

Experimental Evidence on Panel Conditioning Effects when Increasing the Surveying Frequency in a Probability-Based Online Panel

Carina Cornesse^{1,2,3}, Annelies Blom^{2,3,4}, Marie-Lou Sohnius^{2,5}, Marisabel González Ocanto², Tobias Rettig², and Marina Ungefucht²

¹DIW Berlin

²University of Mannheim

³University of Bremen

⁴University of Bergen

⁵University of Oxford

We investigate panel conditioning effects in a long-running probability-based online panel of the general population through a large-scale experiment conducted in 2020. Our experiment was specifically designed to study the effect of intensifying the surveying frequency for the treatment group ($N = 5598$ panel members) during a 16-week corona study while keeping the control group ($N = 799$ panel members) at the usual bi-monthly surveying frequency. While many panel surveys have conducted add-on studies with increased surveying frequency during the COVID-19 pandemic, ours is among the very few which can provide experimental evidence regarding the potential impact of the add-on study on the underlying panel data infrastructure. Our results show that panel conditioning is only a minor issue when increasing surveying frequency and only matters for survey questions directly related to the corona study. It, therefore, provides reassurance to researchers who have implemented or are planning to implement similar special-topic add-on studies.

Keywords: online panel; panel conditioning; COVID-19; survey experiment; probability sample; panel data

1 Introduction

Panel surveys consist of repeated measurements of the same individuals' attitudes, behaviour, and living conditions through self-completion questionnaires or interviewer-mediated surveys (Kasprzyk et al., 1989). For this purpose, respondents are re-surveyed, typically at regular intervals, for example once a year or every other month (Lynn et al., 2019).

The panel survey design combines many attractive measurement qualities. In contrast to cross-sectional surveys, which only allow studying inter-personal differences, the panel survey design also allows studying intra-personal changes over time (Willson & Shuey, 2016). Furthermore, panel surveys can establish the temporal links between dependent and independent variables needed for detecting causal relationships (Hoffmann & Doblhammer, 2021). Finally, long-running panel surveys collect large amounts of data from the same individuals over time. This provides researchers with a large selection of potential control variables,

thus reducing the risk of unobserved heterogeneity and endogeneity in their analyses (Wooldridge, 2005).

However, panel surveys may have a limitation which results from the very characteristic that makes this form of data collection so attractive: the re-surveying process. Researchers fear that by frequently re-surveying the same individuals the respondents may alter their actual and/or reported behaviour and attitudes. This phenomenon is called panel conditioning (Das et al., 2011) and may have a lasting impact on panel data (Warren & Halpern-Manners, 2012). Thus, the mere act of re-surveying may distort the panel's findings (Sun et al., 2019). Moreover, any change in panel design, especially any increase in surveying frequency or questionnaire length and content, may increase panel conditioning effects.

In this study, we explore a specific type of potential panel conditioning. A large-scale experiment in an existing probability-based online panel during the early phase of the COVID-19 pandemic in Germany allows us to investigate whether increasing the surveying frequency by a factor of nine, from bi-monthly to additional weekly surveys, generates panel conditioning effects and whether these effects have a lasting impact across panel data collection waves. Our data stem from an experiment in which 7/8th of an established panel sample was randomly allocated to additional weekly

Contact information: German Institute for Economic Research (DIW), Mohrenstraße 58, 10117 Berlin, Germany (E-Mail: ccor- nesse@diw.de).

surveys for 16 consecutive weeks, whereas the other random 1/8th of the panel sample remained at the regular surveying frequency that respondents were already used to with surveys every other month.

2 Previous Research

Panel conditioning is the learning effect that participating in a panel study has on the panel participants (Struminskaya & Bosnjak, 2021). The learning effect can be advantageous for the data quality, for example panel participants learn to respond to the survey questions as intended by the researchers (e.g. no rounding in expenditure reports, Eckman and Bach, 2021). They can also be disadvantageous for the data quality, for example, if participants learn to avoid burdensome filter loops by providing less information (the so-called “motivated underreporting,” see Eckman et al., 2014). Moreover, the learning effect can affect panel participants’ actual attitudes and behaviour rather than, or in addition to, their response behaviour. For example, repeatedly being asked about a certain survey topic may encourage panel participants to think about this topic more and to gather information about it. This may be beneficial for the panel respondent personally, but it threatens inference from the panel data, because it makes the panel participants different from the population they are supposed to represent (e.g., Toepoel et al., 2009).

While it is universally acknowledged that panel conditioning effects may bias measurements in panel data collections, it remains unclear to what extent and under which conditions panel conditioning occurs. This may be due to the fact that little research conducted on this topic relies on experimental rather than observational (e.g., Müller & Schmiedeberg, 2020) or quasi-experimental studies (e.g., Binswanger et al., 2013).

In one of the rare experimental studies on panel conditioning, Silber et al. (2019) varied the frequency with which respondents were asked to provide information on their ego-centric social networks. Respondents received a social network assessment module with up to 49 questions either once or twice within little more than a month. Contrary to their expectations, the authors found that neither the reported network size and density nor the general data quality suffered from asking respondents to fill out the network module again. These findings suggest that conditioning is not as prevalent as some researchers might think, even when asking for relatively burdensome and time-consuming information. This is in line with a number of non-experimental studies that found either no, no consistent, or only very minor panel conditioning effects in a variety of research settings (e.g., Bailar, 1975; Pennell & Lepkowski, 1992; Sun et al., 2019).

Other experimental studies, however, found significant panel conditioning effects, although the expected direction of the effects remains unclear. For example, Halpern-Manners

et al. (2014) find that repeatedly asking respondents about deviant behaviour (drunk driving, minor theft) increases their likelihood of reporting such behaviour. This suggests that positive experiences with previous panel waves increase people’s willingness to provide sensitive information (i.e. potential advantageous panel conditioning effects, see also Kühne, 2018; Struminskaya, 2016; Uhrig, 2011). However, Torche et al. (2012), find the opposite: Earlier questions about substance use (alcohol, cigarettes, marijuana and cocaine) decreased the respondents’ likelihood to report substance use in later data collections (i.e. potential disadvantageous panel conditioning effects, see also Battaglia et al., 1996; Struminskaya, 2016; Yan & Eckman, 2012). In addition, the prevalence of conditioning effects across time remains unclear. Halpern-Manners et al. (2014) found conditioning effects only when the baseline and follow-up survey were one month apart but not when they were one year apart, suggesting that conditioning effects fade away with an increasing temporal distance between measurements. Torche et al. (2012), however, found conditioning effects even when the survey waves were one year apart, thus suggesting long-lasting conditioning effects.

The literature also remains unclear about which types of survey questions may be particularly affected or unaffected by panel conditioning. Some studies indicate that behavioural questions and, in particular, those which are directly linked to the main topic of the panel are most affected. Bach and Eckman (2019), for example, find that, over the course of an employment panel, unemployed persons increasingly report participating in governmental labour market programmes. By linking the respondents’ data to administrative employment records, the authors find that the reported increase is not just a change in reporting, but in actual behaviour. Furthermore, each additional exposure to the employment survey questions intensifies the effect on respondents’ labour market programme participation. Through an instrumental variable approach, the authors establish that this increase in programme participation is disproportionate in relation to comparable people not selected for the panel. This is in line with other research which suggests, for example, that participation in election studies stimulates respondents to become more interested and engaged in elections (e.g., Clausen, 1968; Traugott & Katosh, 1979; Yalch, 1976).

While most studies investigate panel conditioning effects on behavioural survey items, some studies indicate that attitudinal questions may be affected, too (for a theoretical framework see Bergmann & Barth, 2018). Sturgis et al. (2009) find support for their hypothesis that attitudes strengthen over the course of panel studies. This is in line with research by Waterton and Lievesley (1989), who find circumstantial evidence suggesting people get “politicized” over the course of the panel. However, both of these studies are observational and thus prohibit causal claims. In

addition, some studies suggest that neither attitudinal nor behavioural survey questions are affected by conditioning, while other question types, such as knowledge questions, show panel conditioning biases (e.g., Das et al., 2011). Furthermore, researchers have suggested that some perceived conditioning effects may be an artifact of differential non-response and attrition between comparison groups (e.g., Struminskaya, 2016; Wooden & Li, 2014)

In summary, the literature typically suspects panel conditioning effects but finds little consistent evidence. Most studies on panel conditioning are either not based on experiments and, thus, hamper causal inference, or they are conducted on a small scale or on nonprobability samples and thus lack external validity. As a consequence, few studies find panel conditioning effects and even fewer find effects consistently for the same types of variables, topics, or surveying frequency. Therefore, still little is known about how re-surveying respondents in a panel survey affects the study findings.

In this article, we report the results of a large-scale experiment on increasing the surveying frequency in a long-running probability-based online panel from bi-monthly to additional weekly surveys. It should be noted that we apply a broad definition of panel conditioning effects, where panel conditioning is any learning effect caused by a panel treatment which is observable in substantive survey responses. The random allocation of panelists to the treatment group of additional surveys or the control group without additional surveys allows us to examine the causal effect of the additional weekly surveys. Moreover, the underlying panel survey, which continued during the treatment and afterwards, allows us to examine a multitude of survey questions which may be affected by the treatment. In this way, we aim to contribute to answering the following research questions: Does increasing the surveying frequency in a panel study lead to conditioning effects? And if so, which types of survey questions are affected?

3 Data and Methods

At the outbreak of the COVID-19 pandemic in Germany, the German Internet Panel (GIP) added an additional high-frequency data collection to its usual bi-monthly surveys: The Mannheim Corona Study (MCS). The goal of the MCS was to study the effect of the pandemic and the accompanying containment measures on individuals' lives. It ran from 20th March to 10th July and surveyed GIP panel members by inviting them to complete weekly 10-minute questionnaires (see Blom et al., 2020; Cornesse, Krieger, et al., 2021). In a rotating panel design, the regular GIP sample of 6,397 panelists was divided into eight random sub-samples of approximately equal size.¹ Seven sub-samples were each assigned to a specific day of the week. On this day of the week, the MCS panelists received an invitation to the current MCS questionnaire. MCS panelists were given 48 hours to complete the

week's survey. They received invitations to the MCS for 16 consecutive weeks, in addition to their regular GIP invitations every other month. The seven GIP sub-samples that were invited to the MCS constitute our treatment sample ($N = 5,598$). The GIP and MCS data collections overlapped in March, May, and July (see Figure 1). The eighth GIP sub-sample did not receive any invitation to the MCS but continued with the standard GIP survey intervals. This control group ($N = 799$) allows us to study panel conditioning effects experimentally in a probability-based sample of the general population.²

3.1 German Internet Panel and Mannheim Corona Study

The GIP is a probability-based online panel of the general population in Germany aged 16–75 at the time of recruitment. The first GIP panel members were recruited in 2012 (see Blom et al., 2015) with additional rounds of recruitment in 2014 and 2018. During the first two rounds of recruitment, interviewers conducted initial face-to-face interviews and non-internet households were provided with the needed equipment and technical support to enable their participation in the GIP (see Blom et al., 2017; Cornesse & Schaurer, 2021; Herzing & Blom, 2019). In 2018, new panel members were recruited without interviewers and exclusively via postal mail (see Cornesse, Felderer, et al., 2021; Friedel et al., 2023). All panel members are invited to each survey wave. Consequently, by the time of the start of the MCS, even the most recent GIP sample members had been invited to ten regular panel waves (in addition to the recruitment surveys), and the earliest recruits had received 46 GIP survey invitations over the course of almost 8 years. All panel members were thus well-acquainted with the GIP procedures and surveying intervals.

The GIP is a multi-purpose panel with a focus on social, political, and economic questionnaire content. Approximately two thirds of its content is repeated at yearly intervals or more frequently, for example, for three consecutive waves. Thus, in contrast to some other probability-based online panels, the GIP is a longitudinal panel survey rather than a pool of respondents (see e.g. Blom et al., 2017, for some alternative approaches in Europe). In addition to general lon-

¹Note that all original panel members that had not actively de-registered were considered GIP panelists, independently of when they last participated in a GIP survey.

²Note that an additional 149 GIP panel members were not part of the experiment but continued with the standard bi-monthly surveying intervals, for survey practical reasons. Most of these panelists live in previously offline households and receive equipment and technical support to participate in the regular GIP surveys. Since some of these offline members do not have an email address, they could not be invited to the MCS at short notice. They were excluded from all analyses.

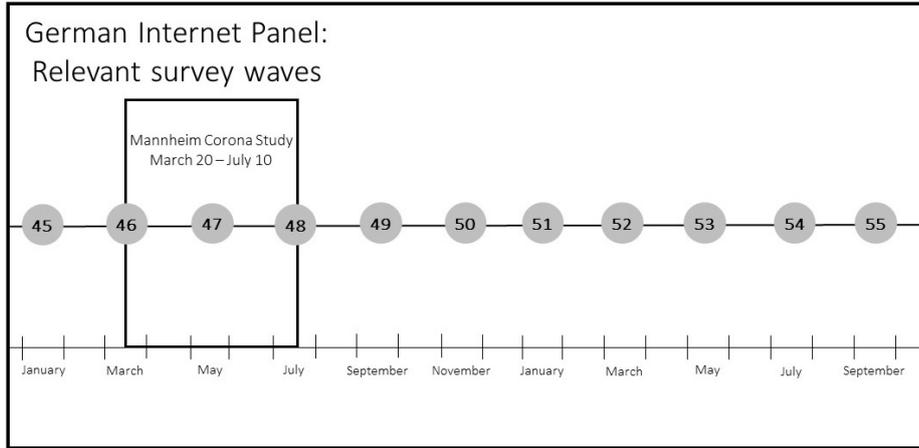


Figure 1

Schematic timeline of GIP survey waves in relation to the MCS fieldwork time

gitudinal survey items, respondents regularly receive questions that investigate their position regarding current events, such as election outcomes, terrorist attacks or natural disasters. Therefore, receiving questions about the COVID-19 pandemic would not have been surprising to GIP panelists. However, the MCS procedures with the increased surveying frequency and the short response window of 48 hours was new for the panelists.

The MCS was run by the GIP team together with a group of substantive researchers from the same research center. The key operational difference between the two studies was the considerably higher pace of the MCS. Instead of compiling questionnaires every two months, the MCS implemented a new questionnaire each week (see Cornesse et al., 2022, for more information on the MCS measurement instruments). Approximately two-thirds of each MCS questionnaire consisted of questions that were repeated each week or that were only slightly adjusted to account for social, economic, political or epidemic developments (e.g., Naumann et al., 2020). The final third of the questionnaire contained questions that were asked only once during the MCS fieldwork (though possibly repeated during the ensuing GIP waves), for example on the introduction of the official contact tracing app (e.g., Blom, Wenz, et al., 2021), or questions that were repeated only once or twice during the 16 weeks of MCS fieldwork (e.g., Mata et al., 2021). To ease the workload and to allow long-term comparisons with GIP data, the MCS adopted previous GIP items wherever possible (e.g., Möhring et al., 2021). The MCS covered topics related to people’s social and economic situation, political attitudes, behaviours and psychological health.

Table 1

Predicting assignment to the treatment

	Invited to MCS		Participated in MCS	
	Coef.	Std. Err.	Coef.	Std. Err.
Female	0.120	0.102	0.127	0.093
Age	0.00004	0.009	-0.002	0.009
Household Size (ref. categ.: 1)				
2	-11.734	308.952	-11.970	309.065
3 or more	-11.568	308.952	-11.799	309.065
Marital Status: (ref. categ.: Married)				
Other	-0.231	0.192	-0.133	0.181
Single	0.180	0.134	0.097	0.120
Education (ref. categ.: Low)				
Medium	-0.093	0.160	-0.068	0.143
High	-0.226	0.164	-0.082	0.148
Citizenship	-0.041	0.327	-0.172	0.315
Constant	13.707	308.952	13.759	309.065

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

3.2 Random Allocation to Treatment and Control Group

The allocation of the 6,397 GIP panelists to one of the MCS samples or the control group was strictly random. We investigated the even allocation of our sample to treatment and control group regarding key socio-demographic characteristics. Neither gender, age, household size, marital status, education, nor citizenship predicts the assignment of GIP panelists to the treatment or control group. Further-

more, comparing the socio-demographic characteristics of those who responded to the MCS invitation in the treatment group to the control group also shows no significant differences (see Table 1).

3.3 GIP Survey Participation

In addition to checking for potential socio-demographic differences, we also examined whether the treatment and control group differed in terms of GIP survey participation. This could either be the case already before the MCS (i.e., as a result of the allocation procedure) or during and after the MCS (i.e., as a potential attrition effect of the MCS). Each of these could potentially impact our conclusions regarding conditioning effects in the GIP. For Figure 2, we fitted bivariate logistic regression models with participation in GIP survey waves as the outcome variable and assignment to the MCS treatment versus control group as the predictor variable. We run these analyses for the three GIP waves before the start of the MCS (waves 43 to 45), the three GIP waves during the MCS (waves 46 to 48) and all available GIP waves after the MCS (waves 49 to 55).

As Figure 2 shows, we find significant differences between the MCS treatment and control group only for GIP wave 46, which coincided with the start of the MCS, and waves 52 and 53, which were conducted more than half a year after the MCS. At each of these three waves, the treatment group is less likely to participate than the control group. We find no significant differences in response propensities before the MCS. Moreover, examining a fixed-effects model for participation across GIP data collection waves (which ignores the inconsistent participation pattern by experimental group established in Figure 2) shows that GIP survey participation generally declines across the observed data collection waves (see Table A1 in the Appendix). This is typical for panel studies and therefore unsurprising. The model also shows systematic differences in participation by socio-demographic characteristics (gender, marital status, education, citizenship).

3.4 Hypotheses

Our unique research setting allows us to assess the impact of increasing the surveying frequency in a panel study with regard to conditioning effects. Unlike many other studies, we can exploit our experimental design, in which some members of the GIP sample were exposed to the increased surveying frequency while others were not. Since the literature on conditioning effects is unclear about whether conditioning effects will occur at all, we first test the following general hypothesis:

H1 Increasing the surveying frequency for the treatment group results in differences in their responses to GIP survey questions as compared to the control group.

If we find any indication of conditioning effects in our experiment, the next question is whether it affects some survey questions more than others. Some of the literature suggests that conditioning effects occur on the dominant topic of the surveys respondents were exposed to. For example, respondents may change their reporting of substance use (Halpern-Manners et al., 2014) or employment programme participation (Bach & Eckman, 2019) if previously asked about these topics. The dominant topic of the MCS is COVID-19 and its impact on all domains of the respondents' lives. We, therefore, test the following hypothesis:

H2 Differences in responses mainly occur on survey questions that are linked to the content of the MCS.

Some of the literature suggests that panel conditioning effects may fade away over time. For example, conditioning effects may be measurable if survey questions are repeated within little more than a month, but not if they are repeated after a year (Halpern-Manners et al., 2014). This is in line with the general evidence on treatment effects in survey experiments, which have often been shown to diminish over time (e.g., Gaines et al., 2007). We, therefore, test the following hypothesis:

H3 Differences in responses mainly occur during or directly after the MCS fieldwork period (i.e. GIP waves 46 to 49).

To get a full picture of when and on which survey questions conditioning effects occur, we need to test a larger number of GIP questions for potential differences between the treatment and control group. Testing many variables for the same hypothesis, however, means that the conducted tests are not independent from one another. If this fact is ignored in the analyses, we may find evidence in support of a hypothesis simply because we tested so many items that, just by chance, we find some statistically significant effects (Savin, 1980). Indeed, not accounting for this multiple-comparison problem may be responsible for some of the rather inconsistent and hard to interpret findings on panel conditioning in the literature. We therefore test the following hypothesis:

H4 Differences in responses between treatment and control group are an artifact of multiple comparison.

3.5 Variable Selection and Analysis Strategy

Panel conditioning effects due to the additional weekly MCS surveys may occur in all GIP survey waves which were implemented either during the 16 weeks of MCS data collection (i.e. GIP waves 46 through 48; see Figure 1) or afterwards (i.e. 49 and after). At the time of our analyses, GIP data are available until wave 55, which means that we can examine 9 data collection waves, spanning a

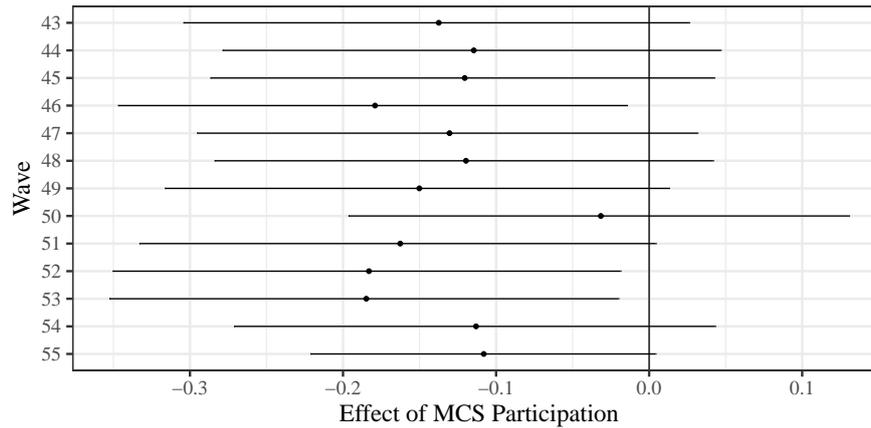


Figure 2

Effects of MCS Participation on GIP Wave Participation

time of 1.5 years since the start of the MCS fieldwork (see <https://paneldata.org/gip/> for documentation of all GIP survey questions by data collection wave). We test all survey questions for potential conditioning effects, except for panel administrative questions (e.g. phone number verification), socio-demographic questions (e.g. age, gender, education), and questions that were not well-suited for our analyses for other reasons, for example because they contained their own experimental splits (e.g. for question wording experiments). Furthermore, we excluded some large item batteries on perceptions of political parties (e.g. left-right placement of each party in the German parliament), because we did not think they would meaningfully contribute to the analyses and only lead to an unnecessary amount of spurious effects. All in all, we tested 86 question items for differences between the treatment and control group. 68 of them were implemented once in the GIP questionnaires during the time span which we examine, 12 items were implemented twice, 4 items three times, and two items four times. Overall, this results in 112 tests for differences between treatment and control group across items and survey waves (see Table A2 in the appendix).

51 tests for differences between treatment and control group were conducted on items that relate to the content of the MCS. 22 of those tests explore differences in questions which explicitly mention the COVID-19 pandemic. This includes, for example, people's fear of the SARS-COV-2 virus, how they weigh the economic damages of the enforced COVID-19 counter-measures against their societal benefits, or whether they downloaded the official COVID-19 contact tracing app. In addition, 29 tests were conducted on items that do not specifically mention COVID-19, but that were replicated, sometimes with small adjustments, from the MCS questionnaires. This includes, for example, a questionnaire module on health behaviours (e.g., snacking, physical activ-

ity) and mental health (e.g. feelings of depression and loneliness) which was included in the MCS and replicated in GIP wave 51. 61 tests were conducted on items unrelated to the MCS, which includes questions on topics such as climate change and data protection.

For each question in our analysis, we examine potential conditioning effects by running regression models with the survey responses as the outcome variables and the treatment versus control group as the central predictor variable.³ Generally, all question items on which we find a significant difference in responses on a 95%-confidence level are considered to be potentially affected by conditioning. To examine the robustness of the findings, we run all models with and without accounting for multiple-comparison testing via Bonferroni corrections (see e.g., Haynes, 2013). For the Bonferroni correction, we divide the alpha-error of the initial estimates by the number of tests in the main comparison group it belongs to (MCS-related vs. MCS-unrelated tests). We also run all models with and without controlling for the socio-demographic characteristics related to differential GIP participation over time (see Table A1 in the Appendix).

4 Results

When accounting for multiple-comparison testing and controlling for socio-demographics, only 5 out of the 112 tests performed for differences between the treatment and

³Note that in the results section, we define the treatment group as people who were both invited to the MCS and also participated in it at least once and the control group as anyone who was never invited to the MCS given that they could have been invited. For robustness analyses where we define the treatment group as anyone who was invited to the MCS, regardless of whether they ever participated in it (see Figures B1 and B2 in the appendix).

control group (i.e. 4% of all tests) suggest an impact of the MCS on responses to GIP survey questions (see Figure 3).⁴

Four of the significant conditioning effects are directly related to the dominant topic of the MCS: Feeling personally threatened by COVID-19, general concern about COVID-19, perceived likelihood of contracting Sars-COV-2, and support for increasing the federal states' competencies in a pandemic. While robust for Bonferroni correction, these effects point into different directions as to how the MCS has impacted the respondents: MCS participants consider COVID-19 as less of a threat to themselves (-0.81 points on 11-point scale) and rate their likelihood of contracting the virus as lower than the control group (-0.31 points on 7-point scale). However, they also report greater general concern about COVID-19 (0.22 points on 4-point scale) and higher support for increasing the German federal states' decision-making competencies during a pandemic (0.47 points on 11-point scale). In addition to these four items that explicitly mention COVID-19 in the question text, one other conditioning effect transpires. It is on feelings of depression replicated from a questionnaire module on mental and physical health implemented in the MCS. The result suggests that MCS participants feel less depressed than the control group (-0.18 points on a 4-point scale). Four out of five robust conditioning effects concern question items measured within six months of the end of the MCS, but none of them are from GIP surveys conducted during (waves 46 to 48) or directly after (wave 49) the MCS. One robust finding is from GIP wave 55, which was conducted more than a year after the end of the MCS.

Since Bonferroni corrections are sometimes judged as too conservative, we additionally examine the results of uncorrected tests as a robustness analysis: only 21 (i.e. 19%) of all tests initially indicate a potential impact of the MCS on responses to GIP survey questions when neither accounting for multiple-comparison testing nor controlling for socio-demographics (see Figure 4).

Among the question items explicitly mentioning COVID-19, 8 tests out of 22 (i.e. 36%) showed significant differences between the treatment and control group. Three of those tests were on repeated measures of the same item (weighing economic damages against societal benefits of COVID-19 counter-measures) and showed significant differences at all three GIP waves in which the item was implemented (waves 50, 52, and 54). The other 5 items (personal COVID threat, general COVID concern, perceived probability of infection, support of increasing state competencies in case of pandemic, and COVID-19 contact tracing app usage) were only implemented once, so it is not possible to say whether effects would continue to surface if repeated.

Among the question items that do not directly mention COVID-19 but that were MCS replications, 6 out of 29 tests (i.e. 21%) initially showed significant differences between treatment and control group. One of those questions (govern-

ment responsibility for public healthcare) was implemented twice since the MCS (waves 51 and 55), but only showed an effect at the first measurement time point, indicating a potential fade-away effect of conditioning after the MCS. All other tests with significant differences were on items only measured once. Three of them were part of an MCS questionnaire module on physical and mental health during the pandemic (i.e., engaging in medium levels of physical activity, snacking, feeling depressed). Two others relate to other MCS topics relevant to the pandemic debate in Germany at the time (working from home legislation, tax evasion).

Among the question items unrelated to COVID-19 and the MCS, 7 out of 61 tests (11%) initially showed significant differences between the treatment and control group. One of those question items (government responsibility for the unemployed) was implemented twice since the MCS (waves 51 and 55) in the same questionnaire module as the question on government responsibility for healthcare replicated from the MCS. Here again, significant differences only occur at the first measurement time point, slightly indicating potential fade-away effects. All other tests with significant differences were on items only measured once (data privacy concerns, data sharing attitudes, opinions on European integration, trust in local government, trust in European Parliament, trust in online social media). Robustness analyses do not indicate that the conditioning effects would be stronger for the newer GIP sample recruited in 2018 than the older (more "trained") samples from 2012 and 2014 (see Figure B4 in the appendix).

The findings generally indicate that conditioning effects occur, but only on a small scale and that more survey questions relating to the dominant topic of the add-on study may be affected than unrelated question items. We find no clear evidence for a fade-away effect of conditioning, since 15 out of 18 tested repeated measures question items (83%) never showed any significant differences in responses. Moreover, the question item on economic harms versus societal benefits of COVID-19 counter-measures displayed significant differences of approximately the same size at all its measurement time points. Most importantly, the vast majority of the effects is likely an artifact of multiple-comparison testing, as indicated by the results from the Bonferroni-corrected tests.

5 Conclusion and Discussion

This article contributes to answering the following research questions: Does increasing the surveying frequency in a panel study lead to conditioning effects? And if so, which types of survey questions are affected? In the study, we apply a broad definition of panel conditioning effects, where

⁴Note that with Bonferroni correction but without socio-demographic controls, only 4 of the initially statistically significant potential conditioning effects remain significant, as can be seen in Figure B3 in the Appendix.

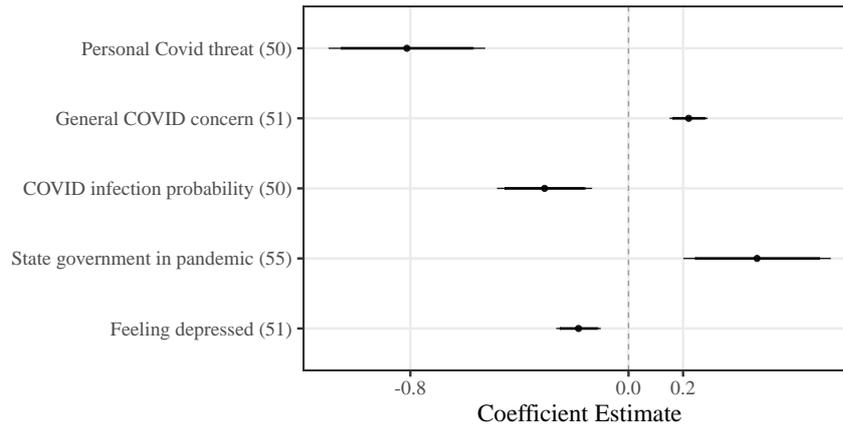


Figure 3

Significant conditioning effects with Bonferroni correction and inclusion of demographic controls (GIP survey wave of the tested item in parentheses)

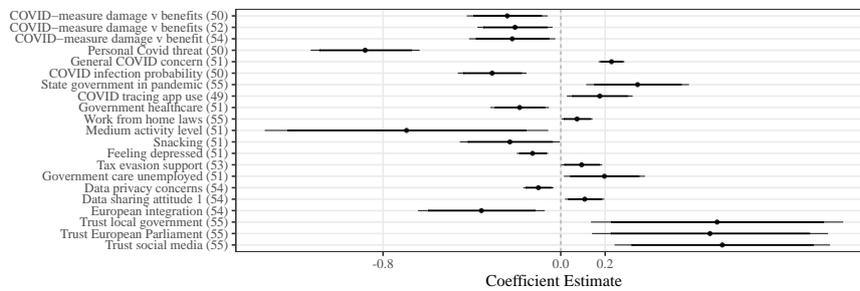


Figure 4

Initially significant conditioning effects (GIP survey wave of the tested item in parentheses)

panel conditioning is any learning effect caused by a panel treatment and which is observable in substantive survey responses. From the literature, we derived 4 hypotheses relating to these questions and examined whether or not we find support for them in our experimental study.

Our experiment consisted of increasing the surveying frequency of a long-running probability-based online panel (i.e., the GIP), which had only ever exposed its panel members to bi-monthly data collection waves, by a factor of 9 during the early phase of the pandemic in Germany for 7/8th of the panel (i.e. the MCS, $N = 5,598$). The remaining randomly selected 1/8th of the panel sample ($N = 799$) was never exposed to the fast-frequency pandemic-related add-on study. By comparing survey responses of the treatment and control group across the regular GIP survey waves during and after the MCS, we could assess whether participating in the MCS had an impact on panel members’ reported feelings, attitudes, and behaviour. In the following, we will address and discuss our findings in relation to our hypotheses.

H1 Increasing the surveying frequency for the treatment group results in differences in their responses to GIP survey questions as compared to the control group.

Overall, we find marginal support for this hypothesis. Not accounting for multiple comparison testing, we find differences in survey responses between the groups in 21 out of 112 conducted tests. With Bonferroni correction and socio-demographic controls, only 5 robust effects remain, suggesting that conditioning effects of increasing surveying frequency are only a minor issue.

H2 Differences in responses mainly occur on survey questions that are linked to the content of the MCS.

4 out of 5 robust conditioning effects found in our study explicitly mention the COVID-19 pandemic in the question text. The only other robust effect relates to a questionnaire module on mental and physical health, which was implemented in one of the MCS questionnaires. This suggests that

question items unrelated to the dominant topic of the add-on study are unaffected by conditioning, while question items related to the dominant topic may be marginally affected.

H3 Differences in responses mainly occur during or directly after the MCS fieldwork period (i.e. GIP waves 46 to 49).

We find no support for this hypothesis. In the initial analyses without accounting for multiple-comparison testing and controlling for socio-demographic characteristics, we detect only one potential conditioning effect in this time span: COVID-19 tracing app use in GIP wave 49. The other 20 tests with significant differences by group concern later survey waves. Moreover, the five robust effects were nearly all on question items implemented in GIP waves 50 and 51. This is still within half a year of the end of the MCS, so it may be worthwhile to explore this hypothesis again in a study with more repeated measures instruments.

H4 Differences in responses between treatment and control group are an artifact of multiple comparison testing.

We find support for this hypothesis, as only 5 of the 21 initially statistically significant conditioning effects remained robust when adding Bonferroni correction factors. We suspect that at least some of the relatively inconsistent findings in the literature may be an artifact of multiple testing, just as most of the results from our initial tests (Sedgwick, 2014). Moreover, we suspect that more studies like ours, which look into conditioning effects of changing the panel protocol, may have found no or no compelling evidence of conditioning effects in the data. This may also in part be responsible for the quite incoherent picture of whether and under which conditions conditioning effects occur in the existing literature. Publishing null (or marginal) findings is often difficult (Franco et al., 2014). Future research should therefore look into identifying spurious correlations and closing the potential publication bias in the conditioning literature.

In terms of survey practical implications, our study provides reassurance to researchers and survey companies which, like us, increased the surveying frequency of their existing panel study to feature additional survey content during the pandemic (for examples see Burton et al., 2020; Kapteyn et al., 2020; Kühne et al., 2020). The negative impact of doing this seems to be marginal at most, at least in panel infrastructures similar to the GIP. The results seem encouraging for conducting similar add-on studies in the future. However, our study has important limitations in terms of generalizability. First, the MCS was conducted in a very special situation: A pandemic which affected everyone and confined people to their homes where they had the time and willingness to provide additional data. Conditioning effects may have been even weaker in an add-on study on a topic which

affects people only indirectly (e.g. government dissolution) or which keeps many people away from surveys or the internet (e.g. natural disasters). Second, the MCS was implemented in an existing panel study where all participants may already have generally altered their behaviour and/or attitudes as a result of being in the study. Conditioning effects may have been stronger in a fresh panel sample (although our sensitivity analyses by recruitment year displayed in Figure B4 do not suggest this). We may also have found more conditioning effects of increasing the surveying frequency if we had replicated more pre-pandemic GIP items frequently during and after the MCS. A true experiment on the impact of increasing the surveying frequency on panel conditioning would have required this. Third, the GIP is a relatively high frequency panel anyway. So the impact of increasing the surveying frequency may have been stronger in a panel with less frequent data collection because the intervention would have been more dramatic for the panel participants. Last, the GIP and MCS collect data exclusively online. The impact of an interviewer-administered add-on study may have been stronger due to the potentially even higher salience of the study topic in people's minds after discussing it verbally with another person. All of these limitations of our study warrant future research. We, therefore, encourage researchers planning to conduct panel add-on studies on the urgent topics of the future to include and report the findings of similar experiments.

References

- Bach, R. L., & Eckman, S. (2019). Participating in a panel survey changes respondents' labour market behaviour. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 182(1), 263–281. <https://doi.org/10.1111/rssa.12367>
- Bailar, B. A. (1975). The effects of rotation group bias on estimates from panel surveys. *Journal of the American Statistical Association*, 70(349), 23–30. <https://doi.org/10.1080/01621459.1975.10480255>
- Battaglia, M. P., Zell, E. R., & Ching, P. L. (1996). Can participating in a panel sample introduce bias into trend estimates? *Proceedings of the Survey Research Methods Section*, 1010–1013. http://www.asarms.org/Proceedings/papers/1996_175.pdf
- Bergmann, M., & Barth, A. (2018). What was i thinking? A theoretical framework for analysing panel conditioning in attitudes and (response) behaviour. *International Journal of Social Research Methodology*, 21(3), 333–345. <https://doi.org/10.1080/13645579.2017.1399622>
- Binswanger, J., Schunk, D., & Toepoel, V. (2013). Panel conditioning in difficult attitudinal questions. *Public Opinion Quarterly*, 77(3), 783–797. <https://doi.org/10.1093/poq/nft030>

- Blom, A. G., Cornesse, C., Friedel, S., Krieger, U., Fikel, M., Rettig, T., Wenz, A., Axenfeld, J., Möhring, K., Naumann, E., Reifenscheid, M., Lehrer, R., & Juhl, S. (2020). High frequency and high quality survey data collection. *Survey Research Methods*, 14(2), 171–178. <https://doi.org/10.18148/SRM/2020.V14I2.7735>
- Blom, A. G., Fikel, M., Friedel, S., Krieger, U., Rettig, T., & SFB 884 ‘Political Economy Of Reforms’, Universität Mannheim. (2021a). German Internet Panel, Wave 46 (March 2020) [GESIS Data Archive]. <https://doi.org/10.4232/1.13679>
- Blom, A. G., Fikel, M., Friedel, S., Krieger, U., Rettig, T., & SFB 884 ‘Political Economy Of Reforms’, Universität Mannheim. (2021b). German Internet Panel, Wave 48 (July 2020) [GESIS Data Archive]. <https://doi.org/10.4232/1.13681>
- Blom, A. G., Fikel, M., Friedel, S., Krieger, U., Rettig, T., & SFB 884 ‘Political Economy Of Reforms’, Universität Mannheim. (2021c). German Internet Panel, Wave 49 - Core Study (September 2020) [GESIS Data Archive]. <https://doi.org/10.4232/1.13682>
- Blom, A. G., Fikel, M., Gonzalez Ocanto, M., Krieger, U., Rettig, T., & SFB 884 ‘Political Economy Of Reforms’, Universität Mannheim. (2021a). German Internet Panel, Wave 53 (May 2021) [GESIS Data Archive]. <https://doi.org/10.4232/1.13834>
- Blom, A. G., Fikel, M., Gonzalez Ocanto, M., Krieger, U., Rettig, T., & SFB 884 ‘Political Economy Of Reforms’, Universität Mannheim. (2021b). German Internet Panel, Wave 54 (July 2021) [GESIS Data Archive]. <https://doi.org/10.4232/1.13835>
- Blom, A. G., Gathmann, C., & Krieger, U. (2015). Setting up an online panel representative of the general population. *Field Methods*, 27(4), 391–408. <https://doi.org/10.1177/1525822x15574494>
- Blom, A. G., Gonzalez Ocanto, M., Fikel, M., Krieger, U., Rettig, T., & SFB 884 ‘Political Economy Of Reforms’, Universität Mannheim. (2021a). German Internet Panel, Wave 47 (May 2020) [GESIS Data Archive]. <https://doi.org/10.4232/1.13680>
- Blom, A. G., Gonzalez Ocanto, M., Fikel, M., Krieger, U., Rettig, T., & SFB 884 ‘Political Economy Of Reforms’, Universität Mannheim. (2021b). German Internet Panel, Wave 50 (November 2020) [GESIS Data Archive]. <https://doi.org/10.4232/1.13734>
- Blom, A. G., Gonzalez Ocanto, M., Fikel, M., Krieger, U., Rettig, T., & SFB 884 ‘Political Economy Of Reforms’, Universität Mannheim. (2021c). German Internet Panel, Wave 51 (January 2021) [GESIS Data Archive]. <https://doi.org/10.4232/1.13833>
- Blom, A. G., Gonzalez Ocanto, M., Krieger, U., Rettig, T., Ungefucht, M., & SFB 884 ‘Political Economy Of Reforms’, Universität Mannheim. (2022). German Internet Panel, Wave 55 - Core Study (September 2021) [GESIS Data Archive]. <https://doi.org/10.4232/1.13874>
- Blom, A. G., Herzing, J. M. E., Cornesse, C., Sakshaug, J. W., Krieger, U., & Bossert, D. (2017). Does the recruitment of offline households increase the sample representativeness of probability-based online panels? Evidence from the German Internet Panel. *Social Science Computer Review*, 35(4), 498–520. <https://doi.org/10.1177/0894439316651584>
- Blom, A. G., Wenz, A., Cornesse, C., Rettig, T., Fikel, M., Friedel, S., Möhring, K., Naumann, E., Reifenscheid, M., & Krieger, U. (2021). Barriers to the large-scale adoption of a COVID-19 contact tracing app in germany: Survey study. *Journal of Medical Internet Research*, 23(3), e23362. <https://doi.org/10.2196/23362>
- Burton, J., Lynn, P., & Benzeval, M. (2020). How understanding society: The UK household longitudinal study adapted to the COVID-19 pandemic. *Survey Research Methods*, 14(2), 235–239. <https://doi.org/10.18148/SRM/2020.V14I2.7746>
- Clausen, A. R. (1968). Response validity: Vote report. *Public Opinion Quarterly*, 32(4), 588–606. <https://doi.org/10.1086/267648>
- Cornesse, C., Felderer, B., Fikel, M., Krieger, U., & Blom, A. G. (2021). Recruiting a probability-based online panel via postal mail: Experimental evidence. *Social Science Computer Review*, 40(5), 1259–1284. <https://doi.org/10.1177/08944393211006059>
- Cornesse, C., Krieger, U., Sohnius, M.-L., Fikel, M., Friedel, S., Rettig, T., Wenz, A., Juhl, S., Lehrer, R., Möhring, K., Naumann, E., Reifenscheid, M., & Blom, A. G. (2021). From German Internet Panel to Mannheim Corona Study: Adaptable probability-based online panel infrastructures during the pandemic. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 185(3), 773–797. <https://doi.org/10.1111/rssa.12749>
- Cornesse, C., Ocanto, M. G., Fikel, M., Friedel, S., Krieger, U., Rettig, T., & Blom, A. G. (2022). Measurement instruments for fast and frequent data collection during the early phase of COVID-19 in germany: Reflections on the Mannheim Corona Study. *Measurement Instruments for the Social Sciences*, 4(1), 2. <https://doi.org/10.1186/s42409-022-00030-5>
- Cornesse, C., & Schaurer, I. (2021). The long-term impact of different offline population inclusion strategies in probability-based online panels: Evidence from the German Internet Panel and the GESIS panel. *Social Science Computer Review*, 39(4), 687–704. <https://doi.org/10.1177/0894439320984131>

- Das, M., Toepoel, V., & van Soest, A. (2011). Nonparametric tests of panel conditioning and attrition bias in panel surveys. *Sociological Methods & Research*, 40(1), 32–56. <https://doi.org/10.1177/0049124110390765>
- Eckman, S., & Bach, R. (2021). Panel conditioning in the u.s. consumer expenditure survey. *Journal of Official Statistics*, 37(1), 53–69. <https://doi.org/10.2478/jos-2021-0003>
- Eckman, S., Kreuter, F., Kirchner, A., Jäckle, A., Tourangeau, R., & Presser, S. (2014). Assessing the mechanisms of misreporting to filter questions in surveys. *Public Opinion Quarterly*, 78(3), 721–733. <https://doi.org/10.1093/poq/nfu030>
- Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203), 1502–1505. <https://doi.org/10.1126/science.1255484>
- Friedel, S., Felderer, B., Krieger, U., Cornesse, C., & Blom, A. G. (2023). The early bird catches the worm! Setting a deadline for online panel recruitment incentives. *Social Science Computer Review*, 41(2), 370–389.
- Gaines, B. J., Kuklinski, J. H., & Quirk, P. J. (2007). The logic of the survey experiment reexamined. *Political Analysis*, 15(1), 1–20. <https://doi.org/10.1093/pan/mpl008>
- Halpern-Manners, A., Warren, J. R., & Torche, F. (2014). Panel conditioning in a longitudinal study of illicit behaviors. *Public Opinion Quarterly*, 78(3), 565–590. <https://doi.org/10.1093/poq/nfu029>
- Haynes, W. (2013). Bonferroni correction. In W. Dubitzky, O. Wolkenhauer, K.-H. Cho, & H. Yokota (Eds.), *Encyclopedia of systems biology* (pp. 154–154). Springer. https://doi.org/10.1007/978-1-4419-9863-7_1213
- Herzing, J. M. E., & Blom, A. G. (2019). The influence of a person's digital affinity on unit nonresponse and attrition in an online panel. *Social Science Computer Review*, 37(3), 404–424. <https://doi.org/10.1177/0894439318774758>
- Hoffmann, R., & Doblhammer, G. (2021). Approaches and methods for causal analysis of panel data in the area of morbidity and mortality. *Comparative Population Studies*, 46. <https://doi.org/10.12765/cpos-2021-03>
- Kapteyn, A., Angrisani, M., Bennett, D., Bruine De Bruin, W., Darling, J., Gutsche, T., Liu, Y., Meijer, E., Perez-Arce, F., Schaner, S., Thomas, K., & Weerman, B. (2020). Tracking the effect of the COVID-19 pandemic on the lives of american households. *Survey Research Methods*, 14(20), 179–186. <https://doi.org/10.18148/SRM/2020.V14I2.7737>
- Kasprzyk, D., Singh, M. P., Duncan, G., & Kalton, G. (Eds.). (1989). *Panel surveys*. Wiley.
- Kühne, S. (2018). From strangers to acquaintances? Interviewer continuity and socially desirable responses in panel surveys. *Survey Research Methods*, 12(2), 121–146. <https://doi.org/10.18148/SRM/2018.V12I2.7299>
- Kühne, S., Kroh, M., Liebig, S., & Zinn, S. (2020). The need for household panel surveys in times of crisis: The case of SOEP-CoV. *Survey Research Methods*, 14(2), 195–203. <https://doi.org/10.18148/SRM/2020.V14I2.7748>
- Lynn, P., Couper, M., & Watson, N. (2019). Longitudinal surveys – unique opportunities and unique methodological challenges. *Longitudinal and Life Course Studies*, 10(4), 415–420. <https://doi.org/10.1332/175795919x15683588414527>
- Mata, J., Wenz, A., Rettig, T., Reifenscheid, M., Möhring, K., Krieger, U., Friedel, S., Fikel, M., Cornesse, C., Blom, A. G., & Naumann, E. (2021). Health behaviors and mental health during the COVID-19 pandemic: A longitudinal population-based survey in Germany. *Social Science & Medicine*, 287, 114333. <https://doi.org/10.1016/j.socscimed.2021.114333>
- Möhring, K., Naumann, E., Reifenscheid, M., Wenz, A., Rettig, T., Krieger, U., Friedel, S., Finkel, M., Cornesse, C., & Blom, A. G. (2021). The COVID-19 pandemic and subjective well-being: Longitudinal evidence on satisfaction with work and family. *European Societies*, 23(sup1), S601–S617. <https://doi.org/10.1080/14616696.2020.1833066>
- Müller, B., & Schmiedeberg, C. (2020). Do respondents get used to answering sensitive questions? *Public Opinion Quarterly*, 84(3), 654–674. <https://doi.org/10.1093/poq/nfaa041>
- Naumann, E., Möhring, K., Reifenscheid, M., Wenz, A., Rettig, T., Lehrer, R., Krieger, U., Juhl, S., Friedel, S., Fikel, M., Cornesse, C., & Blom, A. G. (2020). COVID-19 policies in Germany and their social, political, and psychological consequences. *European Policy Analysis*, 6(2), 191–202. <https://doi.org/10.1002/epa2.1091>
- Pennell, S. G., & Lepkowski, J. M. (1992). Panel conditioning effects in the survey of income and program participation. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 566–571.
- Savin, N. E. (1980). The Bonferroni and the Scheffe multiple comparison procedures. *The Review of Economic Studies*, 47(1), 255. <https://doi.org/10.2307/2297112>

- Sedgwick, P. (2014). Multiple hypothesis testing and bonferroni's correction. *BMJ*, *349*(3), g6284–g6284. <https://doi.org/10.1136/bmj.g6284>
- Silber, H., Schröder, J., Struminskaya, B., Stocké, V., & Bosnjak, M. (2019). Does panel conditioning affect data quality in ego-centered social network questions? *Social Networks*, *56*, 45–54. <https://doi.org/10.1016/j.socnet.2018.08.003>
- Struminskaya, B. (2016). Respondent conditioning in online panel surveys. *Social Science Computer Review*, *34*(1), 95–115. <https://doi.org/10.1177/0894439315574022>
- Struminskaya, B., & Bosnjak, M. (2021). Panel conditioning: Types, causes, and empirical evidence of what we know so far. In P. Lynn (Ed.), *Advances in longitudinal survey methodology* (pp. 272–301). Wiley. <https://doi.org/10.1002/9781119376965.ch12>
- Sturgis, P., Allum, N., & Brunton-Smith, I. (2009). Attitudes over time: The psychology of panel conditioning. In *Methodology of longitudinal surveys* (pp. 113–126). Wiley.
- Sun, H., Tourangeau, R., & Presser, S. (2019). Panel effects: Do the reports of panel respondents get better or worse over time? *Journal of Survey Statistics and Methodology*, *7*(4), 572–588. <https://doi.org/10.1093/jssam/smy021>
- Toepoel, V., Das, M., & Van Soest, A. (2009). Relating question type to panel conditioning: Comparing trained and fresh respondents. *Survey Research Methods*, *3*(2), 73–80. <https://doi.org/10.18148/SRM/2009.V3I2.874>
- Torche, F., Warren, J. R., Halpern-Manners, A., & Valenzuela, E. (2012). Panel conditioning in a longitudinal study of adolescents' substance use: Evidence from an experiment. *Social Forces*, *90*(3), 891–918. <https://doi.org/10.1093/sf/sor006>
- Traugott, M. W., & Katosh, J. P. (1979). Response validity in surveys of voting behavior. *Public Opinion Quarterly*, *43*(3), 359. <https://doi.org/10.1086/268527>
- Uhrig, S. C. N. (2011). Understanding panel conditioning: An examination of social desirability bias in self-reported height and weight in panel surveys using experimental data. *Longitudinal and Life Course Studies*, *3*(1), 120–36. <https://doi.org/10.14301/lcs.v3i1.164>
- Warren, J. R., & Halpern-Manners, A. (2012). Panel conditioning in longitudinal social science surveys. *Sociological Methods & Research*, *41*(4), 491–534. <https://doi.org/10.1177/0049124112460374>
- Waterton, J., & Lievesley, D. (1989). Evidence of conditioning effects in the British Attitudes Panel Survey. In D. Kasprzyk, G. J. Duncan, G. Kalton, & M. P. Singh (Eds.), *Panel surveys*. Wiley.
- Willson, A. E., & Shuey, K. M. (2016). Life course pathways of economic hardship and mobility and midlife trajectories of health. *Journal of Health and Social Behavior*, *57*(3), 407–422. <https://doi.org/10.1177/0022146516660345>
- Wooden, M., & Li, N. (2014). Panel conditioning and subjective well-being. *Social Indicators Research*, *117*(1), 235–255. <https://doi.org/10.1007/s11205-013-0348-1>
- Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, *20*(1), 39–54. <https://doi.org/10.1002/jae.770>
- Yalch, R. F. (1976). Pre-election interview effects on voter turnout. *Public Opinion Quarterly*, *40*(3), 331. <https://doi.org/10.1086/268309>
- Yan, T., & Eckman, S. (2012). Panel conditioning: Change in true value vs. change in self-report. *Proceedings of the Survey Research Methods Section*, 4726–36. http://www.asasrms.org/Proceedings/y2012/Files/306203_76099.pdf

Appendix A
Tables

Table A1*Predicting participation in the German Internet Panel*

	Participation in GIP	
	Fixed effects	
	Coef.	Std. Err.
MCS	-0.068*	0.038
Female	-0.025	0.025
Age	0.006***	0.002
Household Size: (ref=1)		
2	-11.949	65.420
3 or more	-12.134	65.420
Marital Status: (ref=Married)		
Other	-0.066	0.055
Single	-0.453***	0.029
Education: (ref=Low)		
Medium	-0.131***	0.039
High	-0.145***	0.040
German Citizenship	0.390***	0.071
Wave: (ref=46)		
47	-0.168***	0.062
48	-0.186***	0.062
49	-0.143**	0.062
50	-0.098	0.063
51	-0.084	0.063
52	-0.202***	0.062
53	-0.234***	0.061
54	-0.673***	0.058
55	-0.641***	0.052
Constant	14.061	65.420

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table A2

Overview of all tested items

Study Wave	45	46	47	48	49	50	51	52	53	54	55
German Internet Panel											
Mannheim Corona Study											
Items											
Economic Damage Greater than Social Benefit											
Personal Covid Threat											
Personal Covid Concern											
Nervousness											
Government’s Responsibility Healthcare											
Government Expenses Healthcare											
Perceived Personal Health											
Government’s Responsibility Healthcare											
Regulation Home Office											
Satisfaction with Government’s Covid Response											
Covid Infection Probability											
More Competencies Federal Government in Pandemic											
More Competencies State Government in Pandemic											
More Competencies Parliament in Pandemic											
Rating Covid-19 Information by Government											
Rating Covid-19 Information by Parliament											
Rating Covid-19 Information by Parties											
Rating Covid-19 Information by RKI											
Rating Covid-19 Information by TV											
Rating Covid-19 Information by Print Media											
Rating Covid-19 Information by Internet											
Rating Covid-19 Information by Scientists											
Perceived Freedom of Speech: Pandemic											
Perceived Freedom of Speech: General											
Contact Tracing App Installed											
Contact Tracing App Use											
Working from Home											
Previous Covid-19 Infection											
Satisfaction with Democracy											
Data Privacy Concerns											
Data Sharing Attitudes (Four Items)											
Feeling of Control over Personal Data											
Protest Participation											
Government’s Responsibility Care											
Government Expenses Care											
Government’s Responsibility Pensions											
Government Expenses Pensions											
Retirement Age											
Government’s Responsibility Unemployed											
Government Expenses Unemployed											
Government’s Responsibility Families											
Government Expenses Families											
Priority Family Policy											
Child Wellbeing Rating (Experiment)											

(continues on next page)

(continued from next page)

Study Wave	45	46	47	48	49	50	51	52	53	54	55
Physical Activity Level (Three Items)							■				
Screentime							■				
Snacking							■				
Depression (Two Items)							■				
Loneliness							■				
Acceptance of Tax Evasion									■		
EU Unification Views										■	
EU Decision-making Authority									■		
Reduce Income Difference									■		
Owning a Mobile Phone				■				■		■	
Owning a Computer				■				■		■	
Using Facebook/Instagram/LinkedIn/Twitter/Xing				■				■			
Trust in Parliament						■					■
Trust in Government											■
Trust in Local Government											■
Trust in Parties											■
Trust in European Commission											■
Trust in European Parliament											■
Trust in TV											■
Trust in Print Media											■
Trust in Social Media											■
Trust in Education System											■
Trust in Police											■
Trust in Healthcare System											■
Trust in Federal Constitutional Court											■
Trust in Justice											■
Labour Impact of Migrants							■	■			
Credibility Migration Studies							■	■			
Family Satisfaction					■			■		■	■
Financial Satisfaction					■						
Computer Usage					■						
Short-time Work					■					■	■

Appendix B Figures

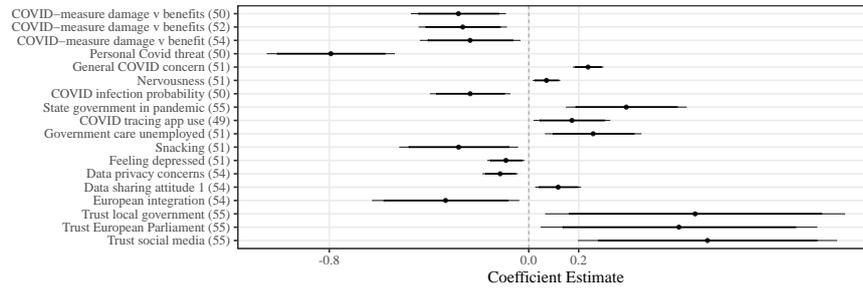


Figure B1

Initially significant conditioning effects with treatment group defined as anyone invited to the MCS

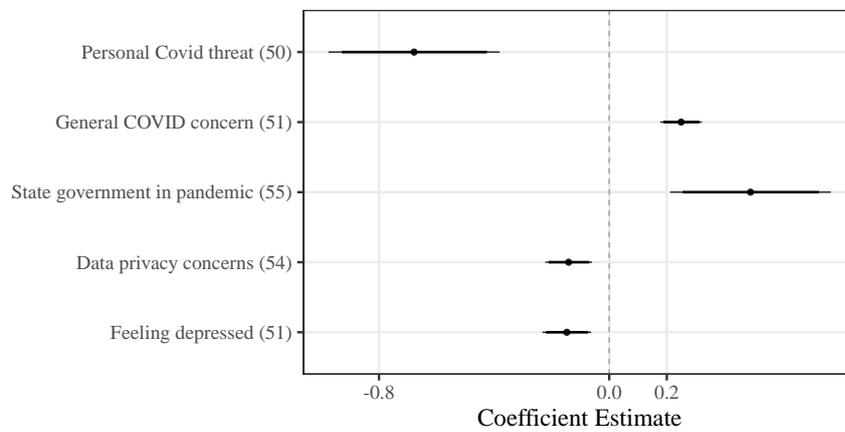


Figure B2

Significant conditioning effects with Bonferroni correction and inclusion of demographic controls and treatment group defined as anyone invited to the MCS

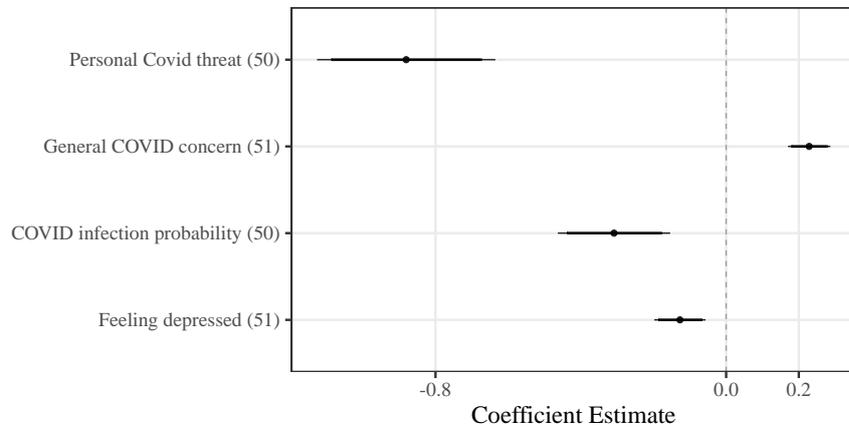


Figure B3

Significant conditioning effects with Bonferroni correction but without demographic controls

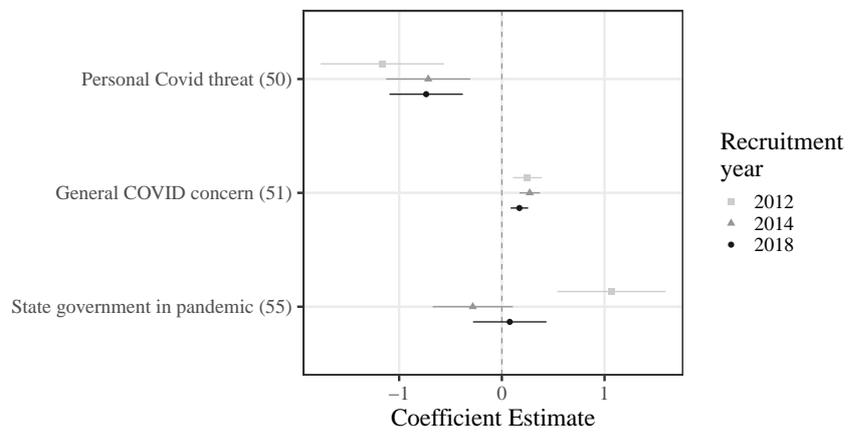


Figure B4

Conditioning effects with Bonferroni correction that were significant in at least one recruitment wave subsample by itself