ORIGINAL ARTICLE

From German Internet Panel to Mannheim Corona Study: Adaptable probability-based online panel infrastructures during the pandemic

Carina Cornesse (iD) | **Ulrich Krieger** | **Marie-Lou Sohnius** |
Marina Fikel | **Sabine Friedel** | **Tobias Rettig** |
Alexander Wenz | **Sebastian Juhl** | **Roni Lehrer** |
Katja Möhring | **Elias Naumann** | **Maximiliane Reifenscheid** |
Annelies G. Blom

University of Mannheim, Mannheim,
Baden-Württemberg, Germany

Correspondence
Carina Cornesse, SFB 884 - Political
Economy of Reforms, University
of Mannheim, Mannheim, Baden-
Württemberg, Germany.
Email: carina.cornesse@uni-mannheim.de

Abstract

The outbreak of COVID-19 has sparked a sudden demand for fast, frequent and accurate data on the societal impact of the pandemic. This demand has highlighted a divide in survey data collection: Most probability-based social surveys, which can deliver the necessary data quality to allow valid inference to the general population, are slow, infrequent and ill-equipped to survey people during a lockdown. Most non-probability online surveys, which can deliver large amounts of data fast, frequently and without interviewer contact, however, cannot provide the data quality needed for population inference. Well aware of this chasm in the data landscape, at the onset of the pandemic, we set up the Mannheim Corona Study (MCS), a rotating panel survey with daily data collection on the basis of the long-standing probability-based online panel infrastructure of the German Internet Panel (GIP). The MCS has

provided academics and political decision makers with key information to understand the social and economic developments during the early phase of the pandemic. This paper describes the panel adaptation process, demonstrates the power of the MCS data on its own and when linked to other data sources, and evaluates the data quality achieved by the MCS fast-response methodology.

1 | INTRODUCTION

Social surveys have been around for many decades and are conducted all across the world (e.g. Schnaudt et al., 2014). They offer political, economic and societal decision makers the opportunity to gain insights into people's attitudes, behaviour and living conditions (e.g. Smith et al., 2006). Social surveys, therefore, serve an important role in democracies by providing data on public opinion and enabling evidence-based policy-making (e.g. Dymond-Green, 2020). However, to reflect reality, social surveys have to accurately represent their population of interest (e.g. Malhotra & Krosnick, 2007). To achieve this, many high-quality social surveys are based on probability samples drawn from sampling frames of the general population, such as address lists or population registers (e.g. Lynn et al., 2004). In addition, they are commonly conducted via face-to-face (e.g. Williams & Brick, 2018) or telephone interviews (e.g. Dutwin & Buskirk, 2020), or via a mix of modes (e.g. Jäckle et al., 2015).

These design features of many probability-based social surveys ensure accurate representation and full inclusiveness of key population subgroups leading to valid inference from the data to the general population (e.g. Yeager et al., 2011). However, this high data quality comes at a price, both financially and in terms of long fieldwork periods (e.g. Beullens et al., 2018). In addition, some of the features responsible for the high data quality have been impossible to uphold during the pandemic, in particular face-to-face interviewing (e.g. Gummer et al., 2020). During lockdown, for example, many social surveys needed to be halted or postponed (e.g. Scherpenzeel et al., 2020). And still afterwards, restarting face-to-face fieldwork in private households proves difficult under social distancing measures (Prior, 2020).

In this context, long-standing panels fared better than cross-sectional surveys. For example, the UK Household Longitudinal Study (UKHLS) adapted its regular fieldwork and additionally started a monthly online add-on survey on the impact of COVID-19 (Burton et al., 2020). Another example is the German Socio-Economic Panel (SOEP), which implemented a weekly COVID-19 telephone add-on survey (Kühne et al., 2020). The panel designs of these studies and their personnel and data infrastructures enabled them to extend their existing pre-COVID-19 data collections with during-COVID-19 measurements. Nonetheless, this adaption was only achieved with a significant time lag. To answer some of the immediately pressing questions during the early phase of the pandemic when people's lives changed rapidly, much faster and more frequent data collection, processing, and reporting was needed.

Non-probability online surveys, some of them with impressive sample sizes and/or multi-national scope, tried to cater to this demand. They often presented their results nearly in real-time to interested audiences (for an overview of both probability-based and non-probability surveys on the impact of COVID-19 see the Oxford Supertracker of COVID-19 surveys; University of Oxford, 2020). These non-probability online surveys, however, rely on pools of self-selected volunteers rather than probability-based population samples. Such non-probability methodology has been repeatedly shown to lead to inaccurate results (see Cornesse et al., 2020 for an overview and Sturgis et al., 2018 for an example). In quiet times, the consequences of such inaccuracies may be limited to some negative press coverage (e.g. Travis, 2017). In times of crisis, though, such inaccurate predictions can be particularly dangerous if decision-makers base their actions on them (Lewis, 2020). Consequently, the COVID-19 crisis has generated a dilemma in which the accurate data we need are not available fast enough, whereas inaccurate and potentially misleading data inundate us in times when scientific evidence is needed the most.

In this paper, we propose probability-based online panel infrastructures as a potential solution to this problem. For this purpose, we demonstrate how the German Internet Panel (GIP) adapted its design by turning into the Mannheim Corona Study (MCS) for a period of 16 weeks between 20 March and 10 July 2020. We describe the online panel adaptation process, demonstrate the power of the MCS from a multifaceted and interdisciplinary research and policy consultation perspective, and evaluate the MCS data quality. We conclude with a discussion of the extent to which adaptable probability-based online panel infrastructures can fill the social data demand during a crisis at short notice.

2 | COMMON FEATURES OF PROBABILITY-BASED ONLINE PANEL INFRASTRUCTURES

Probability-based online panel infrastructures aim to combine the best of two worlds: population inference and innovative online survey methods. In this section, we describe the key features of adaptable probability-based online panel infrastructures. In addition, we provide examples of how these features are applied in existing probability-based online panel infrastructures around the world.

2.1 | Probability sampling and offline recruitment

As their name suggests, probability-based online panels rely on probability sampling procedures. Common approaches applied in practice are address-based sampling (e.g. AmeriSpeak Panel, NORC at the University of Chicago, 2019), sampling from population registers (e.g. GESIS Panel, Bosnjak et al., 2018) and random digit dialling (e.g. Life in Australia Panel, Kaczmirek et al., 2019). Once the samples are drawn, sample members are approached offline, for example via face-to-face interviews (e.g. the early recruitment rounds of the GIP, Blom et al., 2015), telephone (e.g. GESIS Online Access Panel Pilot, Schaurer, 2017) or postal mail (e.g. KnowledgePanel, Ipsos, 2020). In some probability-based online panels, a combination of contact modes is used (e.g. Gallup Panel, GALLUP, 2020), usually sequentially with cheaper modes preceding more expensive modes (e.g. AmeriSpeak Panel, Bilgen et al., 2018).

At the stage of the initial offline contact, many probability-based online panels collect data on sample members, typically in an interviewer-mediated recruitment survey (e.g. LISS Panel, Scherpenzeel

& Toepoel, 2012). Panels may also be recruited as extensions of traditional social surveys, the so-called piggy-backing approach (e.g. NatCen Panel, Jessop, 2017). And some skip the interviewer-mediated process altogether and recruit directly via postal mail with letters containing log-in information to the online panel (e.g. Norwegian Citizen Panel, Høgestøl & Skjervheim, 2014).

2.2 | Online survey data collection

Probability-based online panels share the rigorous sampling procedures and the offline recruitment of the traditional social surveys, but subsequently switch from the offline mode to online data collection (Blom et al., 2016). The online surveying makes for fast (Couper, 2011) and cost-efficient data collection (Kaminska & Lynn, 2017). Furthermore, the online mode enables researchers to collect data at a high frequency (Blom et al., 2020), use visual and audio cues (Haan et al., 2017), collect respondent paradata (McClain et al., 2019) and include elaborate experiments (Kunz & Fuchs, 2019). Moreover, it allows panel participants to fill out the questionnaires at their own pace, convenience and location (Couper et al., 2017). Finally, the self-completion aspect of online surveys eliminates undesirable interviewer effects (West & Blom, 2017) and reduces the social desirability bias in sensitive questions (Kreuter et al., 2008).

However, this transfer from offline recruitment to online data collection also comes at the price of potentially systematic exclusion of sample members who do not use the internet (Blom et al., 2017). Therefore, a number of probability-based online panels implement offline population inclusion strategies. Two common approaches to this can be differentiated (Cornesse & Schaurer, 2021): Either sample members without internet access receive the necessary equipment to participate in the online panel (e.g. Understanding America Study, USC 2017, and ELIPSS Panel, Revilla et al., 2016) or they are surveyed in an alternative mode (e.g. Life in Australia Panel, Kaczmirek et al., 2019, and GESIS Panel, Bosnjak et al., 2018).

With regard to the COVID-19 pandemic, the most important advantages of the online mode of data collection are the speed and flexibility at which data can be collected. Consequently, during the initial phase of the pandemic, probability-based online panel infrastructures across the globe adapted their designs to satisfy the demand for fast, frequent and accurate data. Some collected data monthly (e.g. LISS Panel, van Tilburg et al., 2020), weekly (e.g. COVID Impact Survey based on the AmeriSpeak panel, Data Foundation, 2020), or even daily (e.g. Understanding Coronavirus in America based on the Understanding America Study, Kapteyn et al., 2020).

3 | ADAPTING A PROBABILITY-BASED ONLINE PANEL INFRASTRUCTURE DURING THE PANDEMIC

In the following, we describe how one particular probability-based online panel infrastructure, the German Internet Panel (GIP) was adapted to the need for fast, frequent and accurate data during the early stages of the COVID-19 pandemic in the Mannheim Corona Study (MCS).

3.1 | The German Internet Panel

The GIP is a multitopic, probability-based online panel of the general population in Germany. To date, the GIP has seen three independent recruitment rounds: in 2012, 2014 and 2018. In 2012

and 2014, samples were drawn using a three-stage area sampling procedure with a random-route approach and full listing of all households along the route. From the thus collected addresses, households were randomly drawn and approached for face-to-face recruitment interviews. Subsequently, all age-eligible household members were invited to become GIP panel members (Blom et al., 2015). During the recruitment process, persons without Internet access were provided with the necessary equipment and support (Blom et al., 2017). In 2018, the GIP sample was drawn in a two-stage sampling procedure where, first, municipalities were sampled and, subsequently, individuals were sampled from municipal population registers. All sampled individuals were approached via postal mail and asked to register to the GIP online (Cornesse et al., 2021). In each recruitment round, people were only sampled if they were between 16 and 75 years old. The upper age limit was chosen because it was expected that the GIP's design (web survey mode only, no sampling of people who live in retirement homes) would be unlikely to lead to an adequate representation of older people (e.g. low online panel registration rates and potential biases regarding health and skills; Hunsaker & Hargittai, 2018). However, once recruited, participants remained in the panel sample even as they grew older than 75. The oldest GIP participants were 83 years old in 2020.

After their recruitment, the 2014 and 2018 samples were pooled with the initial 2012 sample. In total, 21.8% of all people ever drawn into the GIP gross samples were recruited to the panel. Every other month, all active panelists are invited to online surveys covering a variety of social, political and economic topics. In 2019, on average, 72.2% of all invited GIP panelists (or 62.1% of all panelists ever recruited to the GIP, i.e. including people who de-registered from the panel over time) responded to the bimonthly surveys. Questionnaires are designed to take 20–25 min and respondents receive 4 Euros for each completed survey plus a 10 Euro bonus, if they fill out all six surveys in a year. Incentives are credited towards respondents' panel accounts and paid out twice a year as online vouchers, bank transfers or charitable donation according to the panelists' preferences.

An established panel operations infrastructure lies at the heart of the GIP and was crucial for the successful and quick data collection adaptation during the pandemic. Some of the most important aspects of this infrastructure include access to additional financial resources and administrative support, the existence of well-rehearsed and largely automated operations processes, and an operations research team with experience in the day-to-day running of the panel (Blom et al., 2020).

3.2 | The Mannheim Corona Study

When the pandemic hit Germany, the GIP was quickly adapted to the demand for data on the social and economic consequences of the pandemic. Data were collected every day for 16 weeks from 20 March until 10 July 2020 using a rotating panel design. For this purpose, the GIP participants were randomly allocated to one of the eight subsamples. Seven of those subsamples were assigned to a weekday on which they received survey invitations during the 16 weeks of the MCS. GIP participants allocated to the eighth group were not part of the MCS sample, but were kept as a control group (see Figure 1 for a schematic depiction of the GIP and its adaptation to the MCS). The reason for excluding a random control group from the MCS was that we wanted to assess whether and to what extent conducting the MCS had an impact on the GIP as its underlying panel infrastructure (e.g. in terms of response rates, see Section 5.2).
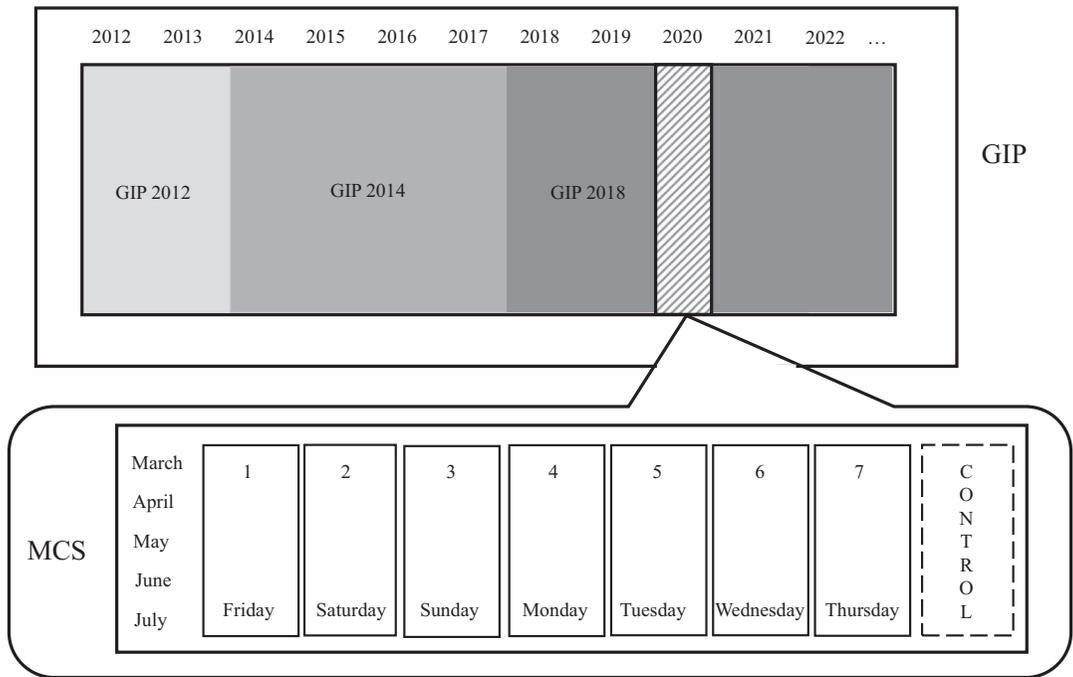
FIGURE 1 Overview of the Mannheim Corona Study

In the MCS, each weekday a different subsample received an email invitation to the day's survey. Contacted panel members were given 48 h to participate, but encouraged to take part on the assigned day of the week, that is within the first 24 h. Indeed, on average, 86.6% of all respondents participated on the day that they received the invitation (Blom et al., 2020). Persons who responded directly on the first day (e.g. Monday) were included in the analysis of that specific day (Monday). Answers of respondents, who participated on the next day (Tuesday), were analysed together with the answers on the day of the next subsample. Within 1 week, the questionnaire remained exactly the same for all participants. Across weeks, we allowed for some variation in the questionnaire to account for changing circumstances and new political debates, such as the roll-out of the Coronavirus warning app in June 2020 (The Federal Government, 2020).

The questions in the MCS revolved around the impact of the COVID-19 pandemic on the German population, covering a range of interdisciplinary aspects, such as whether people were being furloughed or working from home, or how they were organizing childcare when kindergartens and schools were closed. We also asked what people thought about the government's crisis management, whether they adhered to social distancing rules, and whether and how much they were afraid of the virus. In total, 4,387 people responded to at least one of the MCS survey requests and 1,910 participants responded to all of them, leading to a total of 54,696 completed questionnaires over the course of the 16 MCS weeks. The median MCS questionnaire length was 8–9 min. Participation was incentivized with 2 Euros per survey credited towards participants' regular GIP accounts. This amounts to a total of 32 Euros for respondents who participated in every MCS survey. In total, 109,392 Euros were paid out as incentives to the respondents over the course of the MCS study period.

The analyses conducted with the MCS are weighted by a combined propensity and raking weight. First, we calculated a response propensity weight, which projects the characteristics of the MCS participants to the general GIP study using employment status and occupational sector

as weighting variables. Then, a raking weight extrapolated the characteristics of the MCS participants to those of the general population of Germany (based on the German microcensus, Destatis, 2020) in terms of age, gender, marital status, highest level of education, household size and federal state. The propensity weight was used as a pre-weight in the raking process. A chained equation algorithm imputed missing values in the weighting variables. The final weight was trimmed for values >4 and values <1/4. We multiplied the nonresponse weight with the GIP's design weight, which corrects for the unequal inclusion probabilities resulting from the fact that the GIP had three recruitment rounds (2012, 2014 and 2018; see Kolb et al., 2021 for a full description of an equivalent design weighting approach in the GESIS Panel). Each of the GIP recruitment samples by themselves can be regarded as approximately self-weighting (Cornesse et al., 2021).

4 | DEMONSTRATING THE POWER OF THE MCS DATA

The MCS data have been used for many different purposes, including government agency consulting, science communication and in-depth substantive research in different subfields of political science, economics, sociology and psychology. In the following, we demonstrate the power of the MCS data in terms of three aspects relevant to the research community assessing the impact of the pandemic on society: (1) drawing inference to the general population on a daily basis, (2) augmenting MCS data with official COVID-19 statistics, and (3) augmenting MCS data with prior and subsequent GIP data.

4.1 | Population inference on a daily basis

One essential feature of the MCS is the fast and frequent data collection, data processing and result reporting. Each day of the MCS study period, on average 491 MCS participants completed a survey. The survey data were processed immediately and results were communicated in daily reports on the MCS website (https://www.uni-mannheim.de/en/gip/corona-study/). We considered this speed and frequency of data collection, data processing and result reporting to be important, because we expected people's lives to change at a fast pace. Indeed, our results provide support for this expectation.

For example, Figure 2 shows the estimated population proportion that had met with friends, family or colleagues from outside of their own household socially in the previous 7 days across the MCS study period. On the first day of the MCS (20 March 2020), 62.3% of the population had met with other people socially in the previous week. This number dropped rapidly to 29.8% by the 2 April, at the height of the lockdown in Germany. This low level was maintained for a few days until approximately 5 April, when the share of people meeting others socially was at 62.0%. Subsequently, social contact increased steadily, reaching 80.0% exactly 2 months after the start of the MCS (20 May) and remaining at this level until the end of the MCS with 82.7%. This example of the development in social interactions illustrates the necessity of collecting data frequently, because the dramatic changes in people's social lives would otherwise have been overlooked.

While these results demonstrate how quickly people's behaviour changed during the early stages of the pandemic in Germany, the results from the MCS suggest that the same is true for people's attitudes towards COVID-19-related topics. Figure 3, for example, shows the estimated proportion of the population supporting the closure of public facilities such as universities,
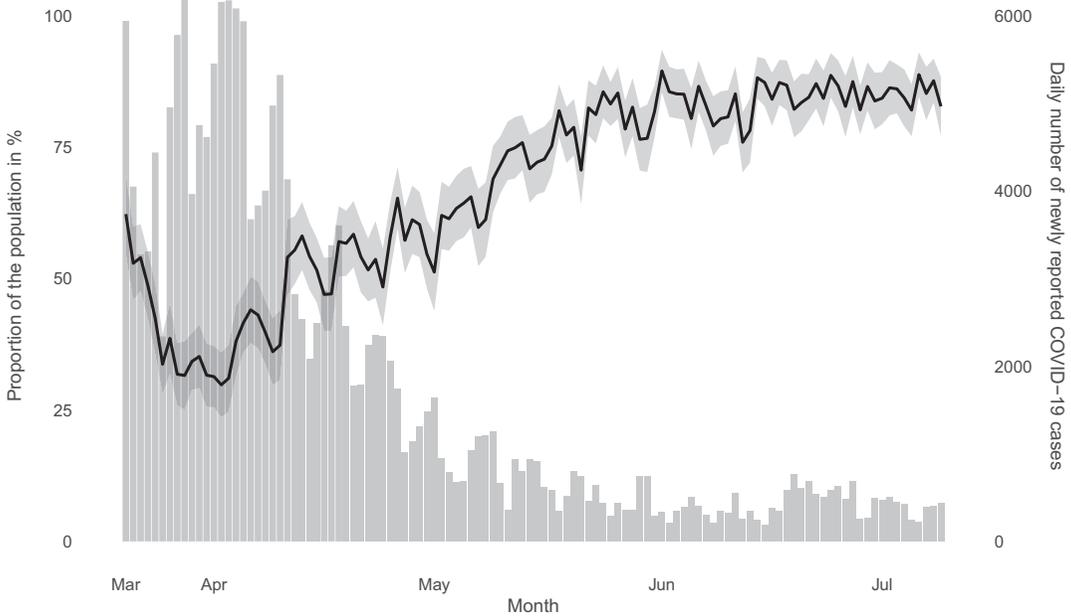
FIGURE 2 Share of Mannheim Corona Study (MCS) respondents meeting with other people from outside of their own household during the previous week across the MCS study period

Note: The black line depicts the share of people who had met with friends, relatives or work colleagues from outside of their own household the previous week across the MCS study period. The shaded area denotes the 95% confidence interval around the estimates. The bars display the reported daily number of new COVID-19 cases in Germany in the official statistics provided by the Robert Koch Institute.

schools and kindergartens. This endorsement was overwhelming at the beginning of the pandemic, with 95.5% supporting closures on 20 March. It decreased slowly over the following days (88.5% on 10 April), but then quickly dropped over the course of the next month. By 15 June, only 35.0% still supported the closure of public facilities, further dropping to 18.2% by the end of the study.

The MCS reports published every workday on the MCS website together with more detailed tailored analyses on specific policy issues were used by the German COVID-19 crisis cabinet, the Federal Ministry of Labour and Social Affairs (BMAS) and the Federal Ministry of Domestic Affairs (BMI) to inform national policy decisions. The probability sampling and offline recruitment of the underlying GIP study provided political and economic decision-makers with the data quality needed for their COVID-19 actions (e.g. Möhring et al., 2020).

Following these, early analyses catered to the specific needs of the crisis situation, ensuing academic research dived deeper into specific research topics. Mata et al. (2020), for example, explore developments in health behaviours and mental health during the early phase of the pandemic. They find that 'mental health and health behaviours worsened as an immediate response to [the spread of COVID-19 and the newly devised counter-measures], but mostly returned to pre-lockdown levels within three months' (Mata et al., 2020:2). They also examine the association between mental health and health behaviour, such as snacking and physical activity, with evidence suggesting that an increase in eating snacks and a decrease in physical activity are significantly associated with a decrease in mental health during the pandemic. Their research thus
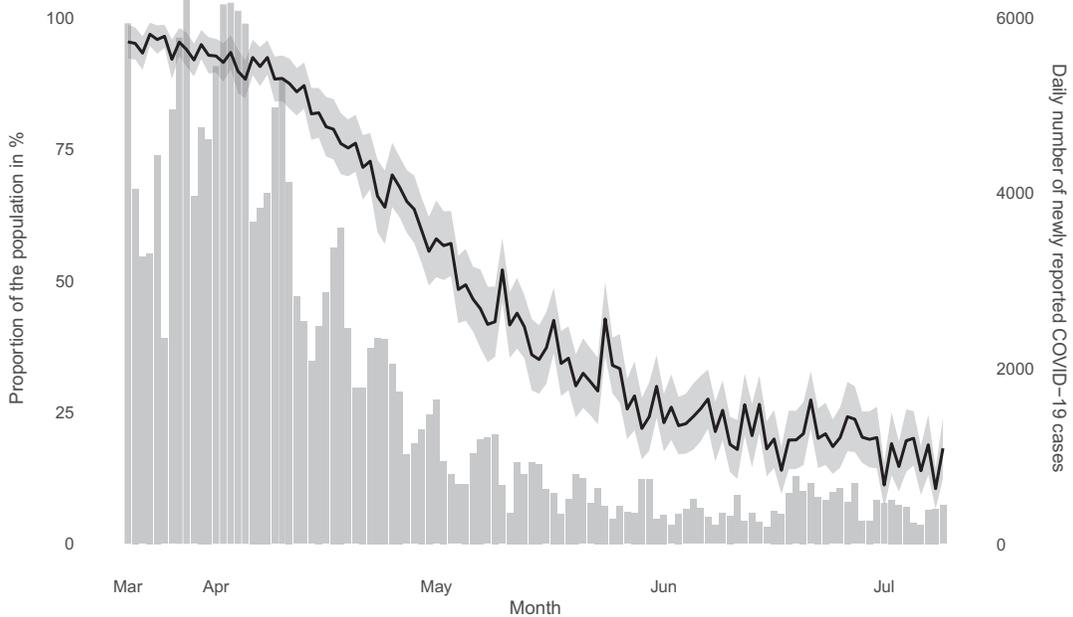
FIGURE 3   Share of Mannheim Corona Study (MCS) respondents endorsing the closure of public facilities such as universities, schools and kindergartens
Note: The black line depicts the share of people who endorse the closure of public facilities across the MCS study period. The shaded area denotes the 95% confidence interval around the estimates. The bars display the reported daily number of new COVID-19 cases in Germany in the official statistics provided by the Robert Koch Institute.

provides valuable insights regarding the societal impact of the pandemic from a psychological health perspective.

4.2 | Augmenting MCS data with official COVID-19 statistics

The daily data collection of the MCS not only allows drawing daily inference to the general population, but also comparing the survey estimates to the daily official COVID-19 statistics. Since 4 March 2020, the Robert Koch Institute (RKI, German centre for disease control) published daily COVID-19 statistics on their website (RKI, 2020a). These official statistics are also available as a time series dataset from the website of the European Centre for Disease Prevention and Control (ECDC, https://opendata.ecdc.europa.eu/covid19/casedistribution/). We downloaded these time series data and merged them with the MCS data to explore associations between the epidemiological development of the pandemic and its social consequences. The RKI provides a number of COVID-19 statistics, such as the cumulative number of known COVID-19 infections, the number of COVID-19-related deaths and the number of new COVID-19 infections per day in Germany.

In order to understand the official COVID-19 statistics, it is important to know how COVID-19 cases are identified and reported in the German health system. If people notice COVID-19 symptoms or have any other reason to belief that they might have contracted the virus (e.g. due to traveling to COVID-19 hotspots or spending time with confirmed COVID-19 cases), their general physician and/or the local public health authorities (LPHA, 'Gesundheitsamt') will urge or even require them to be tested in accordance with regional test-and-trace policies (for information on

the general COVID-19 testing strategy in Germany see Federal Ministry of Health, 2020). Test laboratories report positive COVID-19 tests back to the LPHA.

In total, Germany has more than 400 of these LPHAs (RKI, 2020b). The LPHAs report positive cases and COVID-19-related deaths back to the RKI. This is supposed to happen on a daily basis. In practice, however, time lags exist in the reporting between test laboratories to the LPHAs and from the LPHAs to the RKI, in particular during the weekend and after public holidays (RKI, 2020c). The RKI publishes aggregated statistics on the development of the pandemic every day, retrospectively correcting statistics for the time lags (RKI, 2020c). In addition, the accuracy of the RKI statistics may depend on the national test capacity. The more tests are conducted, the more cases can be detected. In Germany, the testing capacity has continuously increased over the course of the MCS study with 111 laboratories and a capacity of 64,725 tests per day when the MCS started and 145 laboratories and a capacity of 176,898 tests per day at the end of the MCS (Statista, 2020).

The grey bars in Figure 2 show the official new COVID-19 infections plotted against the proportion of the population that had met socially with friends, family or colleagues from outside of their own household during the previous week with a clear negative correlation ($r = 0.4$). Similarly, Figure 3 shows a strong positive association between the endorsement of the closure of public facilities (universities, schools and kindergartens) and new COVID-19 infections ($r = 0.5$). While the results clearly show how the MCS estimates are associated with the development in officially reported infections, it should be noted, however, that the data provide no evidence on causality.

The possibility of augmenting the MCS data with official statistics on COVID-19 has so far already been used in several research contexts. This includes the aforementioned study on mental health and health behaviours by Mata et al., 2020. In addition, a study by Naumann et al. (2020) on COVID-19 policies in Germany and their social, political and psychological consequences puts the MCS estimates on changes in support for COVID-19-related policies into the context of the epidemiological development of the pandemic.

4.3 | Augmenting MCS data with prior and subsequent GIP data

Since MCS participants are directly recruited from the long-standing GIP infrastructure, detailed longitudinal data on its respondents is available from GIP surveys conducted prior to the MCS. This is particularly useful for examining changes in people's lives from before the pandemic to different time points during the pandemic. For example, in the regular GIP survey wave in January 2020, respondents were asked about their current employment status. In January, the pandemic had not yet spread to Germany. Therefore, we can use the GIP data from January as a pre-COVID-19 measurement and examine how the employment situation of MCS participants changed since then. As an example, for all those working fulltime in January, the alluvial diagram in Figure 4 shows changes in employment situation across three time points in the early COVID-19 phase in Germany.

The longitudinal analysis shows that the employment situation of nearly half the people who worked fulltime before the pandemic remained stable across the MCS study period. However, approximately one third of people moved to working from home all or at least some of the time. Since the middle of the MCS study period, around 10% of people went from their fulltime work into the governmentally subsidized short-time work schemes (Federal Ministry of Labour and Social Affairs, 2020). In addition, the share of people who were
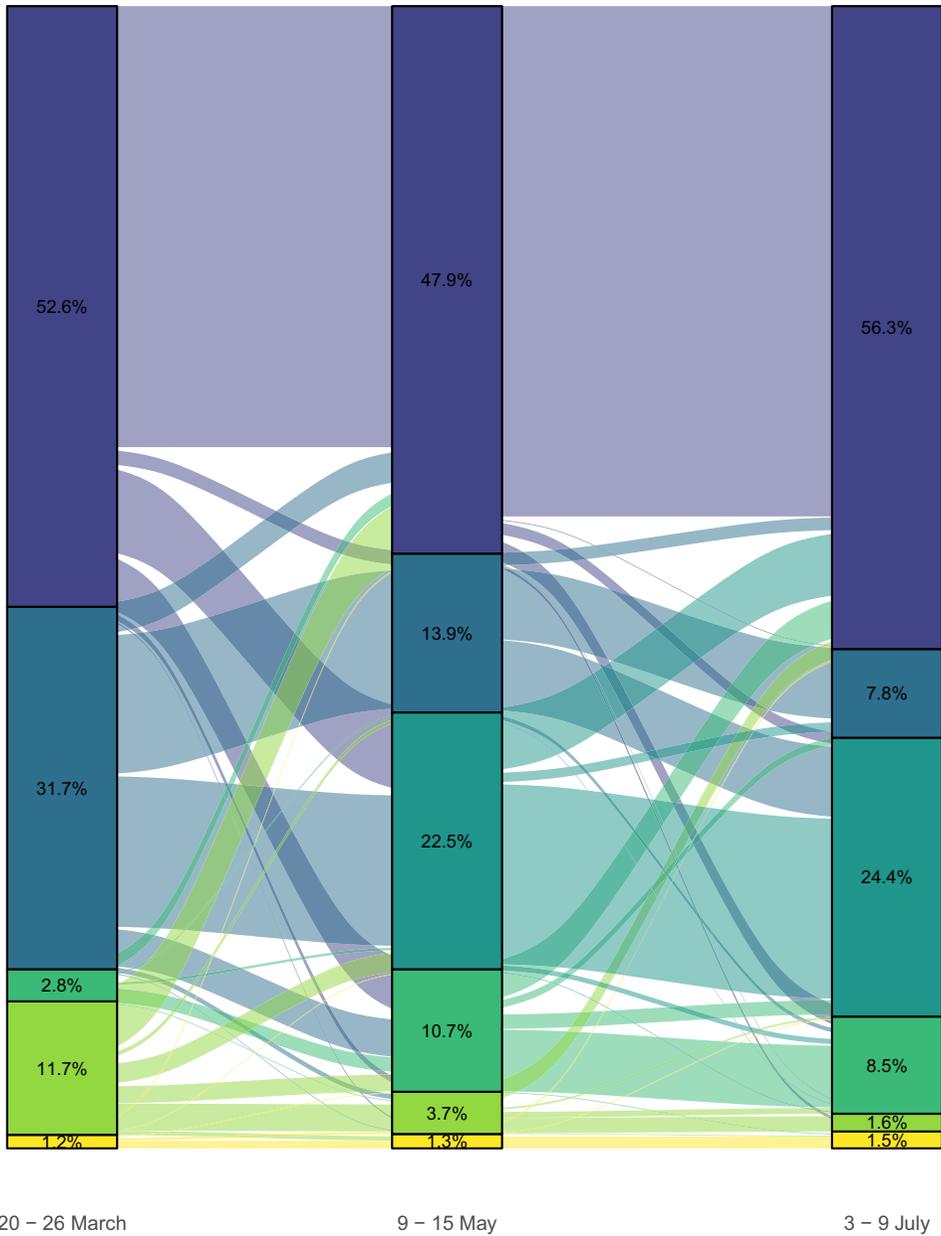
FIGURE 4 Changes in employment situation over the course of the Mannheim Corona Study (MCS)
Note: The three vertical bars represent the share of MCS respondents across the different employment situation categories during the first, eighth, and last week of the MCS. The streams in between the bars represent the changes in employment situation across these three MCS time points. The analysis is restricted to people who worked full-time in January 2020. It should also be noted that the questionnaire in the first week did not yet differentiate between people working from home all the time and working from home only part of the time.

furloughed by the end of March was relatively high (11.7%) but was considerably reduced by the beginning of July (1.6%). Most of the temporarily furloughed persons went back to their pre-COVID-19 employment situation at their regular workplace by mid-May. Throughout the early pandemic phase in Germany, the share of persons who became unemployed since January remained low (1.5%).

The combined MCS-GIP data have been used effectively to inform government agencies about the changes occurring in the German population frequently and in a timely manner, in particular, the German Federal Ministry of Labour and Social Affairs. In addition, these longitudinal data are also used in several interdisciplinary academic research projects. For example, Möhring et al. (2020) investigate inequality in employment during the early stages of the COVID-19 pandemic in Germany. They provide detailed insights into the changes that occurred in people's employment situation from before the pandemic until the end of the MCS. Among other results, they find that, while many highly educated persons with relatively high incomes were able to switch to working from home, most people with relatively low education and income usually had to remain at their usual workplace thus carrying a higher risk of workplace infection.

In addition to augmenting MCS data with prior GIP waves, the continuation of the GIP with its bimonthly data collection will allow research into the future social development as the pandemic continues. For example, data gathered in regular GIP survey waves in 2021 will allow us to assess how people's employment situation developed after the end of the MCS. Similarly, data gathered in the MCS on whether people intended to install the official German contact tracing app once it was launched (Blom et al., 2021a, 2021b, 2021c, 2021d, 2021e, 2021f, 2021g), will be augmented with GIP survey data collected after the end of the MCS on whether people actually installed the app.

5 | EVALUATING THE QUALITY OF THE MCS

While in the previous section, we demonstrated the power of the MCS data, in the following, we will provide insights into the quality of the MCS. We will focus on two aspects: data quality of the MCS data per se and the impact that conducting the MCS had on the GIP infrastructure.

5.1 | MCS response rates and bias assessments

Asking panel participants to complete surveys within 48 h once a week puts a high burden on them. The success of a study like the MCS in terms of data quality, therefore, depends on whether participation can be kept reasonably high across the whole study phase while at the same time keeping the bias low. To study the extent to which this was achieved in the MCS, we examine response rates and biases on each day of the MCS study period. For each day of the MCS, we compute response rates by dividing the number of GIP participants who completed a survey on that day by the number of persons who were invited to do so. In addition, for each day of the MCS, we compare the distribution of a set of socio-demographic characteristics in the MCS data to the official statistics of the German microcensus, which serves as a population benchmark (Destatis, 2020). We aggregated the bias assessments across the socio-demographic characteristics using the average absolute relative bias (AARB, Groves, 2006). The AARB allows us to gain an overview of the extent to which the MCS respondent sample on any given day deviates from the target population. The AARB is given by

$$AARB = \sum_{k=1}^{K} \left(\left| \frac{y_{sk} - y_{bk}}{y_{bk}} \right| / K \right) * 100$$

where y_{sk} stands for the proportion of people in a category k of a given socio-demographic variable for a daily MCS respondent sample s, y_{bk} stands for the corresponding proportion in the benchmark statistic b. In total, y includes seven variables with $K = 20$ categories. The included variables are gender (categories: female and male), age (categories: 18–29, 30–39, 40–49, 50–59, and 60 and older), education (categories: low, middle and high), citizenship (categories: German and non-German), marital status (categories: single, married, divorced and widowed) and household size (categories: one, two, three, four and more household members). In the AARB, the absolute relative biases in the listed categories are averaged to provide an overview bias statistic. The age range of our analyses is restricted to the German population aged 18–78. The MCS data cannot be used to draw inferences beyond this age range.

Apart from the AARB, other measures can be used to assess the bias. This includes a range of disaggregated measures which examine item-specific bias (e.g. Sturgis et al., 2018). Examples include investigating relative biases by variable (see Figure A1 in the Appendix for results from the MCS) or percentage point deviations between survey data and benchmark data (see Figure 6 for results from the MCS). Alternatives to the AARB for examining bias across a set of characteristics on the aggregate level include the average absolute error (Yeager et al., 2011), root mean squared error (MacInnis et al. 2018) and R-Indicators (Luiten & Schouten, 2013). We chose the AARB as an overview statistic because it is well known and commonly used in the context of comparing survey data to external population benchmarks (rather than, e.g. comparing to sample frame information, for which R-Indicators are particularly well established, Schouten et al., 2011, or evaluating statistical models, for which the root mean squared error is popular).

Figure 5 displays the response rates (upper dark grey line) and AARBs (lower black line) across the MCS study period plotted against the newly reported COVID-19 cases from the RKI data. It shows that response rates are very stable across the MCS study period, with 61.1% on average across the 16 weeks of the MCS and a range from 57.6% in week 16 to 64.4% in week 2. If instead of looking at daily response rates among the people who were invited to the MCS (as displayed in Figure 5), we examine the share of MCS respondents among all people who were ever drawn into the GIP gross samples, we can also conclude high stability, albeit at a much lower level (average response rate among all gross sample members: 9.0%; range: 8.5% in week 16 to 9.5% in week 2).

AARBs are similarly stable across the MCS study period, with an average of 19.9% and a weekly average ranging from 18.6% in week 3 to 20.9% in week 10. Examining the development of the response rates and AARBs against the newly reported COVID-19 cases per day (see the bars in Figure 5), we can see that neither of the data quality indicators rise or fall with the epidemiological development.

To gain a deeper understanding of potential biases, we also provide disaggregated descriptive results on the deviation of the MCS sample from the microcensus with regard to the characteristics included in the calculation of the AARB (see Figure 6). Moreover, to put these results in a survey comparative perspective, we compare the deviations from the microcensus found in the MCS data to those from the full GIP sample before the start of the MCS as well as the latest ALLBUS sample from 2018 (GESIS 2019). We chose ALLBUS data for this comparison because the ALLBUS can be regarded as a gold standard probability-based face-to-face social survey of the German population (see also Table A1 in the Appendix).
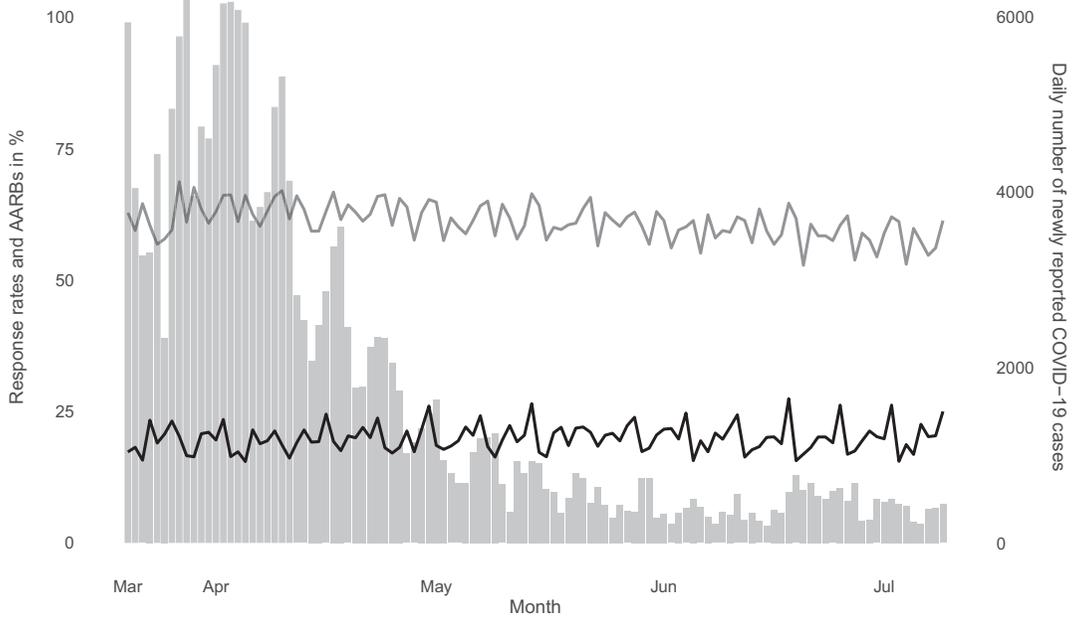
FIGURE 5 Response rates and average absolute relative bias (AARB) in the Mannheim Corona Study (MCS)
Note: The black line depicts the AARB and the grey line depicts the response rate across the MCS study period. The bars display the reported daily number of new COVID-19 cases in Germany in the official statistics provided by the Robert Koch Institute.

Overall, all three samples show similar biases. They all represent gender well, but under-represent younger people, people with low education, people living in one-person households, single persons and non-German citizens. They also all overrepresent older people, people with high education, people living in two-person households, married persons and German citizens. However, some notable differences in the size of the biases can be observed. Regarding age, for example, the MCS bias is smaller than the GIP bias and similar to the ALLBUS bias. This may be due to the fast-and-frequent data collection design of the MCS, which may have been more appealing for young adults and potentially overwhelming for some of the elderly people. Regarding education, however, the MCS bias is similar to the GIP bias and both are higher than the ALLBUS bias. This is likely due to the online data collection mode of the GIP and MCS, which tends to be less appealing for people with low education (see also Cornesse & Schauer, 2021 on this topic).

5.2 | The impact of the MCS on the GIP

While for the MCS per se, it is important to achieve high data quality, from a more general panel perspective, it is essential to keep the GIP infrastructure intact. We expected that panel participants who were suddenly asked to participate in surveys on the impact of COVID-19 every week for a total of 16 weeks might feel overburdened and were thus worried that they may be more hesitant to continue with the regular GIP survey waves every other month. To be able to examine the impact of the MCS on regular GIP survey participation, we excluded a random subset of 1/8th
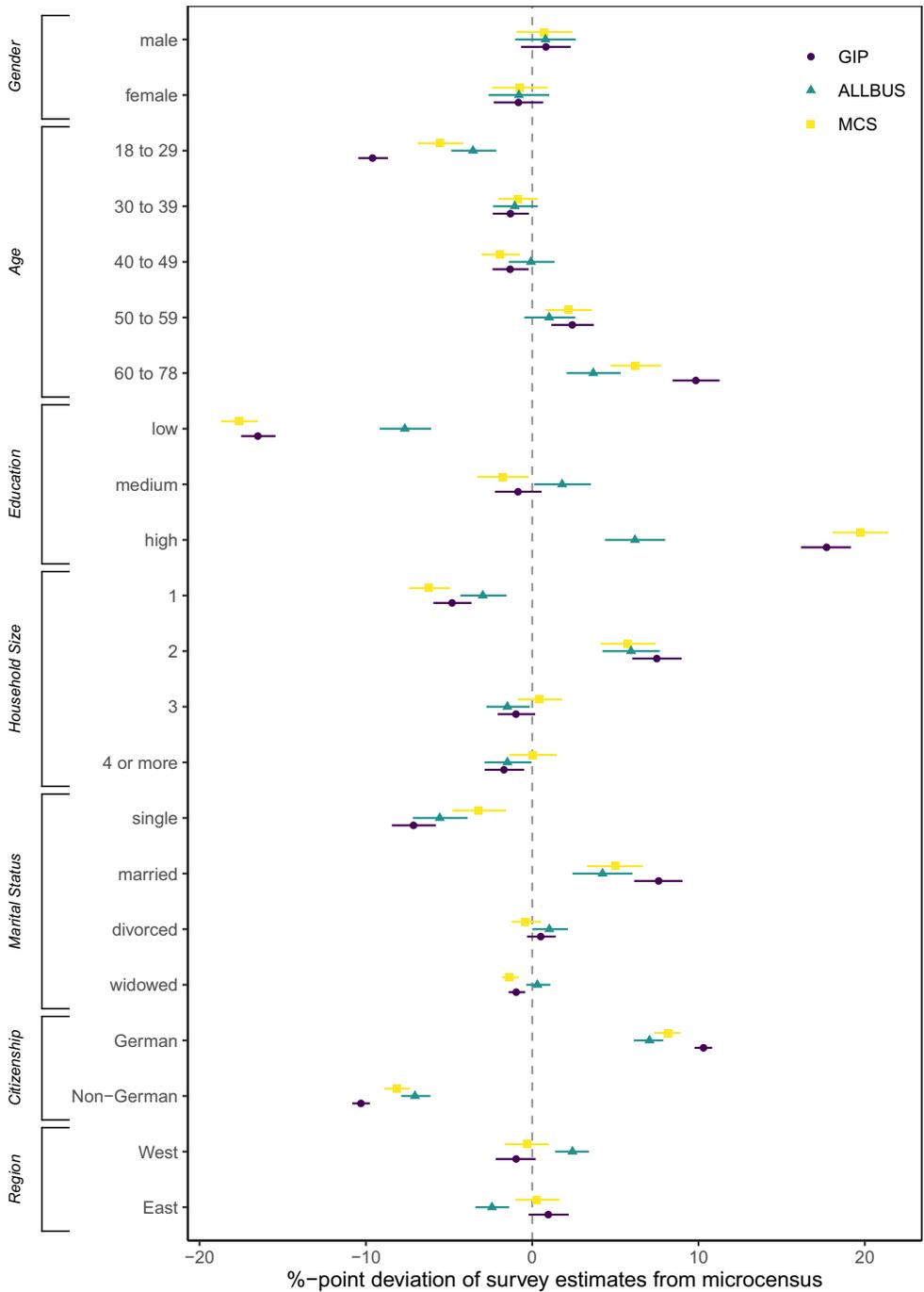
FIGURE 6 Comparison of biases in the German Internet Panel (GIP), ALLBUS and Mannheim Corona Study (MCS)
Note: The circles, triangles and squares display estimates of the deviation of the GIP, ALLBUS and MCS data from the microcensus. The whiskers denote the 95% confidence interval around the estimates. The dashed vertical line represents the point at which a survey estimate would perfectly align with the microcensus values. Survey estimates are design-weighted but not nonresponse-weighted.

of the GIP participants from the MCS. Those GIP participants were never invited to the MCS and were not even informed that the MCS was conducted. Thus, we were able to examine whether response rates in the regular GIP survey waves differed between panelists who were invited to the MCS and those in our control group (Figure 7).

In total, three regular GIP survey waves were conducted during the MCS study period (in March, May and July 2020). In tendency, response rates to the regular GIP survey waves were a little higher in the control group than in the GIP subsample invited to the MCS (control group: 73.0% in March, 71.1% in May and 70.5% in July; MCS group: 69.3% in March, 68.3% in May and 67.9% in July). These response rates are higher than the average MCS response rate (61.1%) but similar to the average GIP response rate in 2019 (72.2%). GIP response rates are likely higher than MCS response rates because GIP respondents have 30 days to respond after each survey invitation while MCS respondents only had 48 h. Only in March is the difference in the response rates between the control group and the MCS group statistically significant (chi2 = 4.3, $p = 0.04$). This might be an artefact of March 2020 being a generally challenging and confusing time for our respondents, given that it coincided with the onset of the pandemic in Germany.

In the three GIP survey waves after the end of the MCS, differences in response rates between the control group and the MCS sample remain statistically insignificant (control group: 71.5% in September, 70.2% in November and 73.3% in January 2021; MCS group: 68.4% in September, 69.5% in November and 70.0% in January 2021). Based on our results, it seems that the impact of
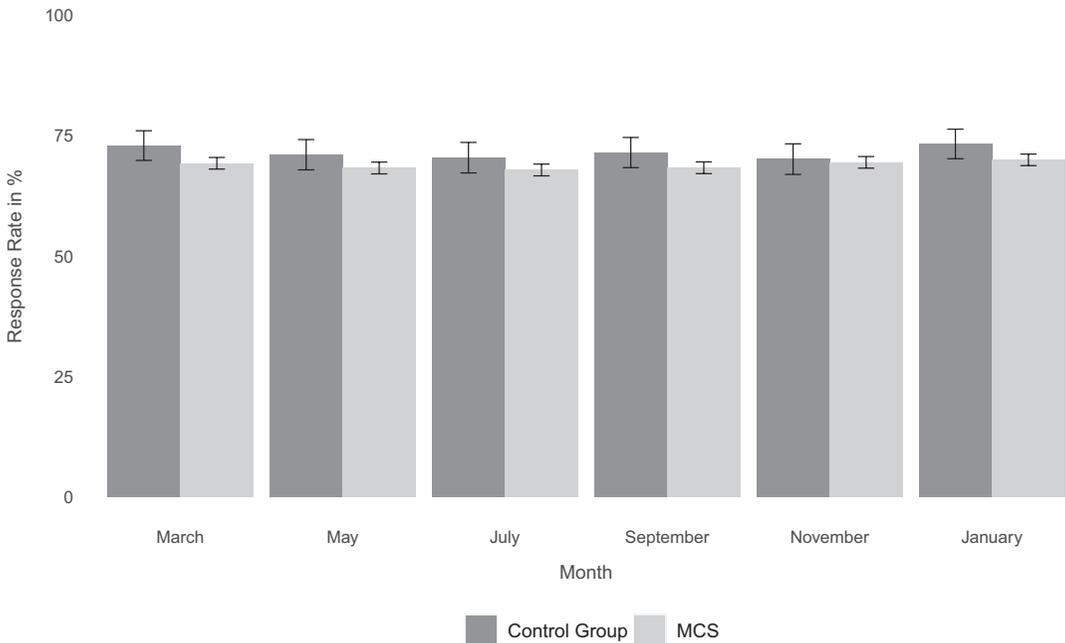

FIGURE 7 Response rates to regular German Internet Panel (GIP) survey waves during and after the Mannheim Corona Study (MCS) study period
Note: The dark grey bars display the response rates to the regular bimonthly GIP survey waves between March 2020 and January 2021 among the random subsample of GIP participants who were not invited to the MCS. The light grey bars display the response rates to the regular GIP survey waves between March 2020 and January 2021 among the random subsample of GIP participants who were invited to the MCS. The shaded whiskers denote the 95% confidence interval around the estimates.

the MCS on the GIP infrastructure is overall negligible, especially given the valuable insights, we gained by conducting the study.

While GIP response rates do not seem to be significantly affected by the MCS, the responses that the MCS participants give to GIP survey questions may be impacted, for example as a result of panel conditioning (e.g. Warren & Halpern-Manners, 2012). To examine this, we compared the answers that MCS participants gave to some of the survey questions in the two GIP waves conducted after the end of the MCS (i.e. in September and November 2020) to the responses given by the GIP participants who were in the MCS control group. For these analyses, we chose political attitudes (satisfaction with democracy, satisfaction with the federal government, political interest, left-right self-placement) as well as an indicator of item nonresponse to the left-right self-placement (which, in the GIP is usually relatively high). These items can generally be expected to be affected by panel conditioning (e.g. Sturgis et al., 2009). Moreover, we include items on whether people installed and used the official German COVID-19 contact tracing app, because these items directly link back to a questionnaire module from the MCS, which asked people whether they intended to install and use this app when it was launched in early July 2020 (for the GIP datasets used in the analyses see Blom et al., 2021a, 2021b, 2021c, 2021d, 2021e, 2021f, 2021g). Because these items are directly related to the content of the MCS, the study may have altered the app-related behaviour of the MCS participants. Our analyses are conducted using bivariate regression models (logits and OLS) with the survey items as the dependent variables and a binary indicator of whether GIP respondents had been invited to the MCS or not as the independent variable (Figure 8).

Overall, we do not find any differences in the responses of the MCS participants and the control group in terms of political attitudes and item nonresponse. However, MCS participants are significantly more likely to state that they installed the COVID-19 tracing app than their counterparts in the control group. This difference cannot be observed for the actual use of the app though. A potential reason for the observed difference in reported app installations may be that MCS participants' awareness of the app and its potential importance was raised by the MCS, thus leading to a higher likelihood of installing the app. However, being included in the MCS may also just have increased the participants' likelihood for socially desirable responding, given that the MCS stressed the severity of the pandemic situation. Future survey waves will include further survey questions which relate back to the content of the MCS. Examining potential differences between MCS participants and the control group on those items may shed more light on the potential influence that conducting the MCS had on GIP survey responses.

Last, we want to note that conducting the MCS of course also had a financial impact on the GIP, which required us to obtain additional funding. For the GIP, the additional fieldwork cost of conducting the MCS was 35,500 Euros plus the 109,392 Euros spent on incentives in the MCS. This is about 15% of the usual annual GIP fieldwork budget. It should be noted that these are rough cost estimates, which, for example, exclude the costs for the GIP and MCS research and survey operations staff as well as the costs for GIP panel maintenance work (for more information on the complexity of calculating survey costs see Olson et al., 2020).

6 | SUMMARY AND DISCUSSION

In this paper, we described how the GIP, an established probability-based online panel infrastructure, was adapted to the need for fast, frequent, and accurate data collection during the early stages of the COVID-19 pandemic in Germany. Between 20 March and 10 July 2020, we
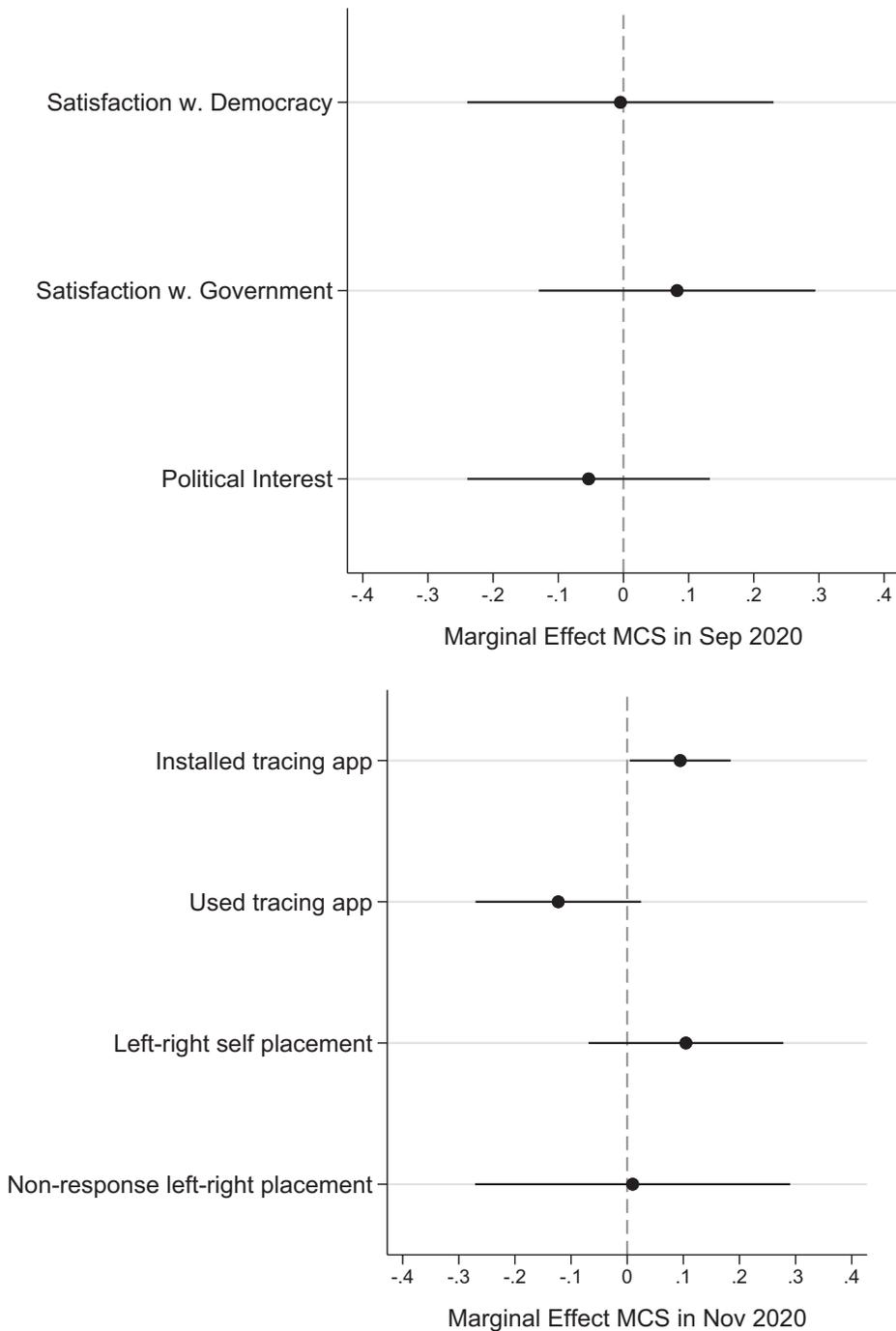
FIGURE 8 Marginal effect of Mannheim Corona Study (MCS) participation on German Internet Panel (GIP) responses

Note: The black dots display the marginal effects of being in the MCS sample versus the control group on the responses to GIP survey questions fielded in September 2020 (satisfaction with democracy, satisfaction with the government, political interest) and November 2020 (installed tracing app, used tracing app, left-right self-placement, item nonresponse to left-right self-placement). The whiskers denote the 95% confidence interval around the estimates. The dashed vertical line indicates no effect.

surveyed nearly 500 people per day in a rotating panel design. The data were processed immediately and reported to interested audiences on a daily basis. The results have informed national policy decisions and were discussed widely in the media. Moreover, the MCS data are used for in-depth research into the societal impact of the COVID-19 pandemic from an interdisciplinary perspective.

We provided examples of the research conducted using MCS data, thus demonstrating the power of the data in terms of three aspects relevant to research on the impact of the pandemic on society: drawing inference to the general population on a daily basis, augmenting survey data with official COVID-19 statistics and linking COVID-19 survey data with data collected prior to the pandemic as well subsequent to the COVID-19 study. In addition, we provided evidence on the stable data quality of the MCS over time as well as in comparison to the full GIP sample and the ALLBUS sample as a gold-standard face-to-face social survey. Furthermore, we showed that conducting the MCS had a negligible impact on the response rates of the GIP as well as a small but noteworthy impact on GIP survey responses in terms of MCS-related content.

Overall, we showed how a probability-based online panel infrastructure can be adapted flexibly to the need for fast, frequent and accurate data collection when unforeseen societal events occur. The adaptability is likely transferable beyond the realm of COVID-19 to other sudden developments, such as stock-market crashes and government dissolutions. However, we do not claim that this approach to data collection is a one-size-fits-all solution for all social research. For example, online data collection remains difficult to conduct in certain subgroups of the population, such as among elderly people or people with low education, at least for the foreseeable future. In this regard, it should be noted, that the GIP age range is limited to ages 16 through 75 at the time of sample members' recruitment to the panel. This feature of the GIP limits the generalizability of our findings, in particular with regard to older adults, who have been especially affected by the pandemic. For the purpose of studying the impact of the pandemic on older people, we recommend using data from the Survey of Health, Ageing and Retirement in Europe (SHARE) Corona Survey (Scherpenzeel et al., 2020). Similarly, the MCS excludes people who do not use the internet or who do not want to use the internet for participating in surveys, which also limits the generalizability of the MCS results, in particular among people with low education. Furthermore, it should be noted that, for probability-based online panel infrastructures to be adaptable, it is necessary to have access to fast and flexible funding opportunities, well-rehearsed and largely automated survey operations processes, and an experienced panel operations team specialized on conducting high-frequency and high-quality data collection.

ORCID

Carina Cornesse http://orcid.org/0000-0002-5437-2999

REFERENCES

Beullens, K., Loosveldt G., Vandenplas C. & Stoop I. (2018) Response rates in the European social survey: increasing, decreasing, or a matter of fieldwork efforts? Survey Methods: insights from the Field. Available from: https://surveyinsights.org?/p=9673 (last accessed 19th September 2020).

Bilgen, I., Dennis, M.J. & Ganesh, N. (2018) Nonresponse follow-up impact on amerispeak panel sample composition and representativeness. Available from http://amerispeak.norc.org/Documents/Research/Bilgen_etal_WhitePaper1_NRFU_SampleComposition_AAPOR_2018.pdf (last accessed 19th September 2020).

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

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

Blom, A.G., 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.

Blom, A.G., Cornesse, C., Friedel, S., Krieger, Fikel, M., Rettig, T., et al. (2020) High-frequency and high-quality survey data collection: the Mannheim corona study. *Survey Research Methods*, 14(2), 171–178.

Blom, A.G., Fikel, M., Friedel, S., Höhne, J.K., Krieger, U., Rettig, T., Wenz, A., SFB 884 'Political Economy of Reforms', Universität Mannheim (2020) German Internet Panel, Wave 45 (January 2020). GESIS Data Archive, Cologne. ZA7642 Data file Version 1.0.0, Available from: https://doi.org/10.4232/1.13588.

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, Cologne. ZA7643 Data file Version 1.0.0, Available from: https://doi.org/10.4232/1.13679.

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 47 (May 2020). GESIS Data Archive, Cologne. ZA7644 Data file Version 1.0.0, Available from: https://doi.org/10.4232/1.13680.

Blom, A.G., Fikel, M., Friedel, S., Krieger, U., Rettig, T., SFB 884 'Political Economy of Reforms', Universität Mannheim. (2021c) German Internet Panel, Wave 48 (July 2020). GESIS Data Archive, Cologne. ZA7645 Data file Version 1.0.0, Available from: 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 (2021d) German Internet Panel, Wave 49 - Core Study (September 2020). GESIS Data Archive, Cologne. ZA7646 Data file Version 1.0.0, Available from: https://doi.org/10.4232/1.13682.

Blom, A.G., Gonzalez Ocanto, M., Fikel, M., Krieger, U., Rettig, T., SFB 884 'Political Economy of Reforms', Universität Mannheim (2021e) German Internet Panel, Wave 50 (November 2020). GESIS Data Archive, Cologne. ZA7647 Data file Version 1.0.0, Available from: https://doi.org/10.4232/1.13734.

Blom, A.G., Cornesse, C., Friedel, S., Krieger, U., Fikel, M., Rettig, T., Wenz, A., Juhl, S., Lehrer, R.., Möhring, K., Naumann, E., Reifenscheid, M., SFB 884 'Political Economy of Reforms', Universität Mannheim (2021f) Mannheim Corona Study. GESIS Data Archive, Cologne. ZA7745 Data file Version 1.0.0, Available from: https://doi.org/10.4232/1.13700.

Blom, A.G., Wenz, A., Cornesse, C., Rettig, T., Fikel, M., Friedel, S. et al. (2021g) Barriers to the large-scale adoption of the COVID-19 contact-tracing app in Germany: survey study. *Journal of Medical Internet Research: JMIR*, 23(3), e23362.

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

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.

Cornesse, C. & Schaurer, I. (2021) Offline population inclusion strategies in probability-based online panels: the impact of providing internet equipment and offering an alternative participation mode. *Social Science Computer Review*, 39(4), 687–704. Available from: https://doi.org/10.1177/0894439320984131

Cornesse, C., Blom, A.G., Dutwin, D., Krosnick, J.A., De Leeuw, E.D., Legleye, S. et al. (2020) A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research. *Journal of Survey Statistics and Methodology*, 8(1), 4–36.

Cornesse, C., Barbara, F., Marina, F., Ulrich, K. & Annelies, B. (2021) Recruiting a probability-based online panel via postal mail: experimental evidence. Social Science Computer Review, Available from: https://doi.org/10.1177/08944393211006059.

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

Couper, M.P., Antoun, C. & Mavletova, A. (2017) Mobile web surveys. In: Biemer, P.P., de Leeuw, E.D., Eckman, S., Edwards, B., Kreuter, F., Lyberg, L.E., Tucker, N.C. & West, B.T. (Eds.) *Total survey error in practice*. Hoboken: John Wiley & Sons, pp. 133–154.

Data Foundation. (2020) COVID impact survey. Available from: https://static1.squarespace.com/static/5e8769b34812765cff8111f7/t/5ea495a78315084d6dc83cd0/1587844524658/Updated_COVID+Impact_Methods_Statement.pdf (last accessed 22nd September 2020).

Destatis. (2020) Haushalte und familien. Was ist der Mikrozensus? Available from: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Haushalte-Familien/Methoden/mikrozensus.html (last accessed 25th September 2020).

Dutwin, D. & Buskirk, T. (2020) Telephone sample surveys: dearly beloved or nearly departed? Trends in survey errors in the era of declining response rates. *Journal of Survey Statistics and Methodology (advance access)*. Available from: https://doi.org/10.1093/jssam/smz044 (last accessed 29th September 2020).

Dymond-Green, N. (2020) COVID-19: social surveys are now more important than ever. Available from: http://blog.ukdataservice.ac.uk/covid-19-social-surveys/ (last accessed 18th September 2020).

Federal Government, The (2020) Coronavirus warning app. Available from: https://www.bundesregierung.de/breg-de/themen/corona-warn-app/corona-warn-app-englisch (last accessed 24th September 2020).

Federal Ministry of Labour and Social Affairs (2020) Questions and answers relating to short-time work (Kurzarbeit) and skills development. Available from: https://www.bmas.de/SharedDocs/Downloads/DE/kug-faq-kurzarbeit-und-qualifizierung-englisch.pdf?__blob=publicationFile&v=2 (last accessed 29th September 2020).

Federal Ministry of Health (2020) Information regarding Covid-19 in Germany. Available from: https://www.zusammengegencorona.de/en/ (last accessed 26th September 2020).

GALLUP (2020) How does the Gallup Panel work? Available from: https://www.gallup.com/174158/gallup-panel-methodology.aspx (last accessed 19th September 2020).

GESIS - Leibniz-Institut für Sozialwissenschaften (2019) German General Social Survey - ALLBUS 2018. GESIS Datenarchiv, Küln. ZA5272 Datenfile Version 1.0.0, Available from: https://doi.org/10.4232/1.13325.

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

Gummer, T., Schmiedeberg, C., Bujard, M., Christmann, P., Hank, K., Kunz, T. et al. (2020) The impact of Covid-19 on fieldwork efforts and planning in pairfam and FReDA-GGS. *Survey Research Methods*, 14(2), 223–227.

Haan, M., Ongena, Y.P., Vannieuwenhuyze, J.T.A. & De Glopper, K. (2017) Response behavior in a video-web survey: a mode comparison study. *Journal of Survey Statistics and Methodology*, 5(1), 48–69.

Hunsaker, A. & Hargittai, E. (2018). A review of Internet use among older adults. *New Media & Society*, 20(10), 3937–3954.

H⊘gest⊘l, A. & Skjervheim, Ø. (2014) Norwegian citizen panel. 2013, first wave. Methodology report. Available from: https://www.uib.no/sites/w3.uib.no/files/attachments/ncp-methodology-report-wave-1.pdfAccessed (last accessed 25th September 2020).

Ipsos (2020) Knowledge Panel. Ensuring accuracy with an industry leading representative random sample, exceptional response rates and coverage. Available from: https://www.ipsos.com/en-us/solutions/public-affairs/knowledgepanel (last accessed 19th September 2020).

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

Jessop, C. (2017) Developing the NatCen panel. August 2015 - July 2017. Available from: https://www.natcen.ac.uk/media/1484228/Developing-the-NatCen-Panel-V2.pdf (last accessed 19th September 2020).

Kaczmirek, L., Phillips, B., Pennay, D., Lavrakas, P.J. & Neiger, D. (2019) Building a probability-based online panel: life in Australia. Available from: https://www.srcentre.com.au/our-research/methods-research/Building%20

a%20probability-based%20online%20panel-Life%20in%20Australia%20-%202.0.pdf (last accessed 19th September 2020).

Kaminska, O. & Lynn, P. (2017) The implications of alternative allocation criteria in adaptive design for panel surveys. *Journal of Official Statistics*, 33(3), 781–799.

Kapteyn, A., Angrisani, M., Bennett, D., Bruine de Bruin, W., Darling, J., Gutsche, T. et al. (2020) Tracking the Effect of the COVID-19 Pandemic on the Lives of American Households. *Survey Research Methods*, 14(2), 179–186.

Kolb, P., Bruch, C., Sand, C. & Weyandt, K. (2021). GESIS panel data manual related to ZA5664 and ZA5665 updated with Wave hc (July / August 2020) Version 39-0-0. Available from: https://dbk.gesis.org/dbksearch/download.asp?id=63527 (last accessed 19th June 2021).

Kreuter, F., Presser, S. & Tourangeau, R. (2008) Social desirability bias in CATI, IVR, and web surveys: the effects of mode and question sensitivity. *Public Opinion Quarterly*, 72(5), 847–865.

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.

Kunz, T. & Fuchs, M. (2019) Using experiments to assess interactive feedback that improves response quality in web surveys. In: Lavrakas, P.J., Traugott, M.W., Kennedy, C., Holbrook, A.L., de Leeuw, E.D. & West, B.T. (Eds.) *Experimental methods in survey research: techniques that combine random sampling with random assignment*, pp. 247–273.

Lewis, N. Jr. (2020) Why coming up with effective interventions to address COVID-19 is so hard. Available from: https://fivethirtyeight.com/features/why-coming-up-with-effective-interventions-to-address-covid-19-is-so-hard/ (last accessed 17th September 2020).

Luiten, A. & Schouten, B. (2013). Tailored fieldwork design to increase representative household survey response: an experiment in the Survey of Consumer Satisfaction. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(1), 169–189.

Lynn, P., Häder, S., Gabler, S. & Laaksonen, S. (2004) Methods for achieving equivalence of samples in cross-national surveys: the European social survey experience. ISER Working Paper Series, No. 2004–09. Available from: https://www.econstor.eu/bitstream/10419/92131/1/2004-09.pdf (last accessed 18th September 2020).

Malhotra, N. & Krosnick, J.A. (2007) The effect of survey mode and sampling on inferences about political attitudes and behavior: comparing the 2000 and 2004 ANES to internet surveys with nonprobability samples. *Political Analysis*, 15(3), 286–323.

Mata, J., Wenz, A., Rettig, T., Reifenscheid, M., Möhring, K., Krieger, U. et al. (2020) Health behaviors and mental healthduring the COVID-19 pandemic: a longitudinal population-based survey [Preprint]. Available from: https://psyarxiv.com/qbgh7/ (last accessed 25th September 2020).

McClain, C.A., Couper, M.P., Hupp, A.L., Keusch, F., Peterson, G., Piskorowski, A.D., et al. (2019) A typology of web survey paradata for assessing total survey error. *Social Science Computer Review* 37(2), 196–213.

Möhring, K., Naumann, E., Reifenscheid, M., Blom, A.G., Wenz, A., Rettig, T. et al. (2020) Inequality in employment during the Corona lockdown: evidence from Germany. Available from: https://uni-tuebingen.de/en/faculties/faculty-of-economics-and-social-sciences/subjects/department-of-social-sciences/ifp/institute/people/comparative-public-policy-professor-seeleib-kaiser/journal-of-european-social-policy/jesp-european-social-policy-blog/newsfullview-jesp/article/inequality-in-employment-during-the-corona-lockdown-evidence-from-germany/ (last accessed 25th September 2020).

Naumann, E., Möhring, E., Reifenscheid, M., Wenz, A., Rettig, T., Lehrer, R. et al. (2020) COVID-19 policies in Germany and their social, political and psychological consequences. European Policy Analysis. Available from: https://onlinelibrary.wiley.com/doi/epdf/10.1002/epa2.1091 (last accessed 28th September 2020)

NORC at the University of Chicago (2019). Panel design: how AmeriSpeak households are sampled. Available from: https://amerispeak.norc.org/about-amerispeak/Pages/Panel-Design.aspx (last accessed 19 September 2020).

Olson, K.M., Wagner, J. & Anderson, R. (2020). Survey costs: where are we and what is the way forward?. *Journal of Survey Statistics and Methodology*, https://doi.org/10.1093/jssam/smaa014.

Prior, G. 2020. I hear you knocking... Available from: https://www.natcen.ac.uk/blog/i-hear-you-knocking%E2%80%A6 (last accessed 18 September 2020).

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

Robert Koch Institut, RKI (2020a) COVID-19: Fallzahlen in Deutschland und weltweit. Available from: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html (last accessed 25th September 2020).

Robert Koch Institut, RKI (2020b) Local health authority day 2020. Available from: https://www.rki.de/EN/Content/Institute/local_health_authority_day/lhad_node.html (last accessed 26th September 2020).

Robert Koch Institut, RKI (2020c) Coronavirus disease 2019 (COVID-19). Daily situation report of the Robert Koch Institute. Available from: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-08-19-en.pdf?__blob=publicationFile (last accessed 26th September 2020).

Schaurer, I. (2017). Recruitment strategies for a probability-based online panel: effects of interview length, question sensitivity, incentives and interviewers. [Doctoral thesis] Available from: https://madoc.bib.uni-mannheim.de/42976/1/Dissertation_Schaurer-1.pdf (last accessed 19 September 2020).

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

Scherpenzeel, A., Axt, K., Bergmann, M., Douhou, S., Oepen, A., Sand, G. et al. (2020). Collecting survey data among the 50+ population during the COVID-19 outbreak: the Survey of Health, Ageing and Retirement in Europe (SHARE). *Survey Research Methods*, 14(2), 217–221.

Smith, T.W., Jibum K., Achim K. & Alison, P. (2006) Social-science research and the general social surveys. *Comparative Sociology*, 5(1), 33–43.

Schnaudt, C., Weinhardt, M., Fitzgerald, R. & Liebig, S. (2014) The European Social Survey: contents, design, and research potential. *Journal of Contextual Economics*, 134(4), 487–506.

Schouten, B., Shlomo, N. & Skinner, C. (2011). Indicators for monitoring and improving representativeness of response. *Journal of Official Statistics*, 27(2), 231–253.

Statista (2020) Laborkapazität für tests auf das Coronavirus (COVID-19) in deutschland bis KW 38 2020. Available from: https://de.statista.com/statistik/daten/studie/1110951/umfrage/testkapazitaeten-fuer-das-coronavirus-covid-19-in-deutschland/ (last accessed 27th September 2020).

Sturgis, P., Kuha, J., Baker, N., Callegaro, M., Fisher, S., Green, J. et al. (2018) An assessment of the causes of the errors in the 2015 UK general election opinion polls. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 181(3), 757–781.

Sturgis, P., Allum, N. & Brunton-Smith, I. (2009). Attitudes over time: the psychology of panel conditioning. In: Lynn, P. (Ed.) *Methodology of longitudinal surveys*, Hoboken: John Wiley & Sons, pp. 113–126.

Travis, A. (2017) Can we still trust opinion polls after 2015, Brexit and Trump? Available from: https://www.theguardian.com/politics/2017/may/08/opinion-polls-general-election (last accessed 19 September 2020).

University of Southern California (USC, 2017) Understanding America Study. Available from: https://uasdata.usc.edu/index.php (last accessed 21 September 2020).

University of Oxford (2020) Oxford Supertracker. The Global Directory for COVID Policy Trackers and Surveys. Surveys. Available from: https://supertracker.spi.ox.ac.uk/surveys/ (last accessed 22 September 2020).

van Tilburg, T.G., Steinmetz, S., Stolte, E., van der Roest, H. & de Vries, D. H. (2020) Loneliness and mental health during the COVID-19 pandemic: a study among Dutch older adults. *The Journals of Gerontology: Series B*, gbaa111.

Warren, J.R. & Halpern-Manners, A. (2012). Panel conditioning in longitudinal social science surveys. *Sociological Methods & Research*, 41(4), 491–534.

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

Williams, D. & Brick, J.M. (2018) Trends in US face-to-face household survey nonresponse and level of effort. *Journal of Survey Statistics and Methodology*, 6(2), 186–211.

Yeager, D.S., Krosnick, J.A., Chang, L., Javitz, H.S., Levendusky, M.S., Simpser, A., et al. (2011) Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*, 75(4), 709–747.

APPENDIX A

TABLE A1 Comparison of sample composition in the Mannheim Corona Study (MCS), German Internet Panel (GIP), ALLBUS and microcensus among the German population aged 18–78

| Category	MCS	GIP	ALLBUS	MZ
Gender				
Female	49.13%	49.04%	49.07%	49.87%
Male	50.87%	50.96%	50.93%	50.13%
Age				
18 to 29	13.97%	9.94%	15.98%	19.54%
30 to 39	16.29%	15.86%	16.12%	17.17%
40 to 49	15.15%	15.74%	17.00%	17.10%
50 to 59	23.96%	24.20%	22.82%	21.80%
60 and older	30.63%	34.26%	28.08%	24.42%
Education				
Low	12.61%	13.76%	22.60%	30.76%
Medium	30.49%	31.43%	34.08%	33.10%
High	56.90%	54.81%	43.32%	36.00%
Household size				
1	16.79%	18.17%	20.02%	23.00%
2	42.95%	44.68%	43.13%	37.19%
3	18.70%	17.29%	16.78%	18.27%
4 or more	21.56%	19.86%	20.07%	21.55%
Marital status				
Single	30.70%	26.77%	28.35%	33.91%
Married	58.70%	61.30%	57.94%	53.71%
Divorced	8.15%	9.06%	9.57%	8.55%
Widowed	2.45%	2.87%	4.14%	3.83%
Citizenship				
German	94.63%	96.79%	93.50%	86.46%
Other	5.37%	3.21%	6.50%	13.55%
Region				
East	19.79%	20.47%	17.08%	19.50%
West	80.21%	79.53%	82.92%	80.50%
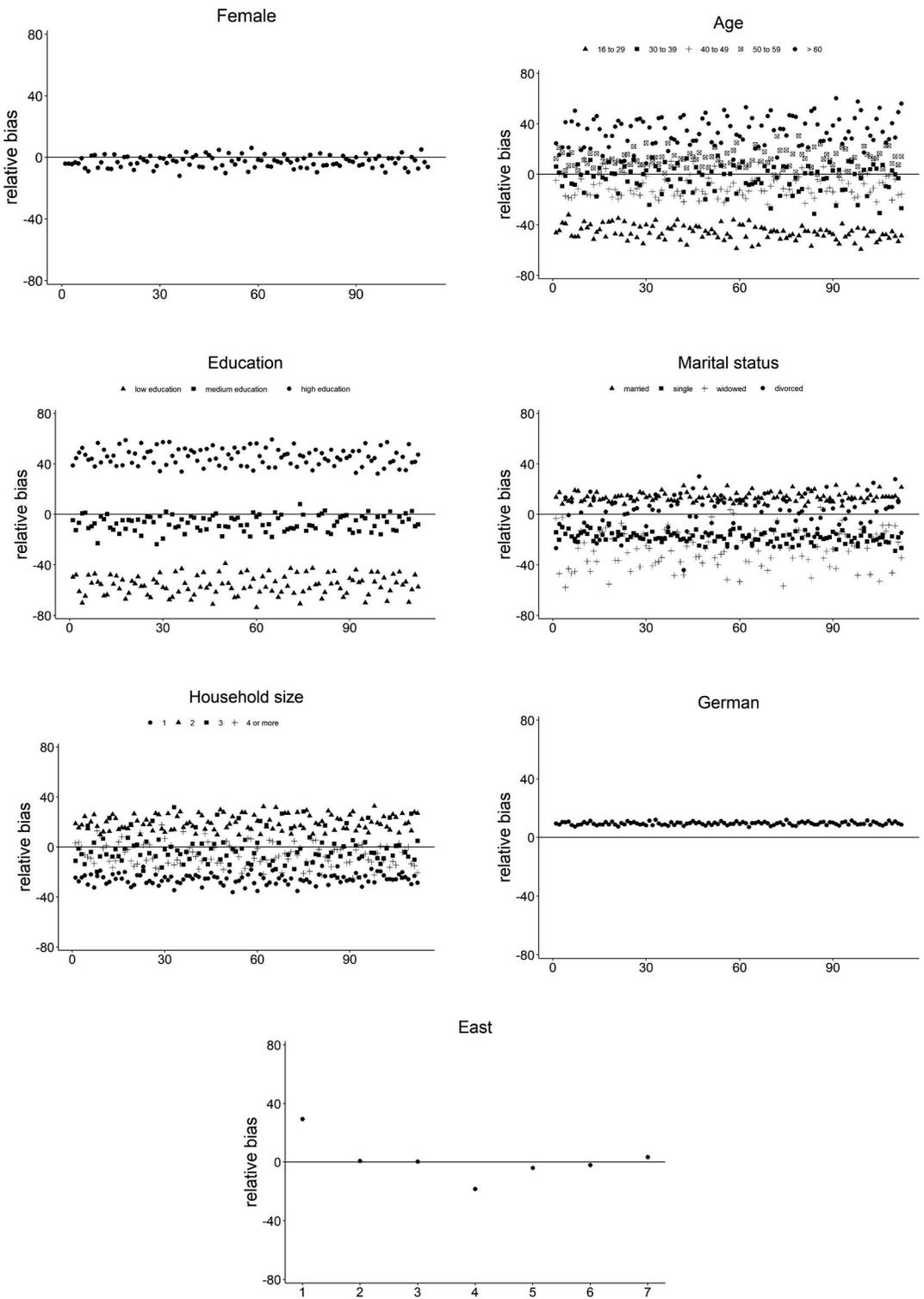
FIGURE A1 Relative biases over days by socio-demographic variable variable