

Risk of Nonresponse Bias and the Length of the Field Period in a Mixed-Mode General Population Panel

Struminskaya, Bella; Gummer, Tobias

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Struminskaya, B., & Gummer, T. (2021). Risk of Nonresponse Bias and the Length of the Field Period in a Mixed-Mode General Population Panel. *Journal of Survey Statistics and Methodology*, 10(1), 161-182. <https://doi.org/10.1093/jssam/smab011>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC Lizenz (Namensnennung-Nicht-kommerziell) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nc/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC Licence (Attribution-NonCommercial). For more information see: <https://creativecommons.org/licenses/by-nc/4.0>

RISK OF NONRESPONSE BIAS AND THE LENGTH OF THE FIELD PERIOD IN A MIXED-MODE GENERAL POPULATION PANEL

BELLA STRUMINSKAYA*

TOBIAS GUMMER

Survey researchers are often confronted with the question of how long to set the length of the field period. Longer fielding time might lead to greater participation yet requires survey managers to devote more of their time to data collection efforts. With the aim of facilitating the decision about the length of the field period, we investigated whether a longer fielding time reduces the risk of nonresponse bias to judge whether field periods can be ended earlier without endangering the performance of the survey. By using data from six waves of a probability-based mixed-mode (online and mail) panel of the German population, we analyzed whether the risk of nonresponse bias decreases over the field period by investigating how day-by-day coefficients of variation develop during the field period. We then determined the optimal cut-off points for each mode after which data collection can be terminated without increasing the risk of nonresponse bias and found that the optimal cut-off points differ by mode. Our study complements prior research by shifting the perspective in the investigation of the risk of nonresponse bias to panel data as well as to mixed-mode surveys, in particular. Our proposed method of using coefficients of variation to assess whether the risk of nonresponse bias decreases significantly with each additional day of fieldwork can aid survey practitioners in finding the optimal field period for their mixed-mode surveys.

KEYWORDS: Coefficient of variation; Field duration; Mixed-mode panels; Nonresponse bias; Response propensity.

BELLA STRUMINSKAYA is an assistant professor in the Department of Methodology and Statistics, Utrecht University, Padualaan 14, Utrecht 3584CH, the Netherlands. TOBIAS GUMMER is a team leader and senior researcher in the Department Monitoring Society and Social Change, GESIS Leibniz Institute for the Social Sciences, B2, 1, Mannheim 68159, Germany.

*Address correspondence to Bella Struminskaya, Department of Methodology and Statistics, Utrecht University, Padualaan 14, Utrecht 3584CH, the Netherlands; Email: b.struminskaya@uu.nl

1. INTRODUCTION

Survey practitioners are often faced with making a decision regarding the length of the fielding period. Longer fielding periods might lead to greater participation yet demand additional time and effort from survey managers and may require additional costs since interviews collected at the end of the field period are relatively expensive due to the effort exerted in gaining cooperation (Kennickell 2008). Furthermore, the resulting data cannot be produced in as timely a manner as with shorter fielding periods. However, shortening the field period might introduce the risk of nonresponse bias, which has been proposed as an alternative dimension for data quality that goes beyond response rates (Groves and Peytcheva 2008). If specific groups of respondents tend to participate either early or late in the field period, the length of the field period might impact nonresponse bias.

In this study, we aim to facilitate informed decisions on an optimal fielding period length by studying the risk of nonresponse bias during the field period in a panel survey in which all respondents were interviewed in either an online or a mail mode in each wave. Thus far, only a few studies have investigated nonresponse bias during survey recruitment. For example, Moore, Durrant, and Smith (2018) have examined the call records from six UK household surveys to assess the representativeness of these surveys at each additional call as compared with the final achieved sample distribution for an array of variables. Sturgis et al. (2017) have demonstrated that the reduction of potential nonresponse bias for these surveys is rather high during the first additional call attempts and subsides after about the fifth call attempt across the analyzed surveys. Kreuter, Müller, and Trappmann (2010) have reached similar conclusions regarding the significant reduction of nonresponse bias resulting from increased call attempts. Still other studies have focused on the post-survey evaluation of the productivity of fieldwork per unit in time—the so-called fieldwork power (Vandenplas and Loosveldt 2017; Vandenplas, Loosveldt, and Beullens 2017). However, the practical implications involved in deciding on field period duration remain understudied. Most importantly, previous studies have not addressed the issue of determining optimal field duration in the context of mixed-mode surveys, which is unfortunate as Gummer and Struminskaya (2020) have demonstrated that the timing of survey participation differs between modes. Furthermore, with the exception of a few studies (such as those of Sturgis et al. 2017 and Moore, Durrant, and Smith 2018), previous research has mostly focused on potential nonresponse stemming from respondents' participation timing by using individual variables rather than aggregate measures (i.e., indicators that combine multiple variables).

In this article, we describe the impact of the length of the field period on the coefficient of variation (CV) of response propensities, a proxy indicator for nonresponse bias. Our approach is similar to that of Moore, Durrant, and Smith (2018). However, in contrast to previous studies, most of which have

predominantly focused on cross-sectional data or on single waves of panel surveys, our focus lies on panel survey data. We thus investigated potential nonresponse bias over multiple waves and explored the potential consequences that shortening the fieldwork period by mode could have on panel attrition. Consequently, our study addresses two research questions: (1) Does a longer fielding time reduce the risk of nonresponse bias? and (2) Is the optimal field period length different for the online and the mail mode?

2. BACKGROUND

The extant literature often focuses on the differences between respondents who respond quickly after receiving a survey request and those who take longer with regard to socio-demographic and substantive variables. Late respondents and early respondents have been reported to differ in terms of age (e.g., [Bates and Creighton 2000](#); [Díaz de Rada 2005](#); [Kruse et al. 2010](#)), nonminority status (e.g., [Voigt, Koepsell, and Daling 2003](#); [Kruse et al. 2010](#); [Sigman et al. 2014](#)), employment status ([Bates and Creighton 2000](#); [Kennickell 2008](#)), and household composition (e.g., [Díaz de Rada 2005](#); [Kruse et al. 2010](#)). Some studies have reported mixed results relating to gender ([Irani, Gregg, and Telg 2004](#); [Kennickell 2008](#); [Rao and Pennington 2013](#); [Sigman et al. 2014](#)) and home ownership ([Bates and Creighton 2000](#); [Kennickell 2008](#); [Rao and Pennington 2013](#)).

In addition, previous studies have reported differences for substantive variables. In a survey on computer skills, [Irani, Gregg, and Telg \(2004\)](#) found differences between early and late respondents in terms of their computer skills. Furthermore, in a consumer behavior survey, [Díaz de Rada \(2005\)](#) found differences between early and late respondents in terms of interest in products and services. [Kypri, Stephenson, and Langley \(2004\)](#) found that early and late respondents varied in terms of drinking behavior, whereas [Voigt, Koepsell, and Daling \(2003\)](#) reported differences in smoking behavior. Moreover, respondents with lower cognitive ability have been found to be more likely to be late or reluctant respondents ([Borg and Tuten 2003](#); [Kaminska, McCutcheon, and Billet 2010](#)). By contrast, studies focusing on the survey effort have found that either the higher number of call attempts or refusal conversion ([Curtin, Presser, and Singer 2000](#))—or adding call attempts that used the same protocol ([Peytchev, Baxter, and Carley-Baxter 2009](#))—did not substantially influence survey estimates. In that sense, a longer field period resembles the use of the same survey protocol for a longer period, and the studies cited above demonstrate that keeping a survey in the field for longer might yield different respondents as opposed to “more of the same.”

The mixed results produced by studying differences in demographic and substantive variables between early and late respondents might stem from the fact that studies have focused on different variables of interest. An alternative

approach would be to investigate indicators that aggregate information across multiple variables of interest rather than focusing on individual variables. Such indicators may be used as a proxy measure for nonresponse bias for large shares or even for the whole range of measures covered in the survey. Recently developed and well-known examples of such indicators include the R-indicator and the CV (Schouten, Cobben, and Bethlehem 2009; Shlomo, Skinner, and Schouten 2012).

The differences between early and late respondents—measured either by aggregate indicators or by using individual characteristics—indicate that the length of the field period may be related to the risk of nonresponse bias in the variables under investigation. The goal of setting the optimal fielding period is to minimize the risk of bias while keeping costs and timeliness in mind. Nonresponse can have different causes: noncontactability, refusal, or incapacity of the respondent (Groves and Couper 1998), and these causes require the implementation of different strategies to reduce nonresponse and the potential risk of nonresponse bias. In face-to-face or telephone surveys, prolonging the fieldwork period might be accompanied by increased effort on the part of the interviewers in contacting the respondents or in refusal conversion. In web surveys, a prolonged fieldwork period might involve additional email reminders sent to potential respondents. While sending the reminders in web surveys can be automated, the effort of the panel management still can be considerable in mixed-mode self-administered surveys: answering the panel members' calls and emails, monitoring the return of postal mail questionnaires (and managing P.O. boxes), and maintaining after-hours respondent hotline. The length of the field period, however, is not independent from fieldwork effort. For example, longer field periods allow for more effort to be exerted. However, effort may produce diminished returns over time. Prior studies have used both these dimensions either focusing on one or on both. The studies that focused on data quality provided by early versus late respondents have defined respondents as being early or late in terms of time passed since invitation (e.g., Wellman et al. 1980; Kruse et al. 2010; Sigman et al. 2014) or (non)response after certain fieldwork effort such as additional reminders (e.g. Kypri et al. 2011), additional contact attempts (e.g., Ullman and Newcomb 1998; Díaz de Rada 2005; Rao and Pennington 2013; Kreuter et al. 2014), or other combinations of fieldwork efforts (e.g., Donald 1960). Yet, others have used combinations of the time and effort dimensions to connect data quality and response timing, basing the distinctions between early and late respondents on the distributions of completed interviews by date (e.g., Dalecki et al. 1993; Bates and Creighton 2000; Irani et al. 2004).

Most of the studies on early and late respondents have focused on interviewer-administered surveys, and the literature on the effects of the length of the fieldwork period in self-administered modes is thereby scarce. This scarcity could be due to the fact that the means of influencing contactability and the potential for refusal conversion are much lower in self-administered

surveys than in interviewer-administered surveys. However, as a large volume of data is currently collected through web surveys (according to ESOMAR, the share of online and digital research in the global breakdown of the spending by research method is 54 percent (ESOMAR 2019, p. 23)), it is imperative to consider the relationship between the length of the fieldwork period and the risk of nonresponse bias for self-administered surveys. Several reasons exist for keeping the fieldwork period as short as possible—that is, without the risk of nonresponse bias—in self-administered surveys (which are predominately web surveys):

- A shorter fielding period allows for a timely collection of data, which is important for gaining insights into time-critical research questions about societal problems (e.g., elections, a refugee crisis, or a public health crisis) and about market research questions regarding which clients can base management decisions using these timely data.
- A short field duration enables a panel study to fit more survey waves per year, thereby utilizing the same number of respondents to answer more research questions.
- Keeping a web survey in the field for a longer period goes hand in hand with greater effort from panel management (which oversees data collection and technical issues), thereby driving up survey costs.

As compared with interviewer-administered telephone and face-to-face surveys, web-based surveys, whose operational procedures are quite standardized, nevertheless vary in field period duration. For example, the Dutch LISS Panel, the German Internet Panel, and the French ELIPSS panel keep their surveys open for one month (Blom et al. 2017), the US-based Knowledge Panel has a field period tailored to its clients' needs (ranging from a few hours to several weeks) (Ipsos 2021), and the Understanding America Study sends one or more surveys to its respondents per month (Alattar, Messel, and Rogofsky 2018). The German mixed-mode GESIS Panel (online and mail) keeps surveys open for two months (GESIS 2017), whereas the mixed-mode (online and telephone) Life in Australia Panel has a fielding period of about two weeks (Social Research Center 2021). The question of the optimal length of the field period is thus not trivial since field periods from large-scale panels from various countries and those that use different mode mixes vary considerably. Due to the rather limited options of influencing fieldwork effort in self-administered surveys beyond sending out reminders for which it has been shown that there is a ceiling effect on the increase of response rates after three reminders in both postal and web surveys (Heberlein and Baumgartner 1978; Dillman 2000; Deutskens et al. 2004; Muñoz-Leiva et al. 2010), we focus on the time dimension.

3. DATA

We used data from the GESIS Panel (GESIS 2017), a mixed-mode probability-based panel of the general population in Germany aged 18–70 at the time of recruitment. The GESIS Panel was recruited in 2013 via face-to-face interviews based on a sample drawn from the population register. The response rate (AAPOR RR1) for the face-to-face interviews was 35.5 percent (AAPOR 2016). The recruitment interview was followed by a self-administered welcome survey (active panel $n = 4,938$ by the end of the recruitment phase).

For the regular panel waves, respondents who used the Internet were asked to participate online, whereas non-Internet users were asked to participate via the mail mode. Internet users who were not comfortable with completing online questionnaires could also opt for the mail mode. Interviewers were asked to encourage willing respondents to participate online but not at the cost of potential dropout. In the ongoing GESIS Panel, about 65 percent of respondents complete surveys online, while about 35 percent respond via mail (Bosnjak et al. 2018). To assess the representativeness of the GESIS Panel with respect to key socio-demographic characteristics, Bosnjak et al. (2018) compared the first wave of the GESIS Panel with the German Microcensus 2013 and other general-population surveys that had been administered face to face. The authors assessed the differences in sample composition between the GESIS Panel and the Microcensus with respect to gender, age, citizenship, marital status, household size, place of birth, education, and household income as being comparable to the differences between other probability-based German general-population studies (the German General Social Survey “ALLBUS” and the German sample of the European Social Survey) and the Microcensus. For household income and education, Bosnjak et al. (2018) reported the greatest deviation from the population reference in the GESIS Panel.

The field period for each wave was set to two months for both modes, with six waves per year. All active panel members always receive an invitation letter with an EUR 5 incentive. For the mail mode, a paper questionnaire and a return envelope are always enclosed. Online panel members additionally receive an email invitation and an email reminder both one and two weeks after the invitation. Panel members who participate via the mail mode thus did not receive any reminders. This design decision to not implement reminders reflects the logistical difficulties associated with tracking, at a given point of time, which respondents have already sent their filled in questionnaires and which have not. The GESIS Panel uses an external service provider to send out invitation documents and code questionnaires that were returned by postal mail. As these design differences do not allow us to compare the effort among the modes, we will treat the results as those of a mode system (cf. Biemer and Lyberg 2003, p. 208; Struminskaya, De Leeuw and Kaczmirek 2015). That is, an entire data collection process designed around a specific mode. In our view, this reflects how mixed-mode panels often operate in terms of optimization within each

mode. For our study, the data collected via the web and mail mode were then combined into a single database that enabled any individual panelist to be distinguished in terms of the mode through which they had participated. Switching between modes was possible, albeit quite uncommon.

We used six waves of the panel conducted between 2014 and 2016, beginning with the first wave after the completion of the recruitment (we only use the data for the original cohort excluding the refreshment sample whose recruitment was carried out in 2016). The first waves of each year were chosen to exclude the influence of the Christmas holiday season, during which response times might follow somewhat different patterns. We chose the fourth wave of each year to ensure the six-month interval between each wave that we chose to analyze. The six waves chosen for the analysis were February–March 2014 (*Wave 1, GESIS Panel designation “ba”*), 2015 (*Wave 7, “ca”*), and 2016 (*Wave 13, “da”*) as well as August–September 2014 (*Wave 4, “bd”*), 2015 (*Wave 10, “cd”*), and 2016 (*Wave 16, “dd”*). For these waves, the completion rates varied from 90.7 percent to 92.8 percent (web) and from 82.4 percent to 90.1 percent (mail), respectively (see the completion rates and the number of panelists invited to each wave in [Table A.2](#) in the Appendix). The cumulative response rates (CUMR1) varied from 20.6 percent to 21.2 percent in the web mode and from 19.3 percent to 21.1 percent in the mail mode, respectively (see <https://www.gesis.org/gesis-panel/documentation/>). It should be noted that in the GESIS Panel, all panel members were invited to complete a wave survey, and no subsample of respondents were selected for some studies but not others.

The date of each respondent’s participation in a panel wave was provided through an automatically generated time stamp for web respondents and through the self-reported date of completion of the mail questionnaire for mail respondents. We chose to use the date of completion rather than the date of receipt of the questionnaire since the completed GESIS Panel questionnaires were sent back to the external service provider. Thus, the tasks with respect to panel management ended for each particular respondent when they filled out the questionnaire and not when the external service provider received the questionnaire.

4. METHODS

Representativeness indicators are a common measure used to assess the risk of nonresponse bias as they quantify sample balance in terms of response propensities. The most frequently used representativeness indicator is the R-indicator, which is defined as the transformed standard deviation of response propensities ([Schouten, Cobben, and Bethlehem 2009](#)). To estimate sample response propensity variation, a statistical model of response is calculated, given a covariate set derived from auxiliary information, such as administrative data or

information from a previous wave (Moore, Durrant, and Smith 2018). Schouten et al. (2016) have demonstrated that reducing the degree of variation (i.e., increasing the sample balance) can help to reduce the risk of nonresponse bias with respect to variables that are used to compute response propensities. Without a direct measure of nonresponse bias, proxy measures including R-indicators, CV, and the fraction of missing information are relied upon. In comparison, coefficients of variation and R-indicators can be used for comparisons within surveys both across waves and over time (Wagner 2012). Due to the association of R-indicators with sample sizes (Shlomo, Skinner, and Schouten 2012), the present study relied on the CV, which is standardized to a sample's response rate and is thus better suited for comparisons. The CV is defined as the standard deviation of response propensities (σ_ρ) divided by the mean response propensity (Schouten, Cobben, and Bethlehem 2009): $CV = \sigma_\rho / \bar{\rho}$. Lower values of the CV indicate higher sample balance and a lower risk of nonresponse bias.

We calculated response propensities (ρ) separately for the mail and web modes using logistic regressions, with participation as the dependent variable (0 = nonresponse; 1 = response) and 28 independent variables that were collected during the recruitment interview and in the first self-administered survey. If respondents attrited from the panel, their participation status was set to missing, and response propensities were calculated for the active panel only. Break off and partial participation were treated as a response. The selected covariates reflect the mechanisms of survey participation, in general (e.g., Groves and Couper 1998), and in panel surveys, specifically (e.g., Watson and Wooden 2009). In addition to the variables theoretically linked with participation, we included key variables from the major topics covered in the survey. This comprehensive set of variables thus covered the basic set of topics included in the panel questionnaires that reflect the content of the whole questionnaire (see Appendix Table A.1). Consequently, the response propensities related not only to participation but also to variables of interest—a necessary requirement when assessing nonresponse bias (e.g., Little and Vartivarian 2005).

Based on our propensity model, we calculated the CV for each day during the field period separately for each survey mode (models are described in Appendix Table A.3). We chose to monitor mail and web respondents' CVs separately to allow for the possibility of making independent decisions regarding whether to stop data collection in either mode.

After calculating the CVs for each mode separately, we compared the CV trends between the modes. It should be noted that our analyses do not constitute a classical study of mode effects; rather, we examined the potential differences between the modes from a logistics or operational perspective. For each of the calculations, we included the respondents who had participated in the respective panel wave up to a given day during the field period. For instance, one wave's CV of Day 10 was calculated based on the response propensities

of all respondents who had participated until Day 10 in that wave divided by the response rate achieved until that day.

To answer our second research question, we determined the optimal fielding period length for both modes with respect to minimizing the risk of nonresponse bias. The optimal fielding period length was comparable to the design phase point capacity (Groves and Heeringa 2006, Lewis 2017, 2019), a point (during data collection) after which the continued use of the current method leads to no further increases or even to decreases in data quality. For our purposes, phase capacity points could be identified by comparing CVs across the field period with the CV values after data collection had been stopped or by comparing a CV at a given point during data collection with the previous CV values during the data collection (Moore, Durrant, and Smith 2018). Since we were interested in finding the point during the field period at which interviewing more people would not help to further reduce the risk of nonresponse bias, we computed bootstrapped 95 percent confidence intervals of the CVs for each day during data collection. We then compared each of these CVs with the CV at the end of the data collection period and used the confidence intervals as an approximation of significant differences. Specifically, we investigated whether the confidence intervals of a CV overlapped with the value of the CV at the end of data collection period. In other words, we determined the point in time at which it would have been possible to end data collection without increasing the risk of nonresponse bias.

5. RESULTS

5.1 Risk of Nonresponse Bias over the Field Period

First, we turn our attention to the development of CVs across the field duration (figure 1). Overall, the pattern is similar for all six waves, with the CV decreasing asymptotically over the course of the field period—first steeply and then reaching a turning point, after which it stabilizes and almost resembles a straight line. A decrease in the CV indicates that the risk of nonresponse bias decreases for any given wave. Thus, as long as the CV decreases, the fieldwork should continue since different types of respondents participate. The stabilization and almost-constant CV indicate that the panel receives “more of the same” respondents or no additional respondents at all. A visual inspection of figure 1 suggests that the fieldwork period could be terminated after the point of reaching stability without an effect on the risk of nonresponse bias.

In terms of differences by mode, Waves 1 and 4 yield different patterns between the two modes, but the CVs for both modes converge and significant differences disappear. This result is assumed to be due to the panelists’ development of participation habits (cf. Lugtig 2014) that influence the timing of participation: The first few months are needed to grow accustomed to being

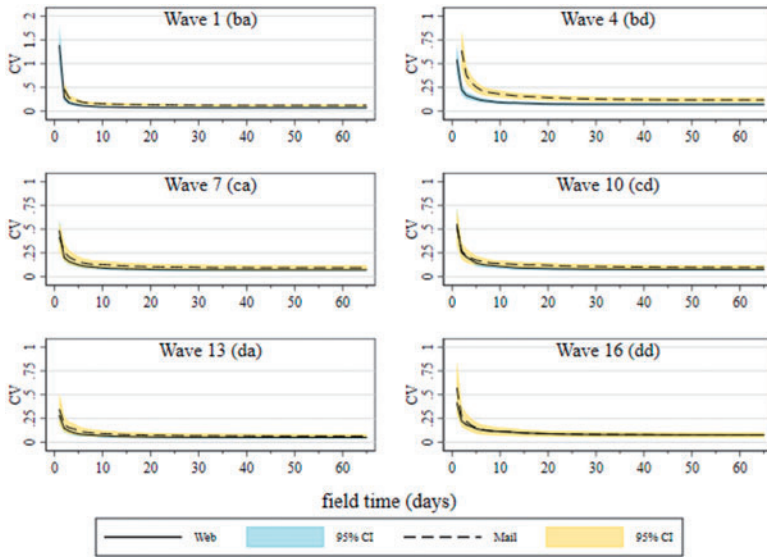


Figure 1. Coefficient of Variation Over the Course of the Field Period in the GESIS Panel.

Note: 65 days was used as an upper limit for the field period to standardize the field length across the different waves and for the purpose of visualizing our results. The maximum nominal field duration in the GESIS Panel was 62 days.

a panel member; afterward, answering questionnaires might become habitual, which could explain why differences in participation between the web and mail modes even out.

5.2 Optimal Fieldwork Cut-Off Points

Next, we take a detailed look at the points after which data collection could potentially be terminated without the risk of increasing nonresponse bias. In addition to visually examining the plots, we estimate the optimal cut-off point for the fielding period in each wave and mode by calculating whether the CV on a given day is significantly different from the CV achieved if all respondents are allowed to participate in the wave without ending the field period early. [Figure 2](#) displays optimal cut-off points for the web and mail modes. The optimal cut-off points are more homogenous in the web mode than in the mail mode. For the GESIS Panel, the optimal cut-off points for the web hover around approximately two weeks after the invitation, while for the mail mode, the point is about three weeks after the invitation. Naturally, these cut-off points will not be exactly the same for every (panel) survey that uses web and

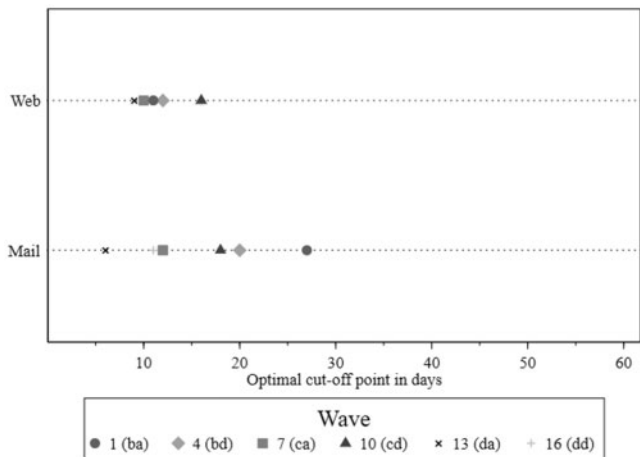


Figure 2. Optimal Field Period by Mode.

mail modes as they are for the GESIS Panel. However, the method that we propose can be used for each survey that has multiple-mode data collection.

5.3 Robustness Check: Taking Potentially Higher Attrition into Account

Our analyses on determining the optimal cut-off points presented above assume that respondents who have not participated up to a certain cut-off point still have a chance to participate in future panel waves. In the GESIS Panel, respondents are excluded at their own will or after three incidents of consecutive nonparticipation. If our strategy is implemented in setting cut-off points of less than two months, the risk of becoming excluded from the panel increases for some of the respondents who have participated late. The literature has found that respondent cooperation in panel surveys is predicted by participation in a previous wave (Olsen 2005), whereas irregular participation and successive nonparticipation predict panel attrition (Watson and Wooden 2009, 2011; Das 2012). Respondents who have missed one wave of a panel survey due to a shortened field period might thus have a greater risk of attrition since they could develop a nonparticipation habit (Lugtig 2014). To investigate this pertinent issue, we performed a robustness check by modeling a situation in which optimal cut-off points were set at 16 days (web mode) and 27 days (mail mode), respectively. Individuals who did not manage to fill out the questionnaires by these points were treated as attriters with no chance of participating in the next wave.¹ This scenario is different from the employed procedures, in

1. It should be noted that for the sake of argument, we assume that the waves for our analysis are consecutive even though we chose only two of the six waves per year.

which respondents were allowed to participate after skipping a wave, and resembles a worst-case scenario. For this scenario, we recalculated the CVs for the web and mail modes based on the known response propensities and compared them with the CVs without reducing the field period length (i.e., in line with the current panel operation). This comparison served to answer the question of what would happen to the risk of nonresponse bias if the fieldwork was shortened according to our optimal cut-off points and people who did not participate until the cut-off permanently dropped out. The results presented in [table 1](#) reveal that the recalculated CVs are as a rule not worse than the CVs that were calculated based on the full available sample, which indicates that even in the worst-case scenario, with cases of late respondents treated as attriters, our proposed method performs reasonably well. In [table 1](#), we also present information on the number of interviews that would be realized under the proposed rules and the percent of panelists that would be lost if these rules were implemented. Whether the loss in sample size would be justified should be decided for each individual panel study. Our results, however, should be interpreted with caution since to assess the effect of attrition, a randomized experiment would need to be conducted in which the fieldwork for part of the sample was shortened and its effect on subsequent attrition was monitored.

6. DISCUSSION

In the present study, we proposed a method that would allow the optimal field period length to be estimated. We also developed an approach that uses the risk of nonresponse bias to inform decisions on the length of the field period. This method can be applied to different surveys and across modes.

Consistent with the existing literature, which has mainly focused on interviewer-administered surveys, our analyses demonstrate that there are non-ignorable differences between early and late respondents in a self-administered panel survey that can be observed based on the fluctuations in the coefficients of variation during the field period. Our analyses are based on data for which the fieldwork had ended before we began the analysis, which allowed us to calculate the potential risk of nonresponse bias at the end of the survey to infer both scenarios in which fieldwork could have been ended earlier on the one hand and the implications of these scenarios for the risk of nonresponse bias on the other hand.

The proposed method can be used not only at the end of the field period but also during fieldwork to adapt the fieldwork processes. This procedure would be reasonable for cross-sectional surveys and panel surveys for which timeliness is an issue (i.e., if test studies using earlier waves are not feasible) and if a survey needs to be terminated early but researchers want to avoid the risk of nonresponse bias.

Table 1. Robustness Check of CVs with Late Respondents as Attriters

Wave	Web				Mail					
	CVcut	CV	N (interviews) with cutoff	N (interviews) without cutoff	% Panelists lost	CVcut	CV	N (interviews) with cutoff	N (interviews) without cutoff	% Panelists lost
1 (ba)	0.082	0.074	2,388	2,770	13.8	0.130	0.109	1,410	1,521	7.3
4 (bd)	0.073	0.072	1,990	2,255	11.8	0.123	0.115	1,120	1,243	9.9
7 (ca)	0.072	0.075	1,806	1,925	6.2	0.090	0.084	881	1,008	12.6
10 (cd)	0.075	0.075	1,611	1,741	7.5	0.087	0.097	767	822	6.7
13 (da)	0.049	0.054	1,468	1,564	6.1	0.059	0.067	683	727	6.1
16 (dd)	0.074	0.078	1,297	1,425	9.0	0.068	0.070	602	646	6.8

NOTE.—CV = coefficient of variation. For CVcut, respondents were coded as attriters if they responded later than the predefined cutoff in an earlier wave.

Our findings contribute to the existing literature in two ways: First, previous studies on the relationship between the field period and nonresponse bias have not focused on panel surveys, and second, they have not specifically addressed mixed-mode surveys. We demonstrated that the optimal field period length differs by mode, which has practical implications for mixed-mode surveys. If an optimal field period is selected for one mode, this field period can be suboptimal for other mode(s). For instance, in the GESIS Panel, the field period was designed to meet the needs of the mail mode, and the web mode field length was thereby adjusted. However, as our findings indicate, field duration in the web mode could be shortened without increasing the risk of nonresponse bias. An alternative, more optimal use of the survey budget would be to allow for field periods to vary between modes.

Some limitations to our study exist and warrant future research. First, our focus lay solely on the risk of bias and not on outcome rates. For panels that have already existed for a long time, the risk of nonresponse (attrition) bias can be secondary to the goal of maximizing participation.

Second, the optimal cut-off points that we identified from our data cannot be generalized to serve as recommendations for all web and mail surveys. Rather, the method of using coefficients of variation to assess whether the risk of nonresponse bias decreases significantly with each additional day of fieldwork can be applied to determine the optimal cut-off points in these surveys. In so doing, the set of variables for estimating response propensities that can be used to calculate coefficients of variation will need to be tailored to the specific survey in terms of mode, topic, and design decisions (for a similar conclusion, see [Moore, Durrant, and Smith 2018](#)). There are multiple factors in survey design, context, and operations that can influence the speed with which panel members may respond (e.g., the speed of mail delivery, whether a deadline for questionnaire return is stated on the invitation, whether sample members have existing relationship with the survey organization and have formed a participation habit, a particular target population, the length of the survey, the ability to resume the web survey after a pause, and many more). Depending on the combination of these features and possible differences between the modes, the cut-off points in other surveys will likely differ from the ones we found to be optimal for the GESIS Panel. We wish to see our study replicated across panels with various design features so that ultimately general recommendations can be made depending on these features and their effect on the field period length. Related to this is our focus on response time while treating fieldwork effort as being fixed. More research is needed on how the length of the fieldwork period is related to the fieldwork effort. Future studies should explore whether the decisions of survey designers about the length of the field period can be made independently from the amount of effort or how such decisions depend on the frequency of effort. For this purpose, a randomized experiment would be required, thus our study is unable to answer this particular question.

Third, due to our focus on the possibility of stopping the data collection for either mode, we calculated the CVs separately for both modes. The combined CV—although unlikely—might follow a different trajectory. Exactly how the decisions that are based on the combined CVs differ from the decisions that are made separately for each mode should be investigated in future studies.

Fourth, we illustrated the method of determining the optimal fieldwork period length by using self-administered surveys. When using this method in interviewer-administered surveys, fieldwork management should be taken into account. For example, our method assumed that all panelists had been contacted simultaneously and had received a comparable number of contact attempts and that the panel management had exerted a comparable amount of fieldwork effort, which is most likely not the case in interviewer-administered surveys, for which interviewers sequentially work the cases assigned to them and might even prioritize specific sampling points or areas. Nevertheless, we see merit in informing decisions on field period length by optimizing key figures (in our case, the risk of nonresponse bias) and conclude that future studies should investigate similar methods for application in interviewer-administered studies.

Fifth, we used bootstrapped confidence intervals to assess differences between CVs during the field period. It should be stressed that the issue of comparing proxy indicators of nonresponse bias during data collection is not trivial, but addressing this topic lay outside the scope of our article. More research in line with [Lewis \(2019\)](#) and [Lewis \(2017\)](#) is required, especially with respect to comparing indicators such as R-indicators and CVs.

Finally, as the GESIS Panel invites all panel members to participate in each wave, our findings might be not generalizable for access panels that carry out subsample studies and in which some panelists receive more survey invitations and have a chance to participate more often than others. In that sense, the GESIS Panel resembles traditional panel surveys, which treat the whole sample in a similar manner. Exactly how the findings generalize to panels that employ the concept of an access panel deserves further investigation.

In our proposed approach, we advocate for making mode-specific decisions on when to stop data collection. This approach represents just one scenario and can lead to various implementations in practice. For example, panel management might decide to calculate the optimal fieldwork length before each wave based on the information from the previous wave(s). This procedure would require additional resources in terms of time and effort. A less resource-intensive alternative would be to calculate the fieldwork length only once after the panel has existed for several waves and to use this information to calculate the coefficients of variation. Another alternative would be to weigh the options as to whether calculating the fieldwork period duration for one mode should be performed after each wave whereas for the other mode it should be calculated once or on a less regular basis (e.g., if the

information needed for the calculations can be extracted easily for the web mode but cannot be acquired in as timely a manner for the mail mode). We thus encourage further research on the implementation strategies of our approach in practical settings.

DATA AVAILABILITY AND QUESTION WORDINGS

The data used in this article are available at the GESIS Data Archive under study number ZA5665: GESIS Panel—Standard Edition, Version 19.0.0, 2017-4-18 Release 19, doi:10.4232/1.12743.

Question wordings (originals in German and in English translation) for all questions asked in the GESIS Panel are available in the “GESIS Panel Codebook Related to ZA5664 and ZA5665” at <https://dbk.gesis.org/dbksearch/download.asp?db=E&id=52375> (last accessed March 7, 2021).

Replication code for this study is available at <https://doi.org/10.7802/2156>.

This study design and analysis was not preregistered.

APPENDIX

Table A.1. Predictors Used to Calculate Response Propensities.

Variable	Coding	Web			Mail				
		M	SD	Min	Max	M	SD	Min	Max
<i>Socio-demographic variables</i>									
Gender: female	0/1	0.488	0.500	0	1	0.567	0.496	0	1
Age	metric	42.271	14.237	18	70	49.888	13.871	18	70
Education: ref. low	0/1	0.337	0.473	0	1	0.390	0.488	0	1
intermediate	0/1	0.522	0.500	0	1	0.227	0.419	0	1
high	0/1	0.949	0.219	0	1	0.944	0.230	0	1
Citizenship: German	0/1	0.556	0.497	0	1	0.616	0.486	0	1
Married	0/1	2.738	1.172	1	5	2.482	1.147	1	5
Household size	metric	0.317	0.465	0	1	0.226	0.418	0	1
Children in household	0/1	0.758	0.429	0	1	0.645	0.479	0	1
Working for pay	0/1	0.264	0.441	0	1	0.397	0.489	0	1
Social class: ref. lower	0/1	0.517	0.500	0	1	0.443	0.497	0	1
working class	0/1	0.138	0.345	0	1	0.071	0.257	0	1
middle class	0/1	0.028	0.164	0	1	0.014	0.116	0	1
upper middle class	0/1	0.321	0.467	0	1	0.463	0.499	0	1
upper class	0/1	5.739	0.751	0	6	5.737	0.840	0	6
Low income	0/1	5.003	1.143	0	6	4.980	1.279	0	6
<i>Substantive variables</i>									
Importance of family	metric								
Importance of free time	metric								

Continued

Table A.1. Continued

Variable	Coding	Web				Mail			
		M	SD	Min	Max	M	SD	Min	Max
Satisfaction with life	metric	3.187	0.843	0	4	3.022	0.956	0	4
Satisfaction with government	metric	4.532	2.163	0	10	4.431	2.076	0	10
Satisfaction with democracy	metric	5.723	2.289	0	10	5.271	2.258	0	10
Satisfaction with economy	metric	6.083	2.053	0	10	5.496	2.189	0	10
Left-right scale	metric	4.520	1.938	0	10	4.678	2.076	0	10
Political engagement	0/1	0.411	0.492	0	1	0.339	0.473	0	1
Political interest: ref. none									
low	0/1	0.175	0.380	0	1	0.182	0.386	0	1
lower intermediate	0/1	0.473	0.499	0	1	0.518	0.500	0	1
upper intermediate	0/1	0.251	0.434	0	1	0.210	0.407	0	1
high	0/1	0.074	0.262	0	1	0.057	0.231	0	1
Civic engagement	factor	0.015	0.773	-2.958	1.473	-0.025	0.799	-2.838	1.544
Trust	0/1	0.722	0.448	0	1	0.625	0.484	0	1
<i>Survey evaluation and participation behavior</i>									
Evaluation of survey (overall)	metric	2.884	0.521	0	4	2.855	0.570	0	4
Perceived survey burden	factor	1.513	0.378	0	3.667	1.531	0.405	0	3.667
Incentive necessary for recruitment	0/1	0.420	0.494	0	1	0.443	0.497	0	1
Likelihood of moving away	0/1	0.598	0.490	0	1	0.460	0.499	0	1
Easily persuaded to participate	0/1	0.862	0.345	0	1	0.779	0.415	0	1
Willingness to become panelist	metric	2.435	0.686	0	3	2.136	0.802	0	3
Experience with surveys	0/1	0.191	0.393	0	1	0.104	0.305	0	1

NOTE.—M = mean, SD = standard deviation, Min = minimum, Max = maximum. All figures based on data from the recruitment and welcome interview.

Table A.2. Number of Invited Panelists and Completion Rates for Each Wave.

Wave	Invited			Completion rate in %		
	Overall	Web	Mail	Overall	Web	Mail
1 (ba)	4,888	3,041	1,847	87.54	90.69	82.35
4 (bd)	4,512	2,871	1,641	88.92	91.36	84.64
7 (ca)	4,249	2,745	1,504	89.95	92.75	84.84
10 (cd)	4,025	2,646	1,379	89.27	90.51	86.87
13 (da)	3,797	2,525	1,272	91.41	92.08	90.09
16 (dd)	3,637	2,436	1,201	89.50	90.35	87.76

NOTE.—Information from GESIS Panel wave reports (available at <https://www.gesis.org/en/gesis-panel/documentation>, last accessed March 7, 2021).

Table A.3. Summary of Logistic Regression on Participation in Panel Waves 1, 4, 7, 10, 13, and 16 for Web and Mail Modes.

Wave	Mode	No. of parameters	Log. Likelihood	Pseudo R^2	N
1 (ba)	web	39	-628.705	0.095	2,313
	mail	39	-485.482	0.094	1,218
4 (bd)	web	39	-541.569	0.109	2,189
	mail	39	-393.399	0.118	1,103
7 (ca)	web	39	-440.821	0.143	2,099
	mail	39	-388.450	0.063	1,005
10 (cd)	web	39	-547.762	0.103	2,037
	mail	39	-296.987	0.109	940
13 (da)	web	39	-480.386	0.069	1,948
	mail	39	-243.201	0.075	874
16 (dd)	web	39	-502.515	0.108	1,878
	mail	39	-268.582	0.064	829

NOTE.—All models are logistic regressions, with participation in the respective wave as the dependent variable. Independent variables are listed in Appendix Table A.1. Listwise deletion was used for cases that had missing values on variables used as predictors in response propensity models.

Note that since the objective of CV indicators is to keep models fixed across surveys, in time or during data collection, it necessitates fixing the selected variables and categories. These models are not boosted to maximize model fit, which is irrelevant in itself. In fact, if nonresponse would be random, model fit would be poor by definition and poor model fit is preferred from the perspective of a survey practitioner.

REFERENCES

- AAPOR (2016). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*: The American Association for Public Opinion Research.
- Alattar, L., M. Messel, and D. Rogofsky (2018). "An Introduction to the Understanding America Study Internet Panel," *Social Security Bulletin*, 78(2), 13–28.
- Bates, N., and K. Creighton (2000). "The Last Five Percent: What Can We Learn From Difficult/Late Interviews?" Proceedings of the Section on Government Statistics and Section on Social Statistics, Alexandria, VA: American Statistical Association, pp. 1–7.
- Biemer, P. P., and L. E. Lyberg (2003). *Introduction to Survey Quality*, Hoboken, NJ: Wiley.
- Blom, A. G., J. M. E. Herzing, C. Cornesse, J. W. Sakshaug, U. Krieger, and D. Bossert (2017). "Does the Recruitment of Offline Households Increase the Sample Representativeness of Probability-Based Online Panels? Evidence from the German Internet Panel," *Social Science Computer Review*, 35, 498–520.
- Borg, I., and T. L. Tuten (2003). "Early versus Later Respondents in Intranet-Based, Organizational Surveys," *Journal of Behavioral and Applied Management*, 4, 1–17.
- Bosnjak, M., T. Dannwolf, T. Enderle, I. Schaurer, B. Struminskaya, A. Tanner, and K. W. Weyandt (2018). "Establishing an Open Probability-Based Mixed-Mode Panel of the General Population in Germany: The GESIS Panel," *Social Science Computer Review*, 36, 103–115.
- Curtin, R., S. Presser, and E. Singer (2000). "The Effect of Response Rate Changes on the Index of Consumer Sentiment," *Public Opinion Quarterly*, 64, 413–428.
- Dalecki, M. G., J. C. Whitehead, and G. C. Blomquist (1993). "Sample Non-Response Bias and Aggregate Benefits in Contingent Valuation: An Examination of Early, Late and Non-Respondents," *Journal of Environmental Management*, 38, 133–143.
- Das, M. (2012). "Innovation in Online Data Collection for Scientific Research: The Dutch MESS Project," *Methodological Innovations*, 7, 7–24.
- Deutskens, E., K. de Ruyter, M. Wetzels, and P. Oosterveld (2004). "Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study," *Marketing Letters*, 15, 21–36.
- Díaz de Rada, V. (2005). "The Effect of Follow-up Mailings on the Response Rate and Response Quality in Mail Surveys," *Quality & Quantity*, 39, 1–18.
- Dillman, D. A. (2000). *Mail and Internet Surveys: The Total Design Method*, New York: Wiley.
- Donald, M. N. (1960). "Implications of Nonresponse for the Interpretation of Mail Questionnaire Data," *Public Opinion Quarterly*, 24, 99–114.
- ESOMAR (2019). *Global Market Research 2019: An ESOMAR Industry Report*, Amsterdam: ESOMAR.
- GESIS (2017). *GESIS Panel – Standard Edition, ZA5665, v19.0.0.*, Cologne, Germany: GESIS Data Archive.
- Groves, R. M., and M. P. Couper (1998). *Nonresponse in Household Interview Surveys*, New York: Wiley.
- Groves, R. M., and S. G. Heeringa (2006). "Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs," *Journal of the Royal Statistical Society*, 169, 439–457.
- Groves, R. M., and E. Peytcheva (2008). "The Impact of Nonresponse Rates on Nonresponse Bias," *Public Opinion Quarterly*, 72, 167–189.
- Gummer, T., and B. Struminskaya (2020). "Early and Late Participation during the Field Period: Response Timing in a Mixed-Mode Probability-Based Panel Survey." *Sociological Methods & Research* online first. 10.1177/0049124120914921.
- Heberlein, T. A., and R. Baumgartner (1978). "Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Publisher Literature," *American Sociological Review*, 43, 447–462.

- Ipsos (2021). "KnowledgePanel: A Methodology Overview," Available at <https://www.ipsos.com/sites/default/files/ipsosknowledgepanelmethodology.pdf>. Accessed March 7, 2021.
- Irani, T. A., J. A. Gregg, and R. Telg (2004). "Choosing to Use the Web: Comparing Early and Late Respondents to an Online Web-Based Survey Designed to Assess IT Computer Skills Perceptions of County Extension Agents," *Journal of Southern Agricultural Education Research*, 54, 168–179.
- Kaminska, O., A. L. McCutcheon, and J. Billiet (2010). "Satisficing among Reluctant Respondents in a Cross-National Context," *Public Opinion Quarterly*, 74, 956–984.
- Kennickell, A. B. (2008). "The Bitter End? The Close of the 2007 SCF Field Period." Paper Read at Joint Statistical Meeting Section on Survey Research Methods, at Alexandria, VA.
- Kreuter, F., G. Müller, and M. Trappmann (2010). "Nonresponse and Measurement Error in Employment Research: Making Use of Administrative Data," *Public Opinion Quarterly*, 74, 880–903.
- . (2014). "A Note on Mechanisms Leading to Lower Data Quality of Late or Reluctant Respondents," *Sociological Methods & Research*, 43, 452–464.
- Kruse, Y., E. Hendarwan, J. M. Dennis, and C. DiSogra (2010). Analysis of Late Responders to Probability-Based Web Panel Recruitment and Panel Surveys. Paper presented at The American Association for Public Opinion Research 65th Annual Conference (AAPOR), Chicago.
- Kypri, K., A. Samaranyaka, J. Connor, J. D. Langley, and B. MacLennan (2011). "Non-Response Bias in a Web-Based Health Behaviour Survey of New Zealand Tertiary Students," *Preventive Medicine*, 53, 274–277.
- Kypri, K., S. Stephenson, and J. Langley (2004). "Assessment of Nonresponse Bias in an Internet Survey of Alcohol Use," *Alcoholism: Clinical and Experimental Research*, 28, 630–634.
- Lewis, T. (2017). "Univariate Tests for Phase Capacity: Tools for Identifying When to Modify a Survey's Data Collection Protocol," *Journal of Official Statistics*, 33, 601–624.
- . (2019). "Multivariate Tests for Phase Capacity," *Survey Research Methods*, 13, 153–165.
- Little, R. J., and S. Vartivarian (2005). "Does Weighting for Nonresponse Increase the Variance of Survey Means?," *Survey Methodology*, 31, 161–168.
- Lutig, P. (2014). "Panel Attrition: Separating Stayers, Fast Attriters, Gradual Attriters, and Lurkers," *Sociological Methods & Research*, 43, 699–723.
- Moore, J. C., G. B. Durrant, and P. W. F. Smith (2018). "Data Set Representativeness during Data Collection in Three UK Social Surveys: Generalizability and the Effects of Auxiliary Covariate Choice," *Journal of the Royal Statistical Society A*, 181, 229–248.
- Muñoz-Leiva, F., J. Sánchez-Fernández, F. Montoro-Ríos, and J. A. Ibáñez-Zapata (2010). "Improving the Response Rate and Quality in Web-Based Surveys through the Personalization and Frequency of Reminder Mailings," *Quality and Quantity*, 44, 1037–1052.
- Olsen, R. J. (2005). "The Problem of Respondent Attrition: Survey Methodology is Key," *Monthly Labor Review*, 128, 63–70.
- Peytchev, A., R. K. Baxter, and L. R. Carley-Baxter (2009). "Not All Survey Effort is Equal: Reduction of Nonresponse Bias and Nonresponse Error," *Public Opinion Quarterly*, 73, 785–806.
- Rao, K., and J. Pennington (2013). "Should the Third Reminder Be Sent? The Role of Survey Response Timing on Web Survey Results," *International Journal of Market Research*, 55, 651–674.
- Schouten, B., F. Cobben, and J. Bethlehem (2009). "Indicators for the Representativeness of Survey Response," *Survey Methodology*, 35, 101–113.
- Schouten, B., F. Cobben, P. Lundquist, and J. Wagner (2016). "Does More Balanced Survey Response Imply Less Non-Response Bias," *Journal of the Royal Statistical Society A*, 179, 727–748.
- Shlomo, N., C. Skinner, and B. Schouten (2012). "Estimation of an Indicator of the Representativeness of Survey Response," *Journal of Statistical Planning and Inference*, 142, 201–211.
- Sigman, R., T. Lewis, N. D. Yount, and K. Lee (2014). "Does the Length of Fielding Period Matter? Examining Response Scores of Early versus Late Responders," *Journal of Official Statistics*, 30, 651–674.

- Social Research Center (2021). "Life in Australia," Available at <https://www.srcentre.com.au/our-research/life-in-australia-panel>. Accessed March 7, 2021.
- Struminskaya, B., E. D. De Leeuw, and L. Kaczmirek (2015). "Mode System Effects in an Online Panel Study: Comparing a Probability-Based Online Panel with Two Face-to-Face Reference Surveys," *Methods, Data, Analyses*, 9, 3–56.
- Sturgis, P., J. Williams, I. Brunton-Smith, and J. Moore (2017). "Fieldwork Effort, Response Rate, and the Distribution of Survey Outcomes: A Multilevel Meta-Analysis," *Public Opinion Quarterly*, 81, 523–542.
- Ullman, J. B., and M. D. Newcomb (1998). "Eager, Reluctant, and Nonresponders to a Mailed Longitudinal Survey: Attitudinal and Substance Use Characteristics Differentiate Respondents," *Journal of Applied Social Psychology*, 28, 357–375.
- Vandenplas, C., and G. Loosveldt (2017). "Modeling the Weekly Data Collection Efficiency of Face-to-Face Surveys: Six Rounds of the European Social Survey," *Journal of Survey Statistics and Methodology*, 5, 212–232.
- Vandenplas, C., G. Loosveldt, and K. Beullens (2017). "Fieldwork Monitoring for the European Social Survey: An Illustration with Belgium and the Czech Republic in Round 7," *Journal of Official Statistics*, 33, 659–686.
- Voigt, L. F., T. D. Koepsell, and J. R. Daling (2003). "Characteristics of Telephone Survey Respondents according to Willingness to Participate," *American Journal of Epidemiology*, 157, 66–73.
- Wagner, J. (2012). "A Comparison of Alternative Indicators for the Risk of Nonresponse Bias," *Public Opinion Quarterly*, 76, 555–575.
- Watson, N., and M. Wooden (2009). "Identifying Factors Affecting Longitudinal Survey Response," in *Methodology of Longitudinal Surveys*, ed. Peter Lynn, pp. 157–182 (Chichester: John Wiley & Sons).
- . (2011). *Re-Engaging with Survey Non-Respondents: The BHPS, SOEP and HILDA Survey Experience*, *SOEP Papers (Vol. 379)*, Berlin: DIW.
- Wellman, J. D., E. G. Hawk, J. W. Roggenbuck, and G. J. Buhyoff (1980). "Mailed Questionnaire Surveys and the Reluctant Respondent: An Empirical Examination of Differences between Early and Late Respondents," *Journal of Leisure Research*, 12, 164–172.