Setting Up an Online Panel Representative of the General Population: The German Internet Panel

Annelies G. Blom¹,³, Christina Gathmann², and Ulrich Krieger³

Abstract
This article looks into the processes and outcomes of setting up and maintaining a probability-based longitudinal online survey, which is recruited face-to-face and representative of both the online and the offline population aged 16–75 in Germany. This German Internet Panel studies political and economic attitudes and reform preferences through bimonthly longitudinal online interviews of individuals. The results presented here demonstrate that a carefully designed and implemented online panel can produce high-quality data at lower marginal costs than existing panels that operate solely in a face-to-face mode. Analyses into the representativeness of the online sample showed no major coverage or nonresponse biases. Finally, including offline households in the panel is important as it improves the representation of the older and female segments of the population.

¹ School of Social Sciences, University of Mannheim, Mannheim, Germany
² University of Heidelberg, IZA and CESifo, Heidelberg, Germany
³ Collaborative Research Center “Political Economy of Reforms” (SFB884), University of Mannheim, Mannheim, Germany

Corresponding Author:
Annelies G. Blom, School of Social Sciences, University of Mannheim, 68131 Mannheim, Germany.
Email: blom@uni-mannheim.de
Keywords
online panel, probability sample, representativeness, offline households, attrition

Introduction
Inferential statistics teach us that to infer from a sample to the population we need a random probability sample. In online survey research, however, volunteer access panels, where respondents self-select into the sample, dominate the landscape. Such panels are attractive due to their low costs; yet, in recent years we have seen increasing debates about their quality in terms of representativeness and measurement error (e.g., Yeager et al. 2011). In this article, we describe the design, sampling, recruitment, and maintenance of a probability-based longitudinal online survey of the general population aged 16–75 living in private households—to wit, the German Internet Panel (GIP). Internationally, the GIP is one among very few online panels that are both based on a gross sample of the general population and include individuals who previously had no or limited access to the Internet (i.e., who would not have been represented in any other online panel). To our knowledge, only two such academically based online panels are currently operating: the LISS Panel in the Netherlands (www.lissdata.nl) and the ELIPSS panel in France (www.elipss.fr). Furthermore, GESIS is currently setting up such a new panel in Germany (www.gesis-panel.org), and KnowledgeNetworks in the United States maintains a similar commercial panel (www.knowledgenetworks.com/knpanel).

Given the novelty of this mode of survey and its potential for the future of survey research, we would like to share key insights into the recruitment and maintenance of one of these probability-based offline-recruited online panels that include previously offline respondents. In the following sections, we describe how the sample for the GIP was drawn, how recruitment was conducted, how sample members who were previously offline were included in the GIP, and how, once recruited, the panel members are interviewed online and are kept motivated. All processes were pretested on a small-scale sample and optimized according to experiences from these pretests. Finally, we provide a first description of sample composition and potential biases.

Recruiting the GIP: Methodology and Results
Sampling
The GIP is based on a three-stage probability sample. In the first stage, we sampled 250 primary sampling units (PSUs) situated in 208 local
administrative units. The drawing of PSUs was stratified by state, govern-
mental district, and level of urbanicity.

During the second sampling stage, interviewers listed 100 households
along a random route with random starting point at each PSU. To prevent
errors and interviewer cheating and to enable detailed checking of the listed
addresses, every household along the predefined route was listed. Different
interviewers conducted the listing and the actual interviewing at all sample
points. The listing yielded a total of 25,000 households, which constituted
the sampling frame in the third sampling stage.

To minimize clustering, 22 addresses per PSU were drawn at set inter-
vals with a random start. The resulting gross sample consisted of 5,500
addresses, of which 622 were found to be ineligible during fieldwork
(American Association for Public Opinion Research [AAPOR] 2011), such
as vacant or commercial housing units.

Recruitment in Two Phases

Recruitment into the GIP online was conducted in two steps: a face-to-face
interview phase and a phase of invitations to the online panel, including
sending out initial invitation letters, reminder letters, and phone reminders
to households in which nobody had registered online yet. Figure 1 illustrates
the complex recruitment process.

Face-to-face recruitment. All sampled addresses were approached by interviewers
who wanted to make contact with the household and conduct a short face-to-face
interview with a household member. A total of 135 interviewers worked on the
GIP face-to-face recruitment. Interviewers were trained face-to-face by the sur-
vey agency and the GIP team in one of the three one-day training sessions.

All sampled households received an advance letter addressed to the house-
hold or, if no household name was available, to the tenant. The advance letter
announced the face-to-face interview as part of the study “Gesellschaft im
Wandel” (Changing Society). The letter was kept very short (Appendix A) and
was accompanied by a one-page color print with additional information
(Appendix B) and a data protection leaflet (Appendix C). All materials iden-
tified the University of Mannheim as the sponsor of the survey. Letters were
sent out approximately one week before interviewers started work in the PSU.

Interviewers conducted short interviews at the household asking about the
household composition (name and year of birth of each household member),
general demographics, simple attitudinal questions, and whether the house-
hold had a computer and a broadband Internet connection. At the end of the
interview, the interviewer asked the respondent for permission to have the University of Mannheim send letters with further information about the online study to all household members born in the years 1937–1996 (i.e., aged approximately 16–75). Households with at least one household member within this age group but without a broadband Internet connection and/or a computer were informed that they were invited to participate in the online study and that, if they agreed, someone would call them in the next week to make an appointment to install Internet and, if no suitable computer was available, a BenPC. Interviewers carried with them materials describing the online survey (Appendix D) to be used during this final phase of the interview. If the interviewer, because of imminent refusal, saw no possibility of conducting a full interview at the household, they were allowed to conduct a short doorstep interview instead, asking only five key questions.

Overall, this phase of face-to-face recruitment yielded 2,121 full interviews, resulting in a response rate of 43.4% (AAPOR RR1 [response rate 1 as defined by AAPOR 2011]). In addition, 422 doorstep interviews were conducted. Including the doorstep interviews, the response rate was 52.5%
These response rates seem rather low by international standards, however, in the German survey climate, where large-scale social surveys such as the Socio Economic Panel (SOEP) and the General Social Survey (ALLBUS) typically achieve response rates of only around 30%, these rates can be considered high.

**Invitations to the online panel.** The list of household members eligible to become members of the online panel was processed weekly by the GIP team. Eligible persons, who had a computer and a broadband Internet connection at home, were sent an invitation letter (Appendix E) to the online panel. This letter contained further information about the study together with a login name and a password. In addition, it contained a small card (credit card sized) with the login details and hotline numbers. Households without Internet access and/or computer were also invited to the study and sent login details. Furthermore, they were informed that someone would call them to make an appointment to install the Internet and/or a BenPC. Households where a doorstep interview had been conducted received an invitation to the online panel addressed to the household. Upon their first online registration, these households were asked about the household composition and additional personal invitations were sent to the other eligible persons in the same household.

Within the 2,121 households where full face-to-face interviews had been conducted, 3,775 eligible individuals were identified, an average of 1.78 eligible persons per household. We got permission to send invitations to the online survey for 3,119 (82.6%) eligible persons. This includes persons living in households without Internet access and/or a computer. A total of 1,578 eligible persons registered for the online panel (AAPOR RR1 41.8%). In addition, 25 persons from households that conducted a doorstep interview registered online (AAPOR RR4 3.3%, assuming 1.78 eligible persons per household). Taking all recruitment stages together, the overall response rate is 18.5%. To find out more about potential selectivities in the panel due to nonresponse, we conducted bias analyses, which are summarized in the section on Representativeness.

**offline Households in the Online Panel**

Equipping previously offline households with a BenPC and/or Internet is a key aspect of the GIP to ensure representativeness of the collected data. So, we did several things to encourage the recruitment of offline households. First, the online panel was not mentioned in the advance letter to prevent households with little computer and Internet affinity to drop out from the
start. Second, interviewers were especially trained to react to the concerns of offline households. They were equipped with materials about the BenPC (the most essential materials are found in Appendix F) to show to the respondents during the interview. Third, the installation of the BenPCs and/or routers was subcontracted to a company that employs computer engineers all over Germany. The computer engineers were trained via video and written materials about the specific challenges of communicating technical aspects to clients with little or no prior experience with computers. Fourth, the local computer engineers made an appointment with the offline respondents and installed the equipment on-site. There, they showed respondents how the equipment worked and how they could fill in the bimonthly questionnaires. Finally, a hotline at the survey organization forwards queries from these households to the information technology company.

Within the 2,121 households where full face-to-face interviews had been conducted, 528 were identified as offline households (24.8%). In these 528 households, 883 eligible sample members were identified, of whom 487 agreed to receive further information about the online panel and to be provided with the necessary equipment to participate. Ultimately, 128 previously offline respondents received equipment and registered for the online panel (AAPOR RR1 14.5%). Thus, the recruitment rate in offline households was considerably lower than in online households.

With respect to key sociodemographic characteristics—age and gender—our analyses show, however, that panel members from online and offline households significantly differ in terms of these characteristics. Their contribution to the representativeness of the panel is discussed subsequently.

**Incentives during Recruitment**

Since the GIP was the first study in Germany to recruit members for an online panel based on face-to-face fieldwork and a probability sample, the most effective incentive strategies had yet to be researched. Moreover, paying incentives unconditionally is rare in German surveys, although it is a widely accepted and well-tested practice in other countries. As part of the survey design of the GIP, we thus implemented two incentives experiments to investigate which strategy maximizes panel response.

The first incentives experiment was conducted during the face-to-face phase. Households where the listing of addresses had yielded one or several household names (i.e., where the advance letter could be addressed directly to the household) were allocated to one of the two experimental conditions. These were an €5 unconditional cash incentive (i.e., the advance letter
contained an €5 bill) or an €10 conditional cash incentive (i.e., the advance letter contained the information that they would receive €10 in cash if they participated in the face-to-face interview). The value of the incentives was chosen such that the costs for the GIP research team would be independent of the incentive, assuming a 50% response rate in the face-to-face interview (including doorstep interviews). We chose this experimental design because we wanted to allocate the available budget for incentives in a way that maximizes the response rate.

With an 8.9% points difference, the unconditional incentives yielded a significantly higher response rate in the face-to-face household interviews ($t = 5.14$). While 50.8% of households with an unconditional incentive responded to the face-to-face recruitment interview, 41.9% of the conditionally incentivized households responded (AAPOR RR1; i.e., not counting doorstep interviews). Moreover, and despite the time lag between face-to-face interviews and the invitations to the online interviews, this effect carries over to and is reinforced at individuals’ online registration for the panel. While 33.8% of the eligible persons registered online when the household had been incentivized for the face-to-face interview with €5 unconditionally, 31.3% of the eligible persons registered for the online interviews when incentivized with €10 conditionally ($t = 2.06$). Households where the name(s) of the inhabitant(s) was not identified during the listing received lower response rates than each of the two experimental groups. Households sampled during the second field phase were excluded from the analyses, since they all received the unconditional incentive.

In the second incentives experiment, we studied the effect of an €5 unconditional incentive in the first mail reminder versus no incentive on the individual-level registration rate to the online study. Given the special two-stage nature of the GIP, the effects of this second-stage incentive have not been empirically studied in the literature. However, given that a variety of studies have found that providing an incentive yields higher response rates than not providing any (e.g., Singer and Ye 2013), we expected a positive incentive effect at this second recruitment stage.

Our analyses showed that the €5 incentive in the reminder letter had a significant effect on the online registration rate. While 30.0% of the cases who received a reminder incentive registered online within two weeks of receiving the reminder letter, 13.7% of the cases without incentive registered online within this time period ($t = 7.27$).

The two experiments were also crossed and their combined effect analyzed. The analysis is restricted to those persons who were part of the recruitment incentive experiment ($N = 3,900$), who responded to the
face-to-face interview, who could be invited to the online panel, and who
received a reminder letter. The sample sizes of the four experimental sub-
groups are displayed in Table 1.

Within both recruitment incentives groups, the incentive in the reminder
letter had a significant effect. Among those respondents who received an
€5 unconditional incentive during the recruitment interview, the €5 incentive
in the reminder letter is associated with an 11.3 percentage points higher
online registration rate ($t = -2.2$). Among respondents who received an
€10 conditional incentive during the recruitment interview, the incentive in
the reminder letter is associated with a registration rate that is 14.8 percentage
points higher ($t = -4.0$) compared to when no reminder incentive was sent.

Representativeness

As a probability-based face-to-face recruited online panel including the
online and offline population, the GIP strives for a quality that is comparable
to the quality of the established face-to-face surveys in Germany, such
as the SOEP or the ALLBUS (see also Bosnjak et al. 2013). The literature
shows that the online mode of data collection is of little concern to data
quality in terms of measurement errors but is of concern in terms of representativeness when recruited as nonprobability panels (e.g., Yeager et al. 2011). To investigate the GIP’s representativeness, we compared GIP respondents in the first wave of online data collection to nonrespondents and population statistics. Data on the complete gross sample and the population stemmed from two data sources (1) linked auxiliary street-level data from a commercial provider (microm) for the gross sample and (2) population statistics from the 2011 census.

Logistic regression analyses with the linked commercial data showed no
effects of urbanicity, region (East/West Germany), age, household type, and
unemployment rate in responses to the online panel. Only three characteristics

### Table 1. The Combined Effect of Two Incentives Experiments.

<table>
<thead>
<tr>
<th>Recruitment Incentive</th>
<th>Reminder Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>€5 Unconditional</td>
<td></td>
</tr>
<tr>
<td>€10 Conditional</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Reminder sent</td>
<td>Reminder sent</td>
</tr>
<tr>
<td>$n$</td>
<td>259</td>
</tr>
<tr>
<td>Percentage</td>
<td>104</td>
</tr>
<tr>
<td>Registration</td>
<td>Registration</td>
</tr>
<tr>
<td>$n$</td>
<td>79</td>
</tr>
<tr>
<td>Percentage</td>
<td>30.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total $n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>950</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total $n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>235</td>
</tr>
</tbody>
</table>
were found to be significant predictors of response: level of education (the online panel overrepresents those living in areas with a high proportion of university graduates), purchasing power (people living in more prosperous areas were more likely to participate), and immigration (areas with higher proportions of immigrants were underrepresented; Krieger and Blom 2013).

Comparing the GIP sample to population distributions of the German census conducted in 2011, we further found that the GIP underrepresents the oldest age groups, while the representation of men and women is quite exact. As Table 2 shows, however, including previously offline respondents in the panel improves the representativeness of the GIP with regard to both age and gender.

### Maintaining the GIP: Methodology and Results

As a longitudinal survey, the GIP not only aspires to recruit a representative and sizable sample, it also aims for a high level of retention throughout the future waves. In this section, we therefore describe the measures taken to ensure high retention rates and report retention rates during the first five waves of the GIP.

#### Panel Maintenance

The literature on attrition typically distinguishes three main components (Lepkowski and Couper 2002), namely, failure to locate the sample unit, failure to make contact, and failure to gain cooperation. Locating the sample unit is of concern especially in face-to-face panel surveys, where geographic mobility of all or part of a household might lead to a loss of respondents (see Fitzgerald et al. 1998; Zabel 1998). The risk of failure to locate a

---

Table 2. The GIP Sample Compared to the German Population Statistics (%).

<table>
<thead>
<tr>
<th></th>
<th>Census</th>
<th>GIP Sample Online</th>
<th>GIP Sample Offline</th>
<th>GIP Sample Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25–29</td>
<td>9.1</td>
<td>12.1</td>
<td>6.2</td>
<td>11.6</td>
</tr>
<tr>
<td>30–39</td>
<td>17.9</td>
<td>17.6</td>
<td>12.4</td>
<td>17.1</td>
</tr>
<tr>
<td>40–49</td>
<td>25.2</td>
<td>27.3</td>
<td>20.4</td>
<td>26.7</td>
</tr>
<tr>
<td>50–64</td>
<td>30.8</td>
<td>33.9</td>
<td>42.5</td>
<td>34.6</td>
</tr>
<tr>
<td>65–74</td>
<td>17.1</td>
<td>9.2</td>
<td>18.6</td>
<td>9.9</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>48.8</td>
<td>51.3</td>
<td>38.3</td>
<td>50.2</td>
</tr>
<tr>
<td>Female</td>
<td>51.2</td>
<td>48.8</td>
<td>61.7</td>
<td>49.8</td>
</tr>
</tbody>
</table>

Note: Since the youngest age group published in the census statistics is 15–24, while in the GIP the youngest participants were 16, we do not provide a comparison for this group. GIP = German Internet Panel.
sample unit is lower in online panels because the primary contact is through e-mail addresses, which typically remain stable as people move, and because of the high frequency of interviews (e.g., bimonthly in the GIP as compared to once a year in the SOEP). Furthermore, GIP members can change the e-mail address where they receive invitations through their account on the study website and through the hotline.

Failure to make contact might, however, also be a problem in the GIP. For example, if a sample unit’s mailbox is full or if spam filters redirect our invitation, e-mails might not reach a potential respondent. In the GIP, these potential problems are counteracted in two ways. First, the e-mail invitation is just one of the two routes through which a sample unit may access the questionnaire; the other is a direct access via the study’s website with login name and password. Our questionnaires are fielded regularly every two months and are always made available on the first day of the uneven months. Therefore, even if the invitation e-mail does not reach our panel members, they can find their way to the questionnaire via the study website. Second, if a panel member has failed to participate in the panel for two consecutive waves or our invitation e-mail bounces, they are called by the survey organization to encourage their renewed participation and to inquire whether there were any technical problems preventing them to participate in the waves.

Finally, failure to gain cooperation may be the greatest concern for longitudinal online surveys. In the GIP, several additional measures are taken to maintain motivation. By motivating sample members, the survey can be made more salient to them (Groves et al. 2000), for example, by providing information about the survey results. In the GIP, we feedback study results bimonthly during even months (i.e., during those months when panel members are not interviewed). Short descriptive reports of results from previous waves are made available through the study website together with a short introduction of the researchers conducting the research. Respondents are alerted to these via e-mail. In this way, we can personalize our contacts with the panel members.

Another form of personalization is to give panel members several means of feeding back their experiences with each interview. The hotline at the survey organization can be reached via e-mail and telephone (a toll-free number). When the participant voices serious concern, these are forwarded to the principal investigator of the GIP, who carefully answers each query. Furthermore, the panel members are asked for feedback at the end of each questionnaire, both in closed rating questions and by providing an open question for more general feedback.

Persistence is another way to gain cooperation from panel members. In addition to the initial invitation e-mail at the start of fieldwork, we send out a first reminder e-mail after approximately one week, a second reminder
e-mail after another week, and attempt contact by phone in the fourth week if the panel member missed two consecutive waves.

Finally, we use monetary incentives at each wave to show the panel members that we value their participation. For each interview of 20–25 minutes, each respondent receives €4 with an annual bonus of €5, if panel members participate in all but one interview, and a bonus of €10, if they participate in all the interviews of that year. As far as attrition is concerned, research shows that cash incentives are more effective than vouchers or incentives in kind (see Booker et al. 2011 for a review). For operational reasons, however, we are not able to send letters with cash to panel members. Instead, when registering for the panel, members are asked whether they would like to receive their incentive via bank transfer—in which case we ask them for their bank details—as an Amazon gift voucher, or whether they would like to donate the money. In 2012/2013, 57% of the panel members preferred to be paid in cash via a bank transfer, 31% opted for the Amazon gift voucher, and 12% chose to donate the money to charitable organizations.

Retention

Figure 2 displays the proportion of respondents participating in waves 1 through 5, conditional on being registered sample members. While 96% of the sample members participated in wave 1, this drops to 73–80% in waves 2–5. Note that these rates are based on the full registered sample. Since panel members can omit one or more waves and then return to the panel again, a wave-on-wave retention rate, as reported by some other panel studies, is less informative for the GIP (see Cheshire et al. 2011 for different types of retention rates).

Figure 2 further shows that differences between previously offline respondents and all registered participants are rather small, except for a slight drop in participation of the former in wave 2, which is recovered again from wave 3 onward. In addition, the retention rates show no evidence of a differential effect of the incentives used during the face-to-face recruitment; the €5 unconditional and the €10 conditional incentives perform similarly well. This means that the higher response rate achieved by the €5 unconditional incentive in the face-to-face recruitment is not counteracted by lower participation rates during the first five waves of the online panel.

Conclusion

The GIP combines the advantages of high-quality offline surveys with the benefits of an online format in many ways. The recruitment stage and first waves of data collection have demonstrated that an online panel based on a probability
sample and face-to-face recruitment can be successfully established and maintained in Germany. The response and retention rates achieved so far are comparable or higher than those of existing data collections in face-to-face mode in Germany. Furthermore, analyses into the representativeness of the online sample showed no major coverage or nonresponse biases. This was achieved in part by including formerly offline households, thus improving the representation of the older and female segments of the population and in part by employing traditional sampling and face-to-face fieldwork methods during the recruitment. Finally, the GIP methodology is more cost-effective than typical face-to-face surveys. This is exemplified by the fact that the 15-minute face-to-face GIP recruitment interview is about as expensive as conducting 12 online GIP waves of 20–25 minutes each, including the costs of incentives.

This article aims to guide researchers in using the GIP data as well as practitioners seeking to set up similar online panels. We describe and analyze some key aspects of the panel. First, the GIP is based on a probability sample of the general population. By conducting the initial recruitment interviews face-to-face, we were able to base the online panel on a traditional area probability sample. Second, offline households which were equipped with a BenPC and/or broadband Internet turned out to be important for the data quality since

Figure 2. Retention rates for the first five waves of the German Internet Panel; percentage of the registered sample at each wave.
this improves the panel’s representativeness. Third, our incentives experiments yield strong positive effects of unconditional cash incentives during recruitment and response did not differentially diminish during the online waves. Finally, various panel maintenance measures have been implemented to increase the stability and longevity of the GIP. They range from incentives at each waves and reminder e-mails/phone calls to personalizing the contact with our panelists by feeding back research results and introducing them to our research team. Although we cannot prove this empirically, the high retention rates in GIP might point toward the joint effectiveness of these measures.

Overall, our experience shows that investing in the recruitment and maintenance of an online panel like the GIP is worthwhile in terms of representativeness of the panel and its longevity.

Appendix A
Advance Letter in Two Versions

conditional incentive

unconditional incentive
Appendix B

Additional Information Accompanying the Advance Letter
Appendix C

Data Protection Leaflet

Appendix D

Leaflet about the Study Used during the Face-to-face Interview
Appendix E
Invitation Letter to the Online Panel with Log-in Details

Appendix F
Leaflet about the BenPC Used during the Face-to-face Interview
Acknowledgments
The GIP is the central data collection of the Collaborative Research Center (SFB) 884 “Political Economy of Reforms” at the University of Mannheim, which studies the determinants and perceptions of political reforms and their consequences for the economy. We especially thank Mick Couper for comments on an earlier version of the article and Dayana Bossert for her contributions to the analyses into the representativeness of previously offline persons. We would also like to thank Annette Holthausen, Carsten Riepe, Frederik Funke, Franziska Gebhard, Rebekka Popp, Guido Ropers, and Jan Veddeler for their work on the GIP as well as TNS Infratest Sozialforschung and LINK Institut for implementing various stages of data collection.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors gratefully acknowledge support from the SFB 884, which is funded by the German Research Foundation (DFB).

Notes
1. All response rates presented in this article were calculated using priority-coded final dispositions (Blom 2013).
2. Note that registration for the panel was possible until April 16, 2013. Late registrants first received the core questionnaire (wave 1) before they could continue with the respective wave. Persons who registered after November 2012 were thus never invited to wave 2, persons who registered after January 2013 were never invited to waves 2 and 3, and persons who registered after March 2013 were never invited to waves 2, 3, and 4.

References


