wave. For example, establishments with 20–49 employees are overrepresented among respondents in the CATI follow-up group when a single-mode mail design was used in the first wave (4%-points). Further, establishments in the service industry are underrepresented in the web-to-mail continuation group by 9%-points and establishments founded in the 2000s by 11%-points in the CATI follow-up group when a web-to-mail design was used in the first wave. In terms of employee characteristics, large nonresponse biases are evident in the single-mode web continuation group, where establishments without apprentices are underrepresented among respondents by 7%-points. The smallest proportion of regular employees is underrepresented by 16%-points in the CATI follow-up group when a web-to-mail design was used in the first wave.

3.2.2 Cumulative Participation

Fig. 5 presents the cumulative average absolute nonresponse bias across both the first and follow-up panel waves. When looking at the cumulative participation perspective, we observe that all experimental groups display nearly identical average nonresponse biases at approximately 4%-points (with the web-to-mail continuation group slightly higher at 6%-points) and there are no statistically significant differences between them. Similar patterns are observed when considering subgroup analyses based on establishment and employee characteristics, as well as when examining the median nonresponse bias across different variables (see Online Supplement C).

Regarding the individual nonresponse bias estimates, it becomes evident how cumulative nonresponse bias can be disentangled into wave-specific nonresponse biases that either align or diverge in direction. For instance, the service industry category exhibits a significant and substantial cumulative nonresponse bias of –16%-points in the continuation of the web-to-mail design. This bias is influenced by effects from both the first wave (–8%-points) and the follow-up wave (–8%-points). On the other hand, establishments with the lowest proportion of employees on full-time contracts display a comparably large and statistically significant nonresponse bias of –12%-points in the follow-up wave for a CATI switch design, which is partially offset by the nonresponse bias estimate in the first panel wave under a single-mode mail design (2%-points), resulting in a statistically insignificant cumulative nonresponse bias of –10%-points. Within the single-mode mail design of the first wave and switching to CATI in the follow-up wave leads to virtually no cumulative nonresponse bias (1%-

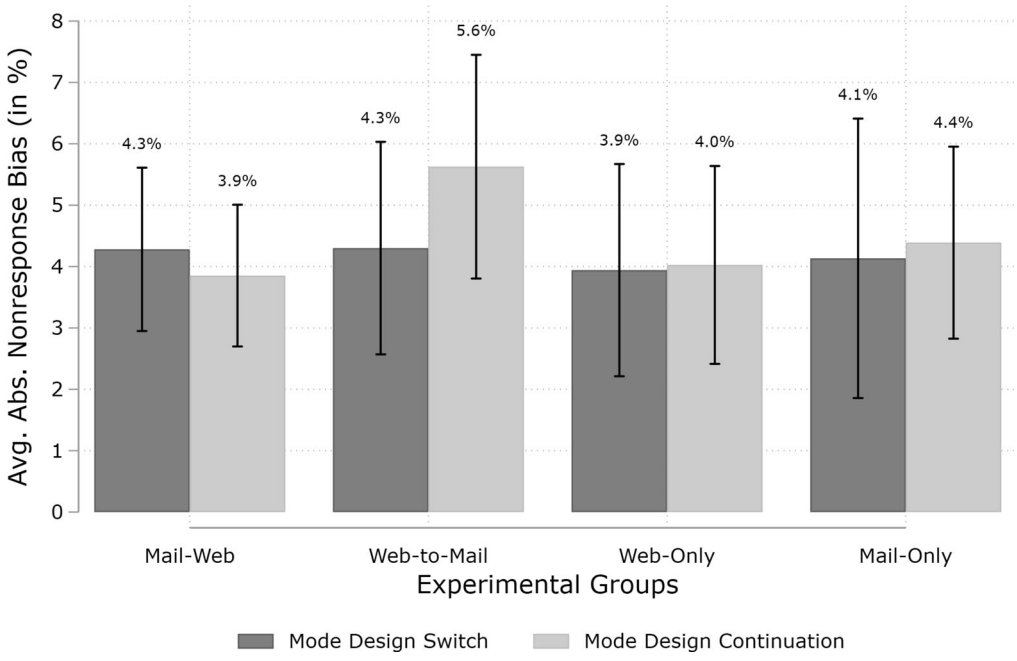


Fig. 5

Average Absolute Nonresponse Bias Cumulative for Both Waves (Design-Weighted) by Experimental Group, All Variables BHP 2020

points) for establishments with 1–9 employees compared with an overrepresentation of these establishments in the continuation group of the single-mode mail design (12%-points). Furthermore, it can be seen that establishments founded in the 2010/20s are statistically significantly underrepresented in the single-mode web continuation group by 15%-points, but are not statistically underrepresented in the single-mode web design who were switched to CATI in the follow-up wave (–2%-points). Establishments without highly-educated employees are statistically significantly underrepresented among respondents in the web-to-mail continuation group (–17%-points), but insignificantly overrepresented in the web-to-mail group that was switched to CATI in the following wave (5%-points).

To sum up the nonresponse bias analysis, the findings indicate that using self-administration in both waves does not have a negative effect on average follow-up and cumulative nonresponse bias compared to switching to a CATI design in the follow-up wave. Furthermore, all self-administered mode designs yield similar levels of average follow-up and cumulative nonresponse bias. However, when analyzing biases for individual variables, several differences between the mode designs and also variations between the two waves were observed.

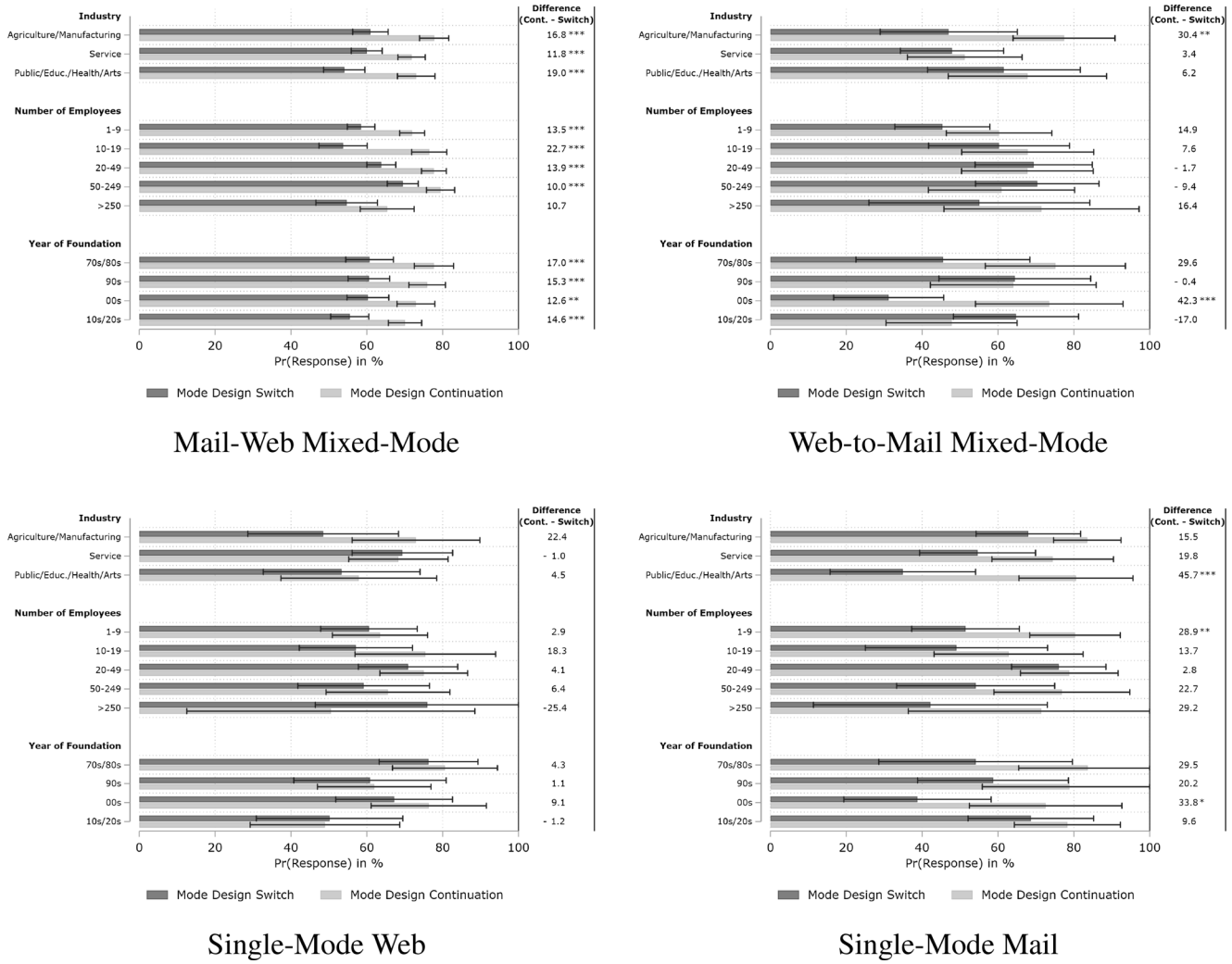


Fig. 6

Predicted Probabilities and 95% Confidence Intervals of Follow-Up Survey Participation in the Mode Design Continuation and Mode Design Switch Groups by First Wave Mode Design, IAB-JVS 2021/Q1. ***a*** *Mail-Web Mixed-Mode,* ***b*** *Web-to-Mail Mixed-Mode,* ***c*** *Single-Mode Web,* ***d*** *Single-Mode Mail*

3.3 Predictors of Survey Participation

3.3.1 Follow-Up Participation

Fig. 6 shows the predicted probabilities of follow-up survey participation for the mode design continuation and switch groups, presented separately by the mode design used in the first wave.

For establishments in the concurrent mail-web mixed-mode design of the first wave (see Fig. 6a), it is evident that all establishment characteristics (industry, number of employees, and year of foundation) are associated with a higher likelihood of response (by at least 10%-points) in the second wave when they are followed up using the same self-administered mode design as opposed to a switch to the CATI design. These differences are statistically significant for each characteristic with the exception of establishments with more than 250 employees. Establishments with 10–19 employees exhibit the most substantial difference of 23%-points between the mode design continuation and switch groups.

For those in the sequential web-to-mail mixed-mode design of the first wave (see Fig. 6b), establishments in agriculture/manufacturing have a 30%-points higher predicted response probability in the follow-up wave under the continuation design than the CATI switch design. The other significant effect is observed for establishments founded in 2000s, which have a 42%-points higher response probability in the continuation design. All other establishment characteristics do not differ significantly in their response probabilities between the continuation and switch designs.

For the single-mode web design of the first wave (see Fig. 6c), there are no statistically significant differences in the follow-up response probabilities between the mode design continuation and switch groups.

Lastly, for the single-mode mail design of the first wave (see Fig. 6d), establishments in the public/education/health/arts industry have a statistically significant 46%-points higher response probability in the follow-up wave under the continuation design than the switch design. Similarly, establishments with 1–9 employees show a 29%-points significantly higher response probability under the continuation design than the switch design. Additionally, the continuation design outperforms the switch design by 34%-points for establishments founded in the 2000s.

In summary, none of the comparisons indicate an adverse effect of replacing the CATI mode design with a continuation of the self-administered mode design on participation in the follow-up wave. The mode design continuation group consistently outperforms (or performs similarly to) the mode design switch group in terms of estimated response probabilities for all establishment characteristics.

3.3.2 Cumulative Participation

Fig. 7 shows the predicted response probabilities for cumulative participation in both waves by different establishment characteristics. The results are presented separately by the eight mode design combinations over the two waves. Overall, the results indicate that response patterns are comparable across all eight mode design combinations. The predicted response probabilities are similar with mostly overlapping confidence intervals for each establishment characteristic across the mode combinations. The results also hold when comparing the average marginal effects based on the same logistic regression models; see Table S70 of the Online Supplement.

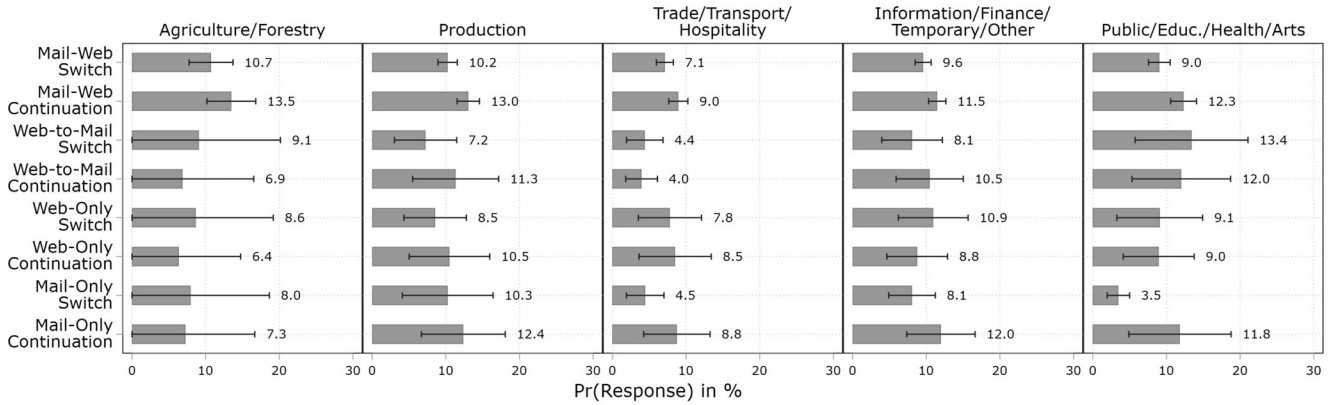
In summary, the analysis of both follow-up and cumulative panel participation shows that replacing CATI in the follow-up wave with a continuation of self-administration from the first wave results in higher (or similar) response probabilities across all establishment characteristics. Thus, we find no negative effects of forgoing interviewer-administered follow-ups on panel participation.

3.4 Survey Costs

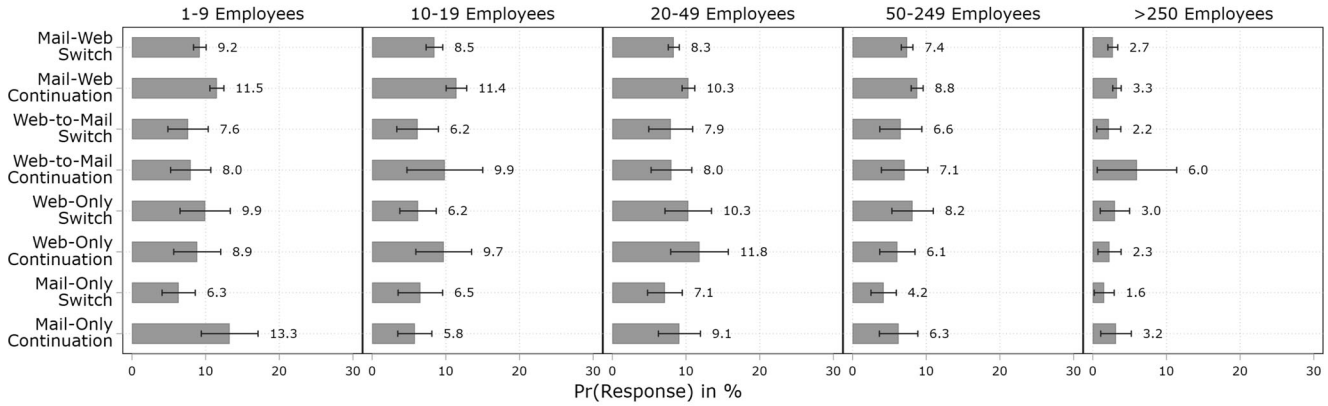
3.4.1 Follow-Up Costs

Fig. 8 presents empirical evidence regarding the cost structure associated with data collection in the follow-up panel wave, specifically comparing the costs of switching to the CATI mode design to that of continuing with the alternative self-administered mode designs.

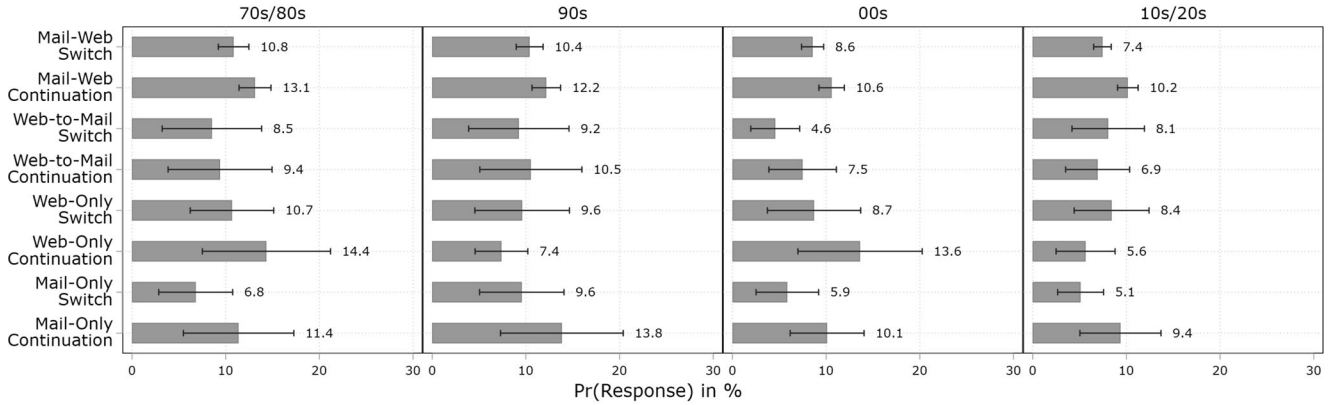
The findings indicate that all self-administered mode designs are associated with lower costs than CATI on a per-respondent basis. Among the self-administered mode designs used in the follow-up wave, the single-mode mail design is the most expensive, with costs per respondent reaching 80% of the costs associated with CATI. This is due to the relatively high costs linked to each contact attempt and data entry required for this mode design. The concurrent mail-web mixed-mode design is the second-most expensive design (66% of CATI costs). The higher costs of this mode design are attributable to the expenses incurred for postage and printing, although these are partially offset by the lower data entry costs associated with this mode design. The sequential web-to-mail mixed-mode design and the single-mode web design, which have higher numbers of web respondents and less expensive mailings, are the least expensive mode designs, with costs per respondent reaching 46% and 33% of the costs associated with CATI, respectively. These findings underscore the potential cost savings associated with self-administered mode designs, particularly those that utilize



Industry



Number of Employees



Decade of Foundation

Fig. 7

Predicted Probabilities and 95% Confidence Intervals of Cumulative Survey Participation in the Mode Design Continuation and Mode Design Switch Groups by First Wave Mode Design, IAB-JVS 2021/Q1

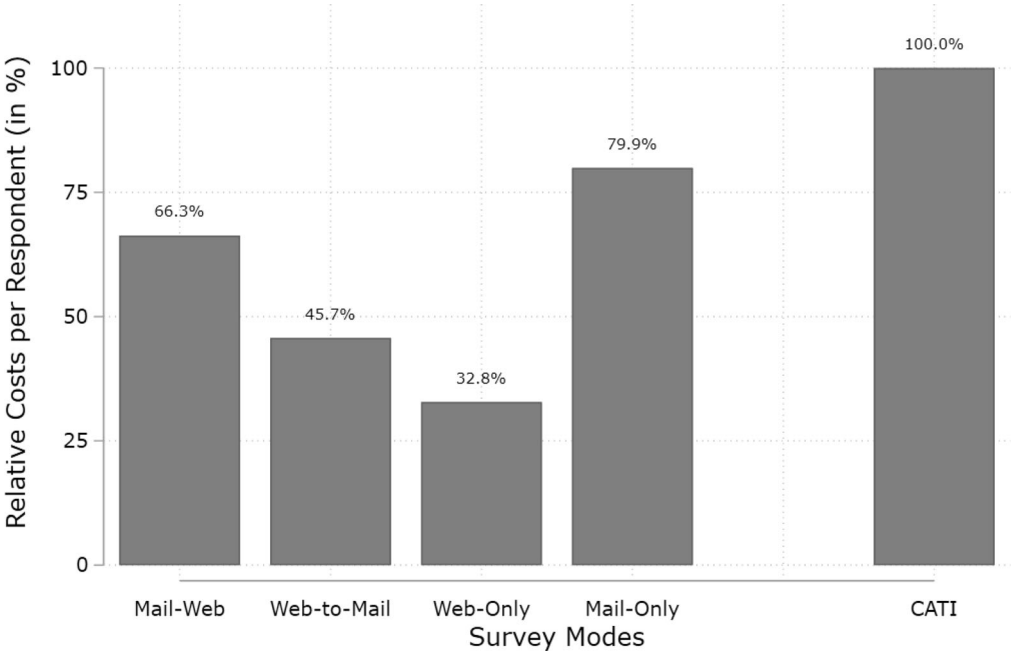


Fig. 8

Relative Follow-Up Survey Costs per Respondent and Mode Design Compared to CATI Mode Design, IAB-JVS 2021/Q1

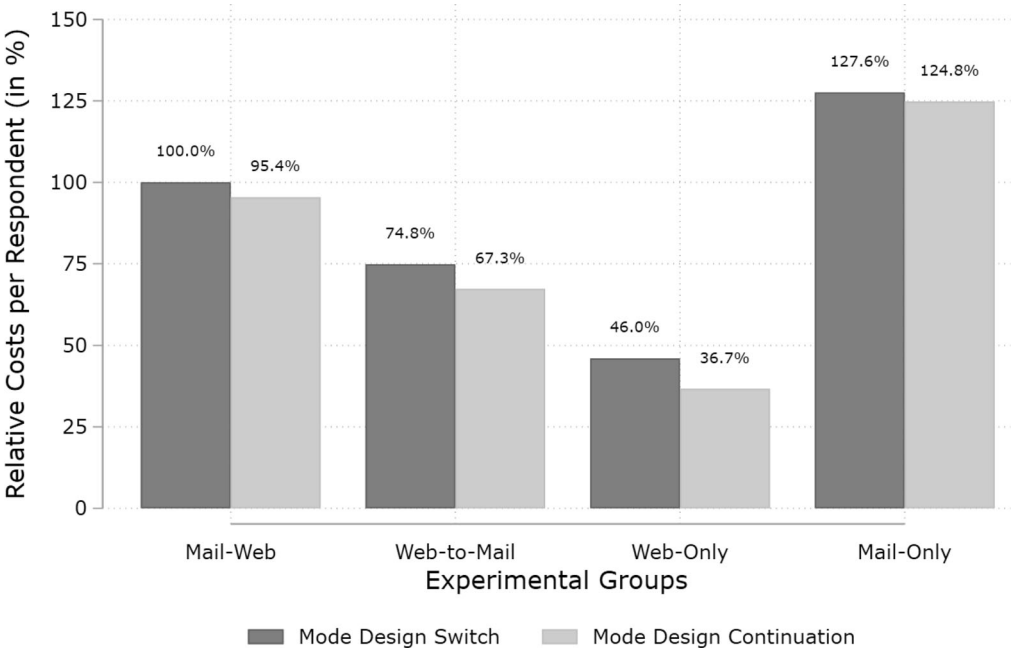


Fig. 9

Relative Cumulative Survey Costs per Respondent and Mode Design Combination Compared to the Current IAB-JVS Mode Design Combination (First Wave: Mail-Web, Follow-Up Wave: CATI), IAB-JVS 2020/Q4–2021/Q1

a push-to-web design, in the context of establishment surveys. Moreover, if the follow-up questionnaire were to be even longer, the cost advantage of these self-administered mode designs would be even more pronounced.

3.4.2 Cumulative Costs

Fig. 9 depicts the cost structure associated with the mode design combinations used across both panel waves. It presents the relative costs of the different mode design combinations relative to the standard mode design combination used in the IAB-JVS (first wave: mail-web, follow-up wave: CATI).

Among the different designs, the single-mode mail design used in the first panel wave exhibits the highest relative per-respondent costs for both the mode design switch and continuation groups (switch: 128%, continuation: 125%) relative to the standard mode design combination used in the IAB-JVS. This is followed by the mode design combination commencing with the current standard of the IAB-JVS first wave (switch: 100%, continuation: 95%). Cost savings are observed for designs employing a sequential web-to-mail design (switch: 75%, continuation: 67%) and a single-mode web design in the first panel wave (switch: 46%, continuation: 37%). The dominance of the cumulative costs associated with the first panel wave can be attributed to the utilization of a substantially longer questionnaire and a larger number of contacted establishments necessary to achieve the desired number of respondents required to meet Eurostat precision targets.

To conclude, replacing the CATI follow-ups with a continuation of self-administered mode designs results in small cumulative cost savings over both waves, ranging from 3%-points (single-mode mail in the first wave) to 9%-points (single-mode web in the first wave). The potential cumulative cost savings would be greater if the follow-up wave used a longer questionnaire.

3.5 Qualitative Insights

To gain a comprehensive understanding of establishments' perceptions of web, mail, and CATI modes and to explore their attitudes and preferences towards changing to a CATI mode design in the follow-up waves, we additionally conducted 8 in-depth semi-structured interviews with respondents of the first and the follow-up panel waves. These participating establishments were selected to represent each experimental group, establishment size class, industry, and region. Recruitment was carried out by the interviewers of the short structured interviews. We used a semi-structured interview guide (see Online Supplement E.3) and the interviews lasted between 31 and 55 min. Each interview was attended by one or two representatives or managers from the human resources department. The interviews were conducted either through video or telephone and were recorded with the participant's consent. Additional methodological details regarding the short structured and in-depth interviews can be found in Online Supplement E.

When asked about their impression of switching mode designs from self-administration to CATI compared to responding in a constant self-administered mode design across two survey waves, respondents strongly preferred keeping the self-administered mode design constant. Most dominantly, respondents explained that a mode design switch would lead to deviating response processes. In the case of a mode design switch to CATI in the follow-up wave, the internal participation process established in the first wave cannot be replicated in the same way: *"Then [i.e., using a self-administered survey mode in the first panel wave] you already know what you have to prepare, and then it [answering in the follow-up wave] goes quickly."* (See Online Supplement Table S75, No. 1). Another respondent emphasized the advantage of keeping the data collection procedure constant: *"It also makes sense that you structure it [i.e., collect data] the same way [...] this uniformity is [...] overall a bit more reasonable."* (See Online Supplement Table S75, No. 2). Another aspect is that a constant mode design and recruitment strategy increases the probability of being recognized as a follow-up to a previous survey: *"That [a constant mode] makes it easier. You have a recognition value and you know that it's not that extensive that you can integrate it well [into your workday]."* (See Online Supplement Table S75, No. 3). This perception as a follow-up survey is of essential importance for panel surveys, since respondents then link the follow-up wave with the first survey wave and hence do not have to be convinced again of the seriousness and importance of their participation.

Next to the mode design switch element, interviewees talked mostly about the advantages and disadvantages of the CATI mode itself. The most highlighted advantage mentioned during the short interviews was that additional clarification for unclear questions could be provided in telephone interviews. A respondent in the in-depth interviews paraphrased it saying: *"One can also ask questions [...]. You can communicate with each other much better."* (see Online Supplement Table S73, No. 1). A CATI mode design, according to some establishments, is linked to less mental burden. Both the preparation for the interview and the actual interview are structured by the interviewer and are not the responsibility of the establishment. In particular, this is true for scheduled interviews: *"It would be for me already better that one says [...] at that time I call and then we do it [the interview]. So otherwise [in the case of self-admin-*

istered modes] it can be that it is lost to me, then I push it further and further out." (see Online Supplement Table S73, No. 2). Some respondents mentioned that they like to talk to people personally and hence enjoy this aspect of telephone interviewing: *"In terms of personal exchange, it has a different character, because HR managers thrive on talking to people."* (see Online Supplement Table S73, No. 3). Related to that, establishments saw an advantage in telephone interviews that they can provide additional information to explain their particular situations, which might help to be seen as a more fruitful contribution: *"Don't you need more info? Or: I would like to give you more info, for example, about this job vacancy. And I'm probably way too boxed in there [when answering self-administered surveys]."* (see Online Supplement Table S73, No. 4).

Interestingly, other establishments perceived CATI and especially telephone calls without an appointment as an interruption in their day-to-day business and, thus, as a very burdensome way of being interviewed. One respondent stated: *"I would find it good if something like this [telephone interview] is announced to me. Otherwise it [without appointment] costs [...] each side time."* (see Online Supplement Table S74, No. 1). Another reported disadvantage of CATI is that respondents have less time to prepare answers to factual or detailed questions (e.g., checking internal systems) when the questions are not available in advance. In contrast, establishments believe that CATI surveys are a good way to grasp data about rough estimations, trends, or opinions. One respondent expressed it in the following way: *"When they [the survey institute] say: O.k. We recorded the following situation at that time [previous wave interview]. Is the trend consistent, worse, or better? You can do that by phone, you don't need precise numbers. If [...][the survey institute asks if] I have 520 to 530 to 540 [job vacancies], I need to look that up. And then it makes sense to go online."* (See Online Supplement Table S74, No. 2). Moreover, some respondents reported privacy concerns when interviewed by telephone: *"We do not provide information by telephone without further notice."* (See Online Supplement Table S74, No. 3) or that the interviewer cannot prove legitimacy, which could also lead to refusals: *"On the phone, anyone can tell you that they are any institution."* (See Online Supplement Table S74, No. 4). One respondent also stated that he got so many advertising calls that he often tries to decline requests by phone: *"So that's sometimes the crux of this unexpectedness [i.e., unannounced calls] because you're always [...] suspecting something else. [...] That's the defense reflex first."* (See Online Supplement Table S74, No. 5). In the in-depth qualitative interviews, some respondents stated additional disadvantages that weren't mentioned during the short interviews. For example, CATI comes with a higher degree of unpredictability as respondents can only control to a limited extent the date and the duration of the interview and do not know the questionnaire in advance. This uncertainty is recognized as a burden of the CATI mode.

In summary, respondents in both the short structured and in-depth interviews perceived more disadvantages of a switch to a CATI mode design in the follow-up waves or the use of the CATI mode in general than when using a constant self-administered mode design in an establishment panel. These disadvantages reflect a higher burden of participation, for example, when established response routines have to be changed because of a change to the mode design from one wave to the next. A summary of the advantages and disadvantages of web, mail, and CATI modes based on the short structured interviews is provided in Online Supplemental Table S72.

4 Conclusion

This study addressed important research gaps related to continuing with the same mode design versus introducing a new mode design in the follow-up wave of establishment panel surveys. Specifically, we analyzed the effects of replacing a CATI mode design in the follow-up wave of a panel with the continuation of a self-administered mode design used in the first panel wave. While previous studies have examined replacing (or supplementing) interviewer-administered modes with self-administered modes in ongoing household panels (e.g., (Jäckle et al. 2015; Bianchi et al. 2017; Allum et al. 2018), there is only limited evidence on such changes in establishment panel surveys (e.g., (Gleiser et al. 2022; Zuckerbraun 2013; Ellguth et al. 2014; DesRoches et al. 2007; Moore 2016). This research gap includes the lack of experimental evidence on continuing or altering the mode design on response rates, nonresponse bias, subgroup participation, and survey costs. Given the recent shift to work-from-home and flexible working arrangements in response to the COVID-19 pandemic and their potential negative effects on CATI survey participation (Küfner et al. 2022b), it was especially important to systematically explore the effects of replacing CATI with self-administered data collection techniques.

To address these gaps, we analyzed an experiment comparing a mode design continuation of various self-administered mode designs with a switch to a CATI mode design in the follow-up wave of the IAB-JVS. The first panel wave started with four experimental mode design groups: a concurrent mail-web mixed-mode, a sequential web-to-mail mixed-mode, a single-mode web, and a single-mode mail design. The continuation of these self-administered mode designs yielded higher (for the mail-web and the single-mode mail groups) or similar (for the web-to-mail and single-mode web groups) response rates compared to

switching to the CATI mode design in the follow-up wave. The continuation of self-administration also performed slightly better than the switch to the CATI mode design in terms of cumulative response rates over both waves. This was especially pronounced for the single-mode mail and concurrent mail-web designs, perhaps due to the short (one-page) follow-up questionnaire that was clearly visible to establishments that received the paper version. The results showed no significant differences in cumulative response rates between all mode designs that used a continuation of self-administration. Moreover, all experimental groups displayed similar average nonresponse biases in the follow-up panel wave and cumulatively over both panel waves, with no statistically significant differences. Nevertheless, significant variations in individual nonresponse bias estimates were identified among the different designs.

The results of the survey participation models showed that replacing the CATI follow-ups with a continuation of self-administration leads to higher (or similar) predicted response propensities concerning industry, establishment size, and decade of establishment foundation in the follow-up wave. There were no meaningful differences between the eight mode design combinations with respect to cumulative response propensities across both waves.

Lastly, the cost analysis illustrated that the continuation of a self-administered mode design could result in cost savings compared to switching to a CATI design in the follow-up wave. The single-mode web and web-to-mail designs demonstrated the greatest potential for relative cost savings, with 33% and 46% of the CATI costs per respondent, respectively. Meanwhile, the mail-web and single-mode mail designs also achieved cost savings but to a lesser extent (66% and 80% of the CATI costs). The cumulative costs for both panel waves combined were dominated by the costs associated with the first panel wave since the recruitment and data collection costs are much higher for the more comprehensive first wave. Hence, the cumulative costs are less for the single-mode web design (37%–46% of the standard IAB-JVS design costs) and the sequential web-to-mail design (67%–75% of the standard IAB-JVS design costs) than the standard IAB-JVS design (first wave: mail-web, second wave: CATI) and the single-mode mail design (125%–128% of the standard IAB-JVS design costs). The potential savings in cumulative costs ranged only from 3%-points (single-mode mail) to 9%-points (single-mode web) when replacing the CATI mode design with a continuation of self-administration in the follow-up wave.

To gain deeper insights into the perceptions of the CATI mode and shifting from it to self-administration in the proceeding waves, qualitative interviews were performed with previous survey respondents. The qualitative study revealed that changing from a self-administered mode design in the first wave to a CATI mode design in the follow-up wave imposes additional burdens on respondents. This is attributed to the inability to replicate established response processes used in the first wave and the lower likelihood of recognizing the survey as a follow-up to the previous one under a different mode design. Regarding the use of the CATI mode itself, interviewees highlighted certain advantages, such as the opportunity to clarify unclear questions (see also (Haraldsen 2013). But respondents also identified disadvantages of the CATI mode, including perceptions of it being burdensome and interrupting their daily business. Privacy and legitimacy concerns with the CATI mode were also reported. Consistent with prior research (e.g., (Haraldsen 2013; Scholtus 2014), respondents remarked that CATI is less suitable for answering factual questions that require checking internal record systems.

A notable strength of this study is the comprehensive examination of different survey outcomes, including response rates (both follow-up and cumulative), nonresponse bias, subgroup participation, and survey costs. Moreover, this study uniquely combines quantitative experiments and qualitative interviews to examine the advantages and disadvantages of carrying out interviewer-administered follow-ups versus employing a fully self-administered panel design from establishments' perspectives, which yields additional insights into establishments' perceptions of alternative mode designs. Additionally, the results of the survey participation models can be used to identify possible auxiliary variables (e.g., number of employees) for nonresponse correction models (Küfner 2022), or for adaptive designs that tailor their recruitment strategies to maximize the uptake of different modes.

We note some study limitations and possibilities for future research. First, like much of the previous research, this study relies on a single case study to evaluate the effects of altering the mode design in an ongoing panel survey. The observed effects might be different in other establishment panels with more detailed follow-up interviews or a longer time interval between waves (e.g., annually). In addition, the sample sizes of the experimental groups crossed with establishment characteristics are, in some cases, not overly large (around 80–100 cases). Larger sample sizes may have resulted in more statistically significant effects. Lastly, it is important to emphasize again that the different mode designs may also affect measurement error and the validity of the survey results (Cernat et al. 2021); hence, this aspect should be considered in future work.

References

Allum, N., Conrad, F. G., & Wenz, A. (2018). Consequences of mid-stream mode-switching in a panel survey. *Survey Research Methods*, *12*(1), 43–58.

American Association for Public Opinion Research (2023). *Standard definitions: final dispositions ofcase codes and outcome rates for surveys* (10th edn.). AAPOR.

Australian Bureau of Statistics (2023). Business surveys. https://www.abs.gov.au/statistics/detailed-methodology-information/concepts-sources-methods/labour-statistics-concepts-sources-and-methods/2023/methods-four-pillars-labour-statistics/business-surveys

Bächmann, A.-C., Bellmann, L., Gensicke, M., Kohaut, S., Möller, I., Schwengler, B., Tschersich, N., & Umkehrer, M. (2023). *IAB-Betriebspanel (IAB-BP) 1993–2022*. FDZ-Datenreport 16/2023. Nürnberg: Research Data Centre of the Federal Employment Agency.

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 ofOfficial Statistics*, *33*(2), 385–408.

Bossler, M., Gürtzgen, N., Kubis, A., Küfner, B., & Lochner, B. (2020). The IAB Job Vacancy Survey: design and research potential. *Journal for Labour Market Research*, *54*(1), 1–12.

Brenzel, H., Czepek, J., Kiesl, H., Kriechel, B., Kubis, A., Moczall, A., Rebien, M., Röttger, C., Szameitat, J., Warning, A., & Weber, E. (2016). *Revision of the iab job vacancy survey: backgrounds, methods and results*. IAB-Forschungsbericht 4/2016. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung.

Buiten, G., Snijkers, G., Saraiva, P., Erikson, J., Erikson, A.-G., & Born, A. (2018). Business data collection: toward electronic data interchange. experiences in Portugal, Canada, Sweden, and the Netherlands with edi. *Journal ofOfficial Statistics*, *34*(2), 419–443.

Cernat, A., & Sakshaug, J.W. (2021). Estimating the measurement effects of mixed modes in longitudinal studies: current practice and issues. In P. Lynn (Ed.), *Advances in longitudinal survey methodology* (pp. 227–249). Wiley.

Couper, M.P. (2008). *Designing effective web surveys*. Cambridge: Cambridge University Press.

De Leeuw, E.D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, *21*(2), 233–255.

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

DesRoches, D., Barton, T., & Ballou, J. (2007). *Understanding web completion in a survey of new businesses*. Quebec. Paper presented at the third ICES-Conference.

Ellguth, P., & Kohaut, S. (2014). *Are they willing to use the web? first results of a possible switch from papi to capi/cawi in an establishment survey*. Paper presented at Statistics Canada Symposium, Quebec.

Ganzer, A., Schmucker, A., Stegmaier, J., & Stüber, H. (2022). *Establishment history panel 1975–2020*. FDZ-Datenreport 3/2022. Nürnberg: Research Data Centre of the Federal Employment Agency.

Gleiser, P., Sakshaug, J.W., Volkert, M., Ellguth, P., Kohaut, S., & Möller, I. (2022). Introducing web in a mixed-mode establishment survey: effects on nonresponse. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *185*(3), 891–915.

Haraldsen, G. (2013). Questionnaire communication in business surveys. In G. Snijkers, G. Haraldsen, J. Jones & D.K. Willimack (Eds.), *Designing and conducting business surveys* (pp. 303–357). Wiley.

Haraldsen, G. (2023). What computerized business questionnaires and questionnaire management tools can offer. In G. Snijkers, M. Bavdaz, S. Bender, J. Jones, S. MacFeely, J.W. Sakshaug, K.J. Thompson & A. van Delden (Eds.), *Advances in business statistics, methods and data collection* (pp. 335–355). Wiley.

Ipsos (2020). *European company survey 2019, Technical and fieldwork report*

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.

König, C., Sakshaug, J.W., Stegmaier, J., & Kohaut, S. (2021). Trends in establishment survey nonresponse rates and nonresponse bias: evidence from the 2001–2017 IAB establishment panel. *Journal ofOfficial Statistics*, *37*(4), 931–953.

Küfner, B., Sakshaug, J.W., & Zins, S. (2022a). Analysing establishment survey nonresponse using administrative data and machine learning. *Journal ofthe Royal Statistical Society Series A*, *185*(S2), 310–342.

Küfner, B., Sakshaug, J.W., & Zins, S. (2022b). Establishment survey participation during the covid-19 pandemic. *Journal for Labour Market Research*, *56*(1), 1–18.

Küfner, B., Sakshaug, J.W., & Zins, S. (2022). Analysing establishment survey non-response using administrative data and machine learning. *Journal ofthe Royal Statistical Society Series A: Statistics in Society*, *185*(Supplement2), S310–S342. https://doi.org/10.1111/rssa.12942.

Küfner, B., Sakshaug, J.W., Zins, S., & Globisch, C. (2024). The impact of mail, web, and mixed-mode data collection on participation in establishment surveys. *Journal of Survey Statistics and Methodology*. https://doi.org/10.1093/jssam/smae033.

Mize, T.D. (2019). Best practices for estimating, interpreting, and presenting nonlinear interaction effects. *Sociological Science*, *6*, 81–117.

Moore, D.L., & Wojan, T. (2016). *From experimental testing to reality: Outcomes from mixed-mode implementation of the 2014 rural establishment innovation survey*. Paper presented at the fifth ICES-Conference, Geneva.

Nicholls, W. II, Mesenbourg, T. Jr., Andrews, S.H., & de Leeuw, E. (2000). *Use of new data collection methods in establishment surveys*. Paper presented at the second ICES-Conference, New York.

Olson, K., Smyth, J.D., Horwitz, R., Keeter, S., Lesser, V., Marken, S., Mathiowetz, N.A., McCarthy, J.S., O'Brien, E., Opsomer, J.D., Steiger, D., Sterrett, D., Su, J., Suzer-Gurtekin, Z.T., Turakhia, C., & Wagner, J. (2021). Transitions from telephone surveys to self-administered and mixed-mode surveys: AAPOR task force report. *Journal of Survey Statistics and Methodology*, *9*(3), 381–411.

Robertson, K.W., & Hatch-Maxfield, J. (2012). *Data collection in the U.S. bureau of labor statistics' current employment statistics survey*. Paper presented at the Conference of European Statisticians, Geneva.

Rosen, R.J., & O'Connell, D. (1997). *Developing an integrated system for mixed mode data collection in a large monthly establishment survey*. Proceedings of the American Statistical Association.

Scholtus, S., van de Laar, R., & Willenborg, L. (2014). The memobust handbook on methodology for modern business statistics. https://ec.europa.eu/eurostat/cros/system/files/NTTS2013fullPaper_246.pdf

Smith, P.A., & Yung, W. (2019). A review and evaluation of the use of longitudinal approaches in business surveys. *Longitudinal and Life Course Studies*, *10*(4), 491–511.

StataCorp (2021). *Stata Statistical Software: Release 17*

U.S. Bureau of Labor Statistics (2023a). Household and establishment survey response rates. https://www.bls.gov/osmr/response-rates/

U.S. Bureau of Labor Statistics (2023b). Job openings and labor turnover survey. https:// www.bls.gov/opub/hom/pdf/jlt-20130314.pdf. BLS Handbook of Methods

Zuckerbraun, S., Hobbs, M., Greene, A., Harris-Kojetin, L., & Sengputa, M. (2013). *When do 'do it yourself-ers' really do it themselves? using paradata to explore the preference for self-completion modes in a multi-mode survey*. Paper presented at the Joint Statistical Meeting of the American Statistical Association, Vancoucer.