

Mode System Effects in an Online Panel Study: Comparing a Probability-based Online Panel with two Face-to-Face Reference Surveys

*Bella Struminskaya*¹, *Edith de Leeuw*² & *Lars Kaczmirek*¹

1 GESIS – Leibniz Institute for the Social Sciences

2 Department of Methodology and Statistics, Utrecht University

Abstract

One of the methods for evaluating online panels in terms of data quality is comparing the estimates that the panels provide with benchmark sources. For probability-based online panels, high-quality surveys or government statistics can be used as references. If differences among the benchmark and the online panel estimates are found, these can have several causes. First, the question wordings can differ between the sources, which can lead to differences in measurement. Second, the reference and the online panel may not be comparable in terms of sample composition. Finally, since the reference estimates are usually collected face-to-face or by telephone, mode effects might be expected. In this article, we investigate mode system effects, an alternative to mode effects that does not focus solely on measurement differences between the modes, but also incorporates survey design features into the comparison. The data from a probability-based offline-recruited online panel is compared to the data from two face-to-face surveys with almost identical recruitment protocols. In the analysis, the distinction is made between factual and attitudinal questions. We report both effect sizes of the differences and significances. The results show that the online panel differs from face-to-face surveys in both attitudinal and factual measures. However, the reference surveys only differ in attitudinal measures and show no significant differences for factual questions. We attribute this to the instability of attitudes and thus show the importance of triangulation and using two surveys of the same mode for comparison.

Keywords: mode system effect, online panel, benchmark survey comparison, data quality



© The Author(s) 2015. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 License. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1 Introduction

Several large-scale panel and repeated cross-sectional surveys incorporate or are planning to incorporate an online mode for data collection to reduce costs or to maximize contact and response rates. Some panels are designed as online panels from the beginning, with an interviewer-administered panel recruitment procedure and web-based data collection for the surveys in the panel (e.g., the LISS Panel¹ in the Netherlands and the KnowledgePanel² of GfK Custom Research, formerly Knowledge Networks in the United States). Other panels switch from interviewer-administered to online mode for their data collection (e.g., the Netherlands Kinship Panel Study³), employ an additional online component (e.g., ANES 2008-2009 Panel Study⁴), or experiment with including an online mode along with interview modes (ESS experiments on mixing modes⁵, Understanding Society Innovation Panel in the UK⁶, and Labor Force Survey and Crime Victimization Survey in the Netherlands⁷).

Data users need to know that irrespective of the mode in which data were collected, it is possible to make valid inferences about the processes that data users study, and that panel or trend data are comparable, that is, not influenced by the mode change.

In the literature on survey data collection modes, key dimensions are discerned on which modes vary that account for differences in responses across modes. In her theoretical model on mode effects, de Leeuw (1992, 2005) identified three sets of factors that explain differences between the modes: (1) media-related factors, (2) factors that are related to information transmission, and (3) interviewer effects. Media-related factors encompass social conventions and customs associated with the media utilized in survey methods. Media-related factors include socio-cultural

1 <http://www.lissdata.nl/lissdata/>

2 <http://www.knowledgenetworks.com/knpanel/>

3 <http://www.nkps.nl/NKPSEN/nkps.htm>; see also Hox, de Leeuw, & Zijlman (2015).

4 http://www.electionstudies.org/study/pages/2008_2009panel/anes2008_2009panel.htm

5 http://www.europeansocialsurvey.org/methodology/mixed_mode_data_collection.html

6 <https://www.understandingsociety.ac.uk/about/innovation-panel>; see also Auspurg et al. (2013)

7 See Schouten et al. (2013)

Direct correspondence to

Bella Struminskaya, GESIS – Leibniz Institute for the Social Sciences, B 2,1,
68159 Mannheim, Germany

E-mail: bella.struminskaya@gesis.org

effects such as familiarity with a medium, patterns of use of the medium, as well as norms of social interaction (e.g., interviewers having more control over the interviewing process and pace because they initiate the interaction). Factors related to information transmission involve more technical aspects of the communication process and include the manner in which information is presented (visual presentation, auditory presentation or both visual and auditory) as well as additional cues that are pertinent to the question-answering process (such as text and lay-out in the self-administered mode, gestures and tone of the interviewer in the face-to-face mode). The third set of factors includes the influence of an interviewer on the responses provided by respondents. Interviewer-administration can influence the respondents' feelings of privacy and lessen their willingness to disclose sensitive information. On the other hand, interviewers can provide clarification or motivate respondents to provide answers. In a meta-analysis of 52 early mode comparison studies, de Leeuw (1992, chapter 3) found that face-to-face and telephone interviews did not differ in response validity through record checks and on social desirability. When both interview modes were compared with self-administered mail surveys, the meta-analysis revealed an interesting picture. It is somewhat harder to have people answer questions in a self-administered mode (both response rate and item missing rates are higher), but the resulting answers show less social desirability and more openness on sensitive topics. These results emphasize the importance of the role of the interviewer.

Groves, Fowler, Couper, Lepkowski, Singer, and Tourangeau (2009) identify five dimensions along which data collection methods differ that partly overlap with the dimensions discussed above: (1) the degree of interviewer involvement, (2) the level of interaction with the respondent, (3) the degree of privacy for the respondent, (4) the channels of communication used, and (5) the degree of technology use. Compared to the interviewer-administered methods of data collection, online surveys eliminate interviewer involvement and offer a high level of privacy, and impose a low cognitive burden, because the questions can be easily reread on the computer screen (Tourangeau, Conrad, & Couper, 2013). In their meta-analysis of studies employing randomized experiments to compare different modes, Tourangeau et al. (2013) conclude that compared to interviewer-administered surveys, online surveys yield more reports of sensitive information and that compared to paper-and-pencil surveys, only small advantages are offered by online surveys. This finding is consistent with the finding of Klausch, Hox, & Schouten (2013), who find measurement differences between interviewer- and self-administration that are absent when comparing face-to-face to the telephone mode and mail to the online mode. It is also in line with the results of the early meta-analysis by de Leeuw (1992), and suggests a dichotomy in modes with and modes without an interviewer.

There are different ways to study mode effects (Groves, 1989). A common form is a field experiment, in which respondents are randomly assigned to a specific

mode (see also the meta-analyses by de Leeuw, 1992 and Tourangeau et al, 2013). These studies focus on a particular source or error (e.g., measurement error), investigate a particular mechanism that produces a mode difference (e.g., social desirability), and aim at estimating pure mode effects (Biemer & Lyberg, 2003, p. 207; Couper, 2011, p. 894-897).

However, in daily practice, associated with each mode is a set of decisions intended to take advantage of the benefits of each mode (Lyberg & Kasprzyk, 1991, p. 249) and in addition to mode other factors will vary, such as number of calls versus number of reminders in interview vs. web modes. Consequently, a good alternative is to measure a mode system effect, that is, to compare whole systems of data collection developed for different modes. A data collection system is defined as an “entire data collection process designed around a specific mode” (Biemer & Lyberg, 2003, p. 208; Biemer, 1988). A small number of studies focuses on the outcomes, examining total survey systems, where the final survey estimates are compared (Couper, 2011), and these studies are especially important for practitioners who want to switch from one mode of data collection to another (e.g. from interview to online survey) or employ a mixed-mode system for data collection.

In this study, we investigate mode system effects of three data collection systems, comparing the data from an online panel to two face-to-face reference surveys. Central is the question whether the different systems produce equivalent results given all the differences in data collection. Data from surveys that are carried out in different modes may differ for three reasons (de Leeuw & Hox, 2011, p. 53). Firstly, differences may be caused by the implementation of different question formats in different modes (question effects), secondly, different modes may lead to a different sample composition (selection effects), and, thirdly, the modes themselves may lead to the different response processes (mode effect).

In order to systematically investigate mode system effects, we focus our comparison on questions with the same question wording and control step-by-step for differences in sample composition. As reference we use data from two face-to-face interview surveys, whose recruitment and administrative procedures are similar. Based on the mode comparison studies and meta-analyses cited above, we do not expect differences in the estimates based on the face-to-face interviews after adjusting for potential sample composition differences. However, we do expect differences in estimates based on data collected online compared to the estimates based on the reference interview surveys (Hypothesis 1).

We concentrate on two types of questions: factual and attitudinal questions. Couper (2011, p. 896) argues that factual questions and nonsensitive questions with only yes/no responses are not likely to be affected by the presence of an interviewer, or by visual or aural presentation. Questions about attitudes and opinions place different demands on respondents' cognitive processes than factual and behavioral questions do. Attitude is defined as a collection of feelings, beliefs, and knowledge

about an issue – considerations that have different levels of accessibility. While forming an answer, respondents process these considerations, which requires deliberation and effort (Tourangeau, Rips, & Rasinski, 2000, pp. 179-180). In a face-to-face survey, an interviewer initiates the interaction and thereby controls the pace of the interview (de Leeuw, 2005), so the time might not be sufficient for a respondent to process the considerations needed to generate responses to an attitudinal question. In a self-administered online survey, by contrast, the respondent controls the survey situation, including the pacing, which allows for taking more time if needed to answer an attitudinal question. Furthermore, according to Tourangeau, Conrad, and Couper (2013), online surveys impose a lower cognitive burden on the respondent due to the visual presentation of information, allowing respondents to consider the question and response alternatives better. In an experimental study specifically designed to evaluate mode effects in which respondents were randomly assigned to mail, web, telephone and face-to-face conditions in the Crime Victimization Survey of the general population in the Netherlands, Klausch, Hox, and Schouten (2014) indeed find measurement differences for attitudinal variables such as questions about the social quality and problems of the neighborhood, and no measurement differences for two factual variables on victimization.

We therefore expect that mode system effects between online and face-to-face interviews are more pronounced in attitudinal questions than in factual questions (Hypothesis 2).

2 Data

We use data from the GESIS Online Panel Pilot⁸, a probability-based online panel of Internet users in Germany; we use two cross-sections from the German General Social Survey (ALLBUS 2010 and ALBUSS 2012) as reference surveys. Table 1 contains an overview of design characteristics for the GOPP and the reference surveys.

GESIS Online Panel Pilot (GOPP)

The GESIS Online Panel Pilot (GOPP) is a telephone-recruited online panel of Internet users aged 18 and older who live in private households in Germany. To recruit participants, the randomized last digit method was used, which is a variation of a random digit dialing (RDD) for Germany (Gabler & Häder, 2002). A dual-frame approach was used: the samples were drawn independently from landline and mobile phone frames, aiming at a final sample with 50% eligible landline numbers and 50% eligible mobile phone numbers. In order to handle the overlap between the

8 <https://dbk.gesis.org/dbksearch/sdesc2.asp?no=5582&db=e&doi=10.4232/1.11570>

two frames, the inclusion probabilities for target persons were calculated according to the formulas of Siegfried Gabler, Sabine Häder and their colleagues under the assumption of independence of the two samples (Gabler, Häder, Lehnhoff, & Mardian, 2012). The inclusion probabilities account for the sample sizes and frame sizes of both landline and mobile phone components, as well as for the number of landline and mobile phone numbers at which a respondent can be reached. For the landline component, household size was also included in the calculation. The target person, that is, the person with the most recent birthday, provided this information. For the mobile phone component, no selection procedures were implemented, since mobile phone sharing in Germany is approximately 2% (Gabler et al. 2012).

The recruitment took place in three sequential study parts using almost identical recruitment protocols. Recruitment periods were in February 2011, in June-July 2011, and in July-August 2011. After a short telephone interview, respondents were asked to provide their email addresses in order to join an online panel. Respondents who agreed would then be sent email invitations to online surveys of 10-15 minutes every month for eight months in total. Prospective panel members were offered incentives of 0, 2, 5, or 10 Euros, which were varied experimentally. An additional bonus of 20 Euros for completing all eight online questionnaires was offered.

In the GESIS Online Panel Pilot questionnaires, a number of questions were replicated from other well-known social surveys. The first goal of such replication was to assess the feasibility of surveying rather complex constructs in an online setting. The second goal was to study data quality by comparing online estimates to external benchmarks. A substantial part of questions originated from the German General Social Survey (“ALLBUS”) 2010. The original question wordings were retained, except for the cases when an adjustment was needed to make questions suitable for a self-administered mode. It is furthermore important to note that every questionnaire in the GOPP had a leading topic and most questions were asked once. Questions for analysis were selected from questionnaires (waves) 1, 2, 3, 4, 5, and 7.

German General Social Survey “ALLBUS”

The German General Social Survey is a general population survey on attitudes, behavior, and social structure in Germany. It has been conducted by GESIS biannually since 1980. The survey mode is face-to-face interviewing. For our analyses, we use ALLBUS 2010⁹ and ALLBUS 2012¹⁰, which both implemented a two-stage disproportionate random sample of individuals living in private households in Germany, aged 18 and older. The data collection for both ALLBUS 2010 and ALLBUS 2012 was conducted by the same fieldwork agency, using similar procedures for contacting and interviewing the respondents. A difference is that in

9 <http://www.gesis.org/en/allbus/study-profiles/2010/>

10 <http://www.gesis.org/en/allbus/study-profiles/2012/>

ALLBUS 2010, an incentive experiment was employed in which respondents could receive 10 Euros, 20 Euros, or no incentive (Wasmer, Scholz, Blohm, Walter, & Jutz, 2012, p. 51), whereas in ALLBUS 2012, all respondents were paid 10 Euros. It should be noted the target population of ALLBUS consists of both Internet and non-Internet users. A question on private Internet use is asked in both ALLBUS surveys. For our analysis, we are therefore able to compare GOPP with full samples of the ALLBUS surveys and with the subsamples of Internet users.

ALLBUS contains item batteries and single questions on opinions, which are repeated over the years in order to analyze social trends. The wording of the demographic questions generally does not change between ALLBUS surveys. Overall, the overlap in questions between ALLBUS 2010 and ALLBUS 2012 is 82 questions, 46 of which are not preceded by one or multiple filter questions¹¹, that is, 46 questions that were asked both in ALLBUS 2010 and in ALLBUS 2012 were posed to the total sample.

3 Measures

For the mode system comparison, we made a careful selection of factual and attitudinal questions from ALLBUS 2010 to avoid question format effects. About 73 questions from ALLBUS 2010 were also asked in the GESIS Online Panel Project. However, only a subset of these questions was asked in ALLBUS 2012. Only questions that were present in both ALLBUS 2010 and 2012 and were replicated in the GOPP are analyzed. These questions were not repeatedly measured in GOPP but distributed over the questionnaires, so that each question that we use for comparison was asked only once in the online panel. This prevents possible confounding of the answers due to learning effects (panel conditioning). We carefully inspected the questions and only included questions that had the same question wording (see Appendix for details) and were asked of the whole sample, that is, were not preceded by filter questions.

In total, 12 attitudinal and 7 factual questions fit these criteria. Attitudinal questions included respondents' assessment of the current economic situation in Germany and the economic situation in one year, the assessment of respondents' own financial situation and prospective financial situation in one year, general health, religiosity, self-assessed social class, four general attitude questions on societal functioning (anomie), and political orientation (right-left).

11 <http://www.gesis.org/fileadmin/upload/dienstleistung/daten/umfragedaten/allbus/dokumente/VariablenlisteBis2012.pdf>. This number does not include the variables with information provided by the interviewer as well as administrative variables.

Table 1 Sample and survey design description for GOPP and ALLBUS surveys

Survey characteristic	GOPP	ALLBUS 2010	ALLBUS 2012
Sample source	Adjusted RDD (landline+mobile)	Two-stage probability: municipalities and individuals from municipal registers	Two-stage probability: municipalities and individuals from municipal registers
Target population	Individuals living in Germany aged 18 and older, who use the Internet	Individuals living in households in Germany aged 18 and older	Individuals living in households in Germany aged 18 and older
Recruitment mode	CATI	CAPI	CAPI
Data collection mode	Online	CAPI	CAPI
Response/Participation rates	CATI AAPOR RR3: 17.1%; Panel participation rates per questionnaire: Q1: 54.92%; Q2: 46.89%; Q3 45.96%; Q4: 44.43%; Q5: 44.08%; Q7: 42.09%	34.4%	37.6%
Fieldwork	February 2011–May 2012	May–November 2010	April–September 2012
Incentives	0, 2, 5, 10, bonus 20 EUR (randomized experiment)	0, 10, 20 EUR (experimental)	10 EUR
N	1114 (started Q1-Q5,Q7)	2827	3480
N Internet users	1114	1869	2525
Percent Internet users ^a	71.5%	67.1%	73.6%

^a Unweighted proportion for GOPP, design-weighted for ALLBUS 2010 and 2012. AAPOR RR3 is short for AAPOR Response Rate 3 (American Association for Public Opinion Research, 2011).

Although self-assessed health is not an attitudinal variable in the classical sense, it is a complex question, and the cognitive process required to answer a self-assessment question about health appears to be more similar to the cognitive processes of the attitudinal variables in the analysis, than to the factual questions. The important difference for the cognitive process is that the factual questions in our analysis only require simple processing with no extensive recall. The factual questions concern employment status, marital status, frequency of church attendance, religious confession, being born in Germany, citizenship, and type of dwelling. We recoded several variables to dichotomous variables. The variable “working for pay” generated from “employment status” makes a distinction between those who are in paid work (working full time, part time, or irregularly) and those who are not (not working). “Legal marital status” contrasts legally married persons (married and living together with their spouse, married and living apart) with persons who are not married (divorced, never married or widowed). Religious confession was recoded into an indicator variable, which takes a value of 1 if a confession was named vs. the value of 0 if no confession was named. The percentage of refusals on this variable is negligible. Citizenship of a specific country was recoded as either having German citizenship or not. A new variable, “owner of dwelling,” was generated from the variable “type of dwelling” (for details on recoding the variables, see Table A5 in the Appendix).

4 Method

We study mode system effects by comparing the estimates from the GESIS Online Panel Pilot (GOPP) and two face-to-face reference surveys. To fully understand the processes when comparing the systems of data collection, we use a stepwise analysis procedure. First, we start with a direct comparison of the data from the GOPP and the two ALLBUS reference surveys. In this analysis, we compare the full samples of the two reference surveys ALLBUS 2010 and ALLBUS 2012, that is, Internet users plus those who do not use the Internet, with each other and with the full sample of the online panel that consists of Internet users only. This allows us to assess the differences between the online panel and the benchmark surveys that arise due to possible coverage bias, nonresponse bias, and mode effects. In this first step, we answer the practical question of what will happen if researchers switch from an interview mode to online surveys and more specifically how this will influence the unadjusted estimates.

In this first step, we do not control for sample composition and the possibility of coverage bias is substantive. Therefore, we expect to find differences between the face-to-face reference surveys and the online panel for both factual and attitudinal variables. Several studies on respondents in online surveys mention differences

in sample composition between those who use the Internet and those who do not (Bandilla, Kaczmirek, Blohm, & Neubarth, 2009; Bosnjak et al. 2013; Mohorko, de Leeuw, & Hox, 2013). Bandilla et al. (2009) found differences in age and education between the groups of Internet users and non-users. Bosnjak et al. (2013) found significant differences related to age, education, and sex between Internet users and non-users when respondents were recruited into an online panel. Mohorko, de Leeuw, and Hox (2013) give an overview of Internet coverage and coverage bias in Europe and point out that even in countries with high Internet coverage the digital divide can be observed, as Internet access is unevenly distributed across the population. In all countries, there are significant differences for age, sex, and education.

Our expectation of finding differences between the mode systems for attitudinal and factual variables is based on the relation of these variables to differences in the covered population, that is, Internet users, and not-covered population (non-Internet-users) in age, sex, and education for the online panel. For instance, due to differences in age, we might find differences in such variables as religiosity, confession, and frequency of church attendance since older people are more likely to be church members and attend religious services (Lois, 2011). In addition, we might find differences in employment status due to its relation with education.

In the second step, we match the reference population, used for the benchmark comparisons, to the target population of the online panel (Internet users). We compare the GESIS Online Panel Pilot to the subsamples of Internet users from the ALLBUS data in order to eliminate coverage as the possible cause of the differences between the GOPP online and the two face-to-face reference surveys.

However, potential differences due to selective nonresponse are not eliminated in this second step, as significant differences have been found between Internet users who are willing to participate in online surveys and Internet users who are not willing to participate. Couper et al. (2007) find that ethnicity, education, and age predict willingness to participate. Bandilla et al. (2009) show that younger and more educated Internet users are more likely to express their willingness to participate in an online survey. For face-to-face interviews, age, sex, and education have been repeatedly found to correlate with nonresponse; for an overview, see Croves & Couper (1998). Hence, differences in demographics between the face-to-face reference surveys and the online panel are expected to persist at this analysis stage. This allows us to assess the differences between the online panel and the benchmark surveys, which arise due to possible nonresponse bias and mode effects.

In the third step, we add weights to compensate for differences in nonresponse. To ensure that differences between the surveys are not caused by sample composition, we use post-stratification weighting based on age, sex, and education.

Post-stratification weighting helps to correct nonresponse errors. Previously we already corrected for coverage bias. In this final step, we are able to assess differences between the online panel and the benchmark surveys due to possible mode

effects as we correct for sample composition. According to hypothesis 1, we do not expect to find differences between the two face-to-face surveys, but we do expect to find differences between the online and the face-to-face surveys. According to hypothesis 2, we expect that differences between the online and the face-to-face surveys are more pronounced for attitudinal questions than for factual questions

To apply post-stratification weighting, classes of the sampled cases are built based on central characteristics (in our case, sex, age, and education), for which the population values are known. The weights are then assigned to the observations in each cell so that the sample data match at least the marginal totals of the population (Gabler & Ganninger, 2010). In standard social surveys, one would use the known population benchmarks to adjust for differences between the sample and the population. However, in our case we use post-stratification to correct for the sample composition bias between the surveys in order to achieve a better assessment of mode effects. Data from the ALLBUS 2010 are used as benchmark data. We treat the distributions of age (five age groups), sex, and education (recoded in three categories, see Table A5 in the Appendix for details) of Internet users in ALLBUS 2010 as reference values (Table A1 in the Appendix). We could have used population values, but unfortunately, neither the German Census nor the German Microcensus includes a question on Internet use. Post-stratification was performed using the iterative proportional fitting (IPF) algorithm also known as raking (Dehming & Stephan, 1940). Raking is an iterative process, the goal of which is to adjust the data furnished by a sample survey to the known marginal distributions obtained from other sources.¹² The weights are obtained stepwise so that the marginal distributions of the weighted data for specified variables match the benchmark marginal distributions.

The weights were calculated in Stata using the `ipfweight` procedure (Bergmann, 2011). Since the questions were spread over multiple waves in the GOPP, we calculated the post-stratification weights separately for each wave (Tables A2 and A3 in the Appendix). This allowed us to control for attrition and other confounding factors such as experiments with incentives.

Since we use data from single waves of the GOPP without making use of the longitudinal component, no additional panel weights were calculated. The demographic variables in the GOPP, which we use for post-stratification here, were all collected during the recruitment interview. In rare cases, when the previous multiple interview appointments with a respondent failed, or if a respondent was near a break-off, interviewers could ask only about Internet usage and proceed straight to the recruitment question. In such cases, demographic questions were asked later in the online questionnaires. For those cases with missing data on demograph-

12 For examples of IPF weighting compared with other weighting methods, see Gabler (1994) and Little & Wu (1991).

ics in the recruitment interview, we replaced the missing values with information obtained online if it was available. For the rare remaining cases with missing values on demographics, the missing values were then imputed using single hot deck imputation (Schonlau, 2012). In this procedure, the algorithm first identifies all observations that have no missing values for the specified variables (donor observations). In the second step, the algorithm replaces all missing values with values from a randomly chosen donor observation that is similar to the observation that has a missing value. The replacement is performed in such a way that correlations between variables are preserved. In our case, the only variable that had no missing values at all was sex.¹³ The imputation was performed before forming the age and educational groups. For an educational group of respondents who were still at school, it could not be known what school-leaving qualification the respondents would obtain; for those who reported having an “other school-leaving degree,” it could not be known how their school-leaving degrees related to the degrees of the German educational system. Those still in school and those with other school-leaving certificates were therefore marked as “missing” and imputed with the values of the variable “education” before the educational groups were formed.

In order to obtain the final weights, post-stratification weights were multiplied with design weights. Design weights correct for differences in selection probabilities. For example, in telephone surveys, persons who live in large households have a lower probability of selection than persons living in smaller households. Persons who have a very low chance of selection but have been selected into the sample “weigh” more than persons who have a high chance of being selected. Hence, a person with a low selection probability receives a high design weight, and a person with a high probability of selection receives a low weight (Gabler & Ganninger, 2010). Design weights for the GOPP, where recruitment was performed by telephone interview, were calculated using the Gabler-Häder method (Gabler, Häder, Lehnhoff, & Mardian, 2012). According to this method, the design weights equal the inverse probabilities of selection, which take into account the number of tele-

13 Unfortunately, we could not use other variables except sex since they contained missing values due to the specified recruitment condition in which respondents would only be asked about the Internet use and basic demographics (age and education – the variables for that we would like to impute missing values). We performed the analysis with post-stratification weighting using two data files: with and without imputed missing values on age and education. The results of these analyses differ for some of the variables we compare. However, the estimates that are based on the data if missing values are not imputed are far less precise due to the reduction of the sample size since cases with missing post-stratification weights but for which the information on the substantial analysis variables is present are discarded. Therefore, we report the analysis results with post-stratification weights for the calculation of which the variables age and education were imputed (Table 4). The analysis with post-stratification weights for the calculation of which variables age and education were not imputed are reported in Table A4 in the Appendix.

phone lines at which a respondent can be reached and household composition. If the design weight was missing after supplementing it with the data collected online, it was imputed with the modal category from the subsample of respondents willing to take part in the online panel. Design weights were normalized, that is, rescaled to have a mean of 1 and a sum that equals the unweighted number of cases. The weights were calculated separately for each questionnaire. Design weights in the ALLBUS account for the oversampling of persons from East Germany and were provided in both ALLBUS surveys.¹⁴

5 Results

We start with the results of the first step of the investigation of mode systems: the direct comparison without any adjustment for differences in coverage. Table 2 presents the estimates obtained from each survey and the results of the tests for statistical significance of the differences. The online GOPP differs from the ALLBUS face-to-face reference surveys in both attitudinal and factual questions. We find statistically significant differences between GOPP and ALLBUS 2010 on 11 out of 12 attitudinal items. The only variable for which no differences are found is the respondent's self-rated financial situation. GOPP differs from ALLBUS 2012 on all attitudinal variables, with the exception of 3 items from the anomie-battery, that is, the differences are found on 9 out of 12 attitudinal variables. For factual questions, GOPP differs from the two ALLBUS surveys on all 7 items at a statistically significant level. When we compare the two ALLBUS surveys, we find differences for 10 out of 12 attitudinal variables: current and prospective state of the German economy, current self-rated financial situation, religiosity, the items of the anomie-battery, left-right orientation, and self-assessed social class. For the factual questions, no significant differences between the two ALLBUS surveys were found.

In the next step, we investigate the effect of coverage bias. Table 3 presents the results of the comparisons between the online GOPP with subsamples of Internet users from the two ALLBUS surveys, excluding the non-Internet-users in ALLBUS from the analysis. For several attitudinal variables that showed statistically significant differences between GOPP and the two ALLBUS surveys, no differences are found when GOPP is compared to ALLBUS Internet users only. The difference with the results of the first step indicates potential coverage bias when measuring attitude questions online. For Internet users only, GOPP and ALLBUS 2010 only differ on 6 instead of 11 out of 12 attitudinal items: current and prospective state of German economy, health status, religiosity, one item of the anomie-battery,

14 For ALLBUS 2010, 69.33% of all interviews took place in West Germany and 30.67% in East Germany, for ALLBUS 2012, these proportions were 67.76% and 32.24% for the West and the East respectively.

and left-right orientation. The differences from the subsample of Internet users in ALLBUS 2012 are found for 4 instead of 9 attitudinal variables: prospective German economy, current self-rated financial situation, religiosity, and one category of the variable self-assessed social class. We find statistically significant differences between GOPP and the subsample of Internet users in ALLBUS 2010 for all factual variables with the exception of the variable "German citizenship." The differences between GOPP and ALLBUS 2012 Internet users are statistically significant for all 7 factual variables. With the exception of prospective financial situation and self-assessed social class, the ALLBUS surveys differ from one another on all attitudinal variables. In sum, when comparing the subsamples of Internet users, ALLBUS surveys differ from each other in 10 out of 12 attitudinal variables and none of the factual variables.

In the last step, we use weighting adjustment to compensate for potential differences in nonresponse between the surveys. Table 4 reports all comparisons using weighted data. The set of variables in which the online GOPP differs from the ALLBUS surveys is different from the set of variables that showed significant differences reported in Table 3, where sample composition bias due to coverage but not due to nonresponse was taken into account. The online GOPP now differs from both ALLBUS surveys in 6 out of 12 attitudinal items. In the following, we refer to the comparisons of GOPP with the subsamples of Internet users in ALLBUS with post-stratification weighting applied to GOPP and to ALLBUS 2012.

The variables in which GOPP differs from ALLBUS 2010 are not the same variables in which GOPP differs from ALLBUS 2012. Significant differences between ALLBUS 2010 and GOPP are now observed for the following attitudinal variables: state of German economy (both current and prospective), self-rated current financial situation, health status, religiosity, and left-right orientation. The set of attitudinal variables with statistically significant differences between GOPP and ALLBUS 2012 excludes the current state of German economy and left-right orientation, but includes two of the items of the anomie-battery. However, the two ALLBUS surveys still show significant mutual differences for the same 10 out of 12 items reported in Table 3.

For factual questions, GOPP differs from ALLBUS 2010 in 6 out of 7 items: in all factual questions with the exception of confession. Estimates from GOPP differ from ALLBUS 2012 in 5 out of 7 items, the exceptions being confession and frequency of going to church. As was the case with the unweighted data in Table 3, there are no significant differences between the two ALLBUS surveys for any of the factual questions. These two findings taken together indicate the presence of a mode system effect in our data. The two face-to-face reference surveys do not show any mutual differences regarding factual questions, but they do differ from the online results, even after adjusting for sample composition differences.

Table 2 Comparison of the GESIS Online Panel Pilot to two ALLBUS surveys: attitudinal and factual questions

Variable	Q	GOPP			ALLBUS 2010 full sample			ALLBUS 2012 full sample		
		Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
<i>Attitudinal:</i>										
German economy	1	3.400	0.026	[3.350, 3.450]	3.037***	0.016	[3.006, 3.068]	3.332*	0.015	[3.303, 3.360]
German economy in 1 year	1	2.868	0.029	[2.811, 2.925]	3.014***	0.017	[2.980, 3.047]	2.662***	0.013	[2.637, 2.688]
Self-rated financial situation	1	3.454	0.027	[3.401, 3.507]	3.411	0.017	[3.379, 3.444]	3.559**	0.014	[3.531, 3.587]
Self-rated financial situation in 1 year	1	3.187	0.027	[3.134, 3.241]	3.090**	0.014	[3.064, 3.117]	3.094**	0.011	[3.071, 3.116]
Health status	1	3.679	0.032	[3.616, 3.743]	3.561**	0.021	[3.521, 3.602]	3.562**	0.018	[3.527, 3.597]
Religiosity	4	4.462	0.125	[4.217, 4.707]	5.724***	0.059	[5.608, 5.840]	6.111***	0.054	[6.006, 6.216]
Self-assessed social class	5									
lower class		2.68%	0.007	[1.66, 4.32]	3.16%	0.003	[2.56, 3.90]	1.94%	0.002	[1.53, 2.46]
working class		19.43%	0.017	[16.34, 22.96]	25.75%**	0.008	[24.13, 27.45]	26.21%***	0.008	[24.73, 27.75]
middle class		61.12%	0.020	[57.18, 64.92]	59.74%	0.010	[57.83, 61.61]	61.12%	0.009	[59.41, 62.80]
upper middle class		16.14%	0.014	[13.56, 19.09]	10.84%**	0.006	[9.67, 12.13]	10.08%***	0.006	[9.05, 11.22]
upper class		0.63%	0.003	[0.25, 1.56]	0.51%	0.001	[0.29, 0.89]	0.65%	0.001	[0.42, 1.00]
Life is getting worse	7	78.66%	0.017	[75.20, 81.75]	83.42%**	0.007	[81.93, 84.81]	77.44%	0.008	[75.91, 78.90]
Irresponsible to have children	7	35.10%	0.020	[31.36, 39.04]	42.07%**	0.010	[40.19, 43.98]	35.70%	0.009	[34.05, 37.39]
Politicians not interested	7	76.86%	0.018	[73.23, 80.14]	80.82%*	0.008	[79.24, 82.31]	76.18%	0.008	[74.62, 77.67]
People don't care about others	7	65.30%	0.020	[61.33, 69.07]	73.99%***	0.009	[72.26, 75.64]	70.91%**	0.008	[69.27, 72.49]
Left-right orientation	7	4.848	0.070	[4.712, 4.985]	5.218***	0.034	[5.152, 5.284]	5.045*	0.031	[4.984, 5.106]

Table 2 continued

Variable	Q	GOPP			ALLBUS 2010 full sample			ALLBUS 2012 full sample		
		Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
<i>Factual:</i>										
Working for pay	2	83.70%	0.014	[80.73, 86.29]	57.80%***	0.010	[55.91, 59.68]	59.76%***	0.009	[58.04, 61.45]
Legal marital status	3	50.47%	0.020	[46.60, 54.34]	57.55%***	0.010	[55.65, 59.42]	57.49%***	0.009	[55.77, 59.20]
Confession	4	64.56%	0.019	[60.74, 68.20]	73.14%***	0.008	[71.49, 74.73]	73.13%***	0.007	[71.65, 74.57]
Frequency of church attendance	4	2.143	0.045	[2.055, 2.231]	2.385%***	0.025	[2.335, 2.434]	2.338%***	0.023	[2.293, 2.384]
Born in Germany	4	90.72%	0.012	[88.17, 92.76]	84.01%***	0.007	[82.53, 85.39]	85.27%***	0.006	[83.96, 86.49]
German citizenship	4	96.25%	0.007	[94.51, 97.45]	94.07%*	0.005	[93.07, 94.94]	93.93%***	0.004	[93.00, 94.73]
Owner of dwelling	5	47.78%	0.020	[43.86, 51.73]	57.11%***	0.010	[55.20, 58.99]	59.35%***	0.009	[57.64, 61.04]

Note: * $p < .05$, ** $p < .01$, *** $p < .001$. Differences between ALLBUS 2010 and 2012 in italic bold, Q short for GOPP questionnaire, SE short for standard error, CI short for confidence interval. N for GOPP (started): Q1=1010; Q2=838; Q3=800; Q4=775; Q5=761; Q7=729. Design weights are applied to all three surveys. Significance tests are pairwise t-tests for continuous/ordinal variables or tests of proportions.

Table 3 Comparison of the GESIS Online Panel Pilot to the subsamples of Internet users from two ALLBUS surveys: attitudinal and factual questions

Variable	Q	GOPP			ALLBUS 2010 Internet users			ALLBUS 2012 Internet users		
		Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
<i>Attitudinal:</i>										
German economy	1	3.400	0.026	[3.350, 3.450]	3.086***	0.019	[3.048, 3.124]	3.355	0.017	[3.321, 3.389]
German economy in 1 year	1	2.868	0.029	[2.811, 2.925]	3.077***	0.021	[3.036, 3.117]	2.663***	0.015	[2.633, 2.694]
Self-rated financial situation	1	3.454	0.027	[3.401, 3.507]	3.471	0.020	[3.433, 3.510]	3.595***	0.016	[3.564, 3.627]
Self-rated financial situation in 1 year	1	3.187	0.027	[3.134, 3.241]	3.182	0.017	[3.148, 3.216]	3.145	0.014	[3.117, 3.173]
Health status	1	3.679	0.032	[3.616, 3.743]	3.823***	0.023	[3.778, 3.868]	3.725	0.020	[3.686, 3.764]
Religiosity	4	4.462	0.125	[4.217, 4.707]	5.974***	0.073	[5.832, 6.117]	6.334***	0.062	[6.213, 6.456]
Self-assessed social class	5									
lower class		2.68%	0.007	[1.65, 4.32]	2.27%	0.003	[1.67, 3.06]	1.58%	0.002	[1.16, 2.14]
working class		19.43%	0.017	[16.34, 22.96]	20.23%	0.010	[18.42, 22.17]	21.81%	0.008	[20.20, 23.52]
middle class		61.12%	0.020	[57.18, 64.92]	63.38%	0.012	[61.08, 65.63]	63.71%	0.010	[61.73, 65.64]
upper middle class		16.14%	0.014	[13.56, 19.09]	13.80%	0.008	[12.22, 15.55]	12.24%*	0.007	[10.93, 13.67]
upper class		0.63%	0.003	[0.25, 1.56]	0.32%	0.001	[0.14, 0.73]	0.66%	0.002	[0.40, 1.10]
Life is getting worse	7	78.66%	0.017	[75.20, 81.75]	82.00%	0.009	[80.10, 83.74]	75.71%	0.009	[73.88, 77.45]
Irresponsible to have children	7	35.10%	0.020	[31.36, 39.04]	37.35%	0.012	[35.10, 39.65]	31.07%	0.010	[29.22, 32.99]
Politicians not interested	7	76.86%	0.018	[73.23, 80.14]	77.42%	0.010	[75.37, 79.35]	73.63%	0.009	[71.75, 75.43]
People don't care about others	7	65.30%	0.020	[61.33, 69.07]	71.72%**	0.011	[69.55, 73.79]	68.69%	0.010	[66.74, 70.57]
Left-right orientation	7	4.848	0.070	[4.712, 4.985]	5.143***	0.041	[5.063, 5.222]	4.937	0.035	[4.868, 5.006]

Table 3 continued

Variable	Q	GOPP			ALLBUS 2010 Internet users			ALLBUS 2012 Internet users		
		Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
<i>Factual:</i>										
Working for pay	2	83.70%	0.014	[80.73, 86.29]	72.41%***	0.011	[70.26, 74.46]	72.11%***	0.009	[70.25, 73.90]
Legal marital status	3	50.47%	0.020	[46.60, 54.34]	56.89%***	0.012	[54.56, 59.19]	57.00%***	0.010	[54.98, 59.00]
Confession	4	64.56%	0.019	[60.74, 68.20]	70.57%***	0.011	[68.47, 72.59]	71.00%***	0.009	[69.20, 72.74]
Frequency of church attendance	4	2.143	0.045	[2.055, 2.231]	2.301**	0.029	[2.244, 2.359]	2.287**	0.026	[2.237, 2.338]
Born in Germany	4	90.72%	0.012	[88.17, 92.76]	87.44%*	0.008	[85.76, 88.94]	86.62%***	0.007	[85.13, 87.99]
German citizenship	4	96.25%	0.007	[94.51, 97.45]	94.53%	0.006	[93.32, 95.52]	93.64%***	0.005	[92.52, 94.59]
Owner of dwelling	5	47.78%	0.020	[43.86, 51.73]	56.64%***	0.012	[54.30, 58.95]	59.19%***	0.010	[57.18, 61.17]

Note: * $p < .05$, ** $p < .01$, *** $p < .001$. Differences between ALLBUS 2010 and 2012 in italic bold, Q short for GOPP questionnaire, SE short for standard error, CI short for confidence interval. N for GOPP (started): Q1=1010; Q2=838; Q3=800; Q4=775; Q5=761; Q7=729. Design weights are applied to all three surveys. Significance tests are pairwise t-tests for continuous/ordinal variables or tests of proportions.

Table 4 Comparison of the GESIS Online Panel Pilot to the subsamples of Internet users from two ALLBUS surveys: attitudinal and factual questions with post-stratification

Variable	Q	GOPP			ALLBUS 2010 Internet users			ALLBUS 2012 Internet users		
		Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
<i>Attitudinal:</i>										
German economy	1	3.336	0.029	[3.279, 3.393]	3.086***	0.019	[3.048, 3.124]	3.351	0.018	[3.317, 3.385]
German economy in 1 year	1	2.844	0.033	[2.779, 2.909]	3.077***	0.021	[3.036, 3.117]	2.667***	0.016	[2.636, 2.698]
Self-rated financial situation	1	3.394	0.030	[3.334, 3.453]	3.471*	0.020	[3.433, 3.510]	3.593***	0.017	[3.560, 3.626]
Self-rated financial situation in 1 year	1	3.168	0.031	[3.107, 3.228]	3.182	0.017	[3.148, 3.216]	3.157	0.015	[3.128, 3.186]
Health status	1	3.606	0.038	[3.531, 3.681]	3.823***	0.023	[3.778, 3.868]	3.748**	0.020	[3.709, 3.787]
Religiosity	4	4.585	0.145	[4.300, 4.870]	5.974***	0.073	[5.832, 6.117]	6.341***	0.063	[6.218, 6.466]
Self-assessed social class	5									
lower class		3.36%	0.010	[1.85, 6.03]	2.27%	0.003	[1.67, 3.06]	1.55%	0.003	[1.13, 2.13]
working class		24.81%	0.023	[20.52, 29.65]	20.23%	0.010	[18.42, 22.17]	22.21%	0.009	[20.55, 23.97]
middle class		59.36%	0.024	[54.60, 63.94]	63.38%	0.012	[61.08, 65.63]	63.24%	0.010	[61.21, 65.22]
upper middle class		12.02%	0.012	[9.87, 14.56]	13.80%	0.008	[12.22, 15.55]	12.36%	0.007	[11.03, 13.83]
upper class		0.45%	0.002	[0.18, 1.16]	0.32%	0.001	[0.14, 0.73]	0.64%	0.002	[0.38, 1.09]
Life is getting worse	7	79.51%	0.019	[75.47, 83.04]	82.00%	0.009	[80.10, 83.74]	75.58%	0.009	[73.70, 77.36]
Irresponsible to have children	7	40.41%	0.024	[35.87, 45.13]	37.35%	0.012	[35.10, 39.65]	31.06%***	0.010	[29.17, 33.01]
Politicians not interested	7	79.28%	0.019	[75.38, 82.70]	77.42%	0.010	[75.37, 79.35]	73.56%***	0.010	[71.64, 75.40]
People don't care about others	7	68.45%	0.022	[64.09, 72.50]	71.72%	0.011	[69.55, 73.79]	68.63%	0.010	[66.64, 70.55]
Left-right orientation	7	4.945	0.080	[4.788, 5.102]	5.143*	0.041	[5.063, 5.222]	4.946	0.036	[4.876, 5.016]

Table 4 continued

Variable	Q	GOPP			ALLBUS 2010 Internet users			ALLBUS 2012 Internet users		
		Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
<i>Factual:</i>										
Working for pay	2	83.55%	0.016	[80.28, 86.37]	72.41%***	0.011	[70.26, 74.46]	73.60%***	0.009	[71.75, 75.37]
Legal marital status	3	49.87%	0.023	[45.30, 54.44]	56.89%***	0.012	[54.56, 59.19]	56.23%*	0.011	[54.16, 58.28]
Confession	4	67.29%	0.021	[63.10, 71.21]	70.57%	0.011	[68.47, 72.59]	71.29%	0.009	[69.46, 73.04]
Frequency of church attendance	4	2.175	0.052	[2.073, 2.277]	2.301*	0.029	[2.244, 2.359]	2.286	0.026	[2.234, 2.337]
Born in Germany	4	91.04%	0.013	[88.08, 93.32]	87.44%*	0.008	[85.76, 88.94]	86.25%***	0.008	[84.69, 87.67]
German citizenship	4	96.50%	0.008	[94.63, 97.73]	94.53%*	0.006	[93.32, 95.52]	93.50%***	0.006	[92.33, 94.50]
Owner of dwelling	5	45.38%	0.023	[40.86, 49.98]	56.64%***	0.012	[54.30, 58.95]	58.06%***	0.010	[56.99, 60.10]

Note: * $p < .05$, ** $p < .01$, *** $p < .001$. Differences between ALLBUS 2010 and 2012 in italic bold, Q short for GOPP questionnaire, SE short for standard error, CI short for confidence interval. N for GOPP (started); Q1=1010; Q2=838; Q3=800; Q4=775; Q5=761; Q7=729. Weighted data: ALLBUS 2010 design, GOPP & ALLBUS 2012 design and post-stratification. Significance tests are pairwise t-tests for continuous/ordinal variables or tests of proportions.

The fact that the two face-to-face ALLBUS surveys differ on attitudinal questions shows that attitudes, being more unstable constructs, would not have allowed us to single out a mode system effect. We see that two surveys with identical recruitment and design features do not differ in factual questions, whereas the difference between the online and the face-to-face surveys is clear.

However, finding statistically significant differences is not the only indicator of mode system effects. In addition to statistical significance, Biemer (1988) recommends examining the effect size, the direction of the difference, and the violations of the underlying assumptions for the mode comparison study, which could explain the magnitude of the difference.

Calculating effect sizes standardizes the comparisons between the means and proportions reported in Table 4 and allows for a better estimation of error, because effect sizes do not only take the difference of the estimates into account, but also the sample sizes and the precision of estimates.

We calculated the standardized mean difference effect sizes that are used to synthesize results from studies that contrast two groups on measures with a continuous underlying distribution for the continuous and ordinal variables (Lipsey & Wilson, 2001, p. 172) and approximated standardized mean difference effect sizes that are calculated differencing the arcsine-transformed proportions (Lipsey & Wilson, 2001, p. 187) for binary variables.¹⁵ The standardized mean difference effect sizes allow us to compare the size and direction of the difference between the GOPP and ALLBUS 2010 with the difference between GOPP and ALLBUS 2012 as well as with the difference between ALLBUS 2010 and ALLBUS 2012 for each variable. In order to compare the differences between the surveys for groups of variables (attitudinal vs. factual), we calculated the mean effect sizes. The mean effect size is computed by weighting each effect size by the inverse of its variance (Lipsey & Wilson, 2001, p. 114). Since the standardized mean effect sizes take the direction of the difference (indicated by the sign of the effect size) into account, the differences between the variables may be underestimated when the sum of the negative

15 The standardized mean difference effect size (ES_{sm}) is the difference between the group means (\bar{X}) divided by the pooled standard deviation (s), which is calculated based on the sample sizes for each group:

$$ES_{sm} = \frac{\bar{X}_1 - \bar{X}_2}{s_{pooled}}, \quad s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

The approximations based on dichotomous data: $ES_{sm} = \arcsine(p_1) - \arcsine(p_2)$, where p is the proportion for each group (Lipsey & Wilson, 2001, p. 198-200).

The mean effect size $\overline{ES} = \frac{\sum_{i=1}^k (w_i \times ES_i)}{\sum_{i=1}^k w_i}$; $w_i = \frac{1}{SE_i^2}$,

where w is the inverse variance weight, ES is the unweighted effect size (d), and SE is the standard error of the difference between the effect sizes (Lipsey & Wilson, 2001, p. 114).

and positive effect sizes is calculated in order to compute the mean effect size. We therefore also calculate the absolute mean effect sizes, which are based on absolute values of the standardized mean difference effect sizes. In Table 5, we report the standardized mean difference effect sizes (d) and mean effect sizes as well as absolute mean effect sizes for attitudinal variables, factual variables and the overall mean effect size, which incorporates both attitudinal and factual variables.

From Table 5, we can conclude that the directions of all mean effect sizes are the same. For weighted mean effect sizes over all questions, the difference between GOPP and ALLBUS 2010 (-0.063) and the difference between GOPP and ALLBUS 2012 (-0.036) are larger than the difference between the two ALLBUS surveys (-0.030). However, the difference between GOPP and ALLBUS 2012 has almost the same magnitude as the difference between the two face-to-face reference surveys.

If we look at the effect sizes by type of questions (factual vs. attitudinal), the results mimic what we have seen when comparing the frequency counts of statistically significant results. For factual questions, the difference between the two face-to-face surveys (ALLBUS 2010 & 2012) is minimal (0.001) and for the online GOPP vs. face-to-face surveys the differences are almost the same (0.048 and 0.050). For attitudinal questions ALLBUS 2012 is closer to GOPP (-0.031) than to ALLBUS 2010 (-0.068). The difference between the two face-to-face surveys on attitudinal questions is even larger than the difference between the online GOPP and the ALLBUS 2012 (-0.042 vs. -0.031). Although effect sizes differ between surveys and between question types, most are very small (Ferguson, 2009, p. 533). The absolute weighted mean effect sizes reported in the end of Table 5 do not account for the direction of the effect, but only for the magnitude. Here, we see, once again, that for attitudinal items GOPP estimates are closer to ALLBUS 2012 estimates, but in the same magnitude as the difference between the face-to-face surveys. However, for factual questions the difference between the two face-to-face surveys is much smaller than between each of the two ALLBUS surveys and GOPP, resulting in an overall difference between GOPP and each individual ALLBUS survey that is larger than that between the two ALLBUS surveys. Although absolute mean effect sizes are larger than mean effect sizes, they are still small and do not exceed 0.2. From a practical point of view, the mode system effects between the surveys are therefore negligible when judged by their effect sizes and small when judged by the absolute effect sizes. Naturally, this is after differences in sample composition have been accounted for.

Table 5 Effect sizes (d), inverse variance weights (w) and mean effect sizes across the surveys

Variable	GOPP vs. ALLBUS 2010 Internet users		GOPP vs. ALLBUS 2012 Internet users		ALLBUS 2012 vs. 2010 Internet users	
	d	w	d	w	d	w
<i>Attitudinal:</i>						
German economy	0.295	625.000	-0.017	688.705	0.307	1051.525
German economy in 1 year	-0.247	623.053	0.206	682.128	-0.488	1019.368
Self-rated financial situation	-0.087	630.120	-0.229	684.932	0.143	1062.699
Self-rated financial situ- ation in 1 year	-0.017	626.566	0.014	684.463	-0.034	1052.632
Health status	-0.206	621.504	-0.136	679.348	-0.075	1064.963
Religiosity	-0.407	530.786	-0.523	568.182	0.117	1059.322
Self-assessed social class						
lower class	0.222	49.998	0.437	48.603	-0.214	64.045
working class	0.145	311.915	0.080	346.380	0.065	571.755
middle class	-0.094	416.840	-0.090	452.284	-0.003	800.000
upper middle class	-0.087	191.022	-0.018	201.086	-0.070	393.236
upper class	0.189	7.004	-0.195	9.100	0.384	13.942
Life is getting worse	-0.088	269.469	0.125	305.344	-0.213	552.181
Irresponsible to have children	0.071	406.669	0.225	428.266	-0.154	776.398
Politicians not interested	0.060	284.738	0.175	310.945	-0.115	626.174
People don't care about others	-0.086	362.450	-0.005	394.477	-0.082	719.424
Left-right orientation	-0.107	509.684	-0.001	549.451	-0.113	1022.495
<i>Factual:</i>						
Working for pay	0.364	280.741	0.331	296.121	0.033	692.042
Legal marital status	-0.156	453.309	-0.141	490.196	-0.015	861.326
Confession	-0.085	383.877	-0.104	413.736	0.019	720.461
Frequency of church attendance	-0.097	539.374	-0.084	582.072	-0.012	1062.699
Born in Germany	0.209	157.109	0.266	169.033	-0.058	397.931
German citizenship	0.258	66.934	0.359	72.417	-0.101	193.836
Owner of dwelling	-0.250	435.730	-0.282	468.604	0.032	856.164

Table 5 continued

Variable	GOPP vs. ALLBUS 2010 Internet users		GOPP vs. ALLBUS 2012 Internet users		ALLBUS 2012 vs. 2010 Internet users	
	<i>d</i>	<i>w</i>	<i>d</i>	<i>w</i>	<i>d</i>	<i>w</i>
Mean effect size:						
attitudinal questions		-0.068 (0.012)		-0.031 (0.012)		-0.042 (0.009)
factual questions		-0.048 (0.021)		-0.050 (0.020)		-0.001 (0.014)
Overall		-0.063 (0.011)		-0.036 (0.010)		-0.030 (0.008)
Mean absolute effect size:						
attitudinal questions		0.157 (0.012)		0.141 (0.012)		0.150 (0.009)
factual questions		0.180 (0.021)		0.185 (0.020)		0.028 (0.014)
Overall		0.163 (0.011)		0.153 (0.010)		0.115 (0.008)

Note: *d* is the unweighted effect size, *w* is the inverse variance weight, standard errors in parentheses.

6 Discussion

We investigated mode system effects by comparing the data from a probability-based telephone-recruited online panel with the data from two face-to-face surveys. All three sources implemented questions with identical wording and we controlled for differences in sample composition due to undercoverage and selective nonresponse in order to single out the mode system effect. We distinguished between factual and attitudinal questions. We hypothesized that the effects are more pronounced for attitudinal questions than for factual questions. This hypothesis finds no support. There are differences between the online collected data and interviewer collected data on both types of questions. However, the distinction between factual and attitudinal data remains important. We did find that for factual questions both face-to-face surveys differ from the online panel, but do not differ significantly from each other. This conclusion is supported by the analysis of effect sizes. For attitudinal questions, the difference between the two interviewer-administered face-to-face surveys is larger than the difference between one face-to-face survey (ALLBUS 2012) and the online panel. We attribute this result to the instability of attitudes. However, alternative explanations are possible. For example, the reference period

of the ALLBUS 2012 is closer to the online panel, which was administered in 2011-2012 than to the ALLBUS 2010. Surprisingly, the mean effect size for differences in attitudinal variables is larger for ALLBUS 2010 to ALLBUS 2012 comparison than for the GOPP to ALLBUS 2012 comparison, although all effect sizes are small. When we look at the content of the questions, we see that effect sizes are relatively large for questions that refer to the economic situation. The differences between the surveys on those variables might be attributed not to the mode system effect, but to the fact that Germany experienced real economic changes between 2010 and 2012. To examine the robustness of our results, we recalculated what the effect sizes are when the four variables that refer to the economic condition are excluded. This did not affect our overall conclusions, but it provided us with more accurate measures of the mode system effect.¹⁶

When comparing data from surveys to data from benchmark studies, Callegaro et al. (2014) advise that the following conditions be met: (1) question wording should be identical across compared surveys and (2) populations represented by each survey need to be comparable. Typically, studies comparing online panel data to benchmarks use demographic and behavioral measures (Callegaro et al., 2014). One of the reasons for this might be that the official statistics ideally used as benchmarks can only provide such data. If benchmarks come from high-quality surveys, using attitudinal measures can be considered. Our study meets both conditions (identical questions wording and comparable populations after adjustment) and shows the importance of the reference period depending on the nature of the measure. Ideally, the reference period should be the same for a benchmark source (survey) and the survey that is being compared to the benchmark – a situation given when a pure mode effect is estimated in a randomized experiment. However, in practical circumstances – when estimating a mode system effect – this might not be the case.

Our findings have important implications for estimating mode system effects. First, we have shown the importance of having more than one reference survey. In our case, both face-to-face surveys serve as control surveys for the online panel and allow us to draw conclusions about whether or not differences could be attributed to the mode. We use two surveys with equal recruitment procedures. Second, we show the importance of distinguishing between types of questions. In past studies, differences between online and interviewer-administered modes are well-documented for sensitive questions (e.g., meta-analysis by Tourangeau, Conrad, & Couper, 2013,

16 Mean effect sizes without the economic items. Overall: -0.082 (0.013), -0.048 (0.012), -0.036 (0.009); absolute: 0.164 (0.013), 0.168 (0.012), 0.073 (0.009); for attitudinal items: -0.102 (0.016), -0.047 (0.015), -0.058 (0.011); absolute for attitudinal items: 0.155 (0.016), 0.157 (0.015), 0.101 (0.011).

p. 142).¹⁷ We draw on another dimension, in which a distinction is made based on the cognitive demands that questions place on a respondent. Our findings are in line with Klausch, Hox, and Schouten (2014), who also report that mode effects depend on the question type. For data analysts, the conclusion to be drawn from our analysis is that mode system effects differ across question types. Data users should therefore have different concerns about mode effects when analyzing only attitudinal, only factual, or both types of items. Third, we find a common denominator for the comparison of mode system effects by reporting effect sizes. The magnitude of the mode system effect when judged by the effect sizes is small. However, researchers who use data from large surveys might be misled if they rely solely on significance testing. We encourage other researchers to report effect sizes. This would allow for comparing our results to similar studies based on different data. Fourth, it is important to realize, that we corrected for sample composition bias, and found indications of coverage and nonresponse bias. In survey practice, one could use mixed-mode approaches to account for the digital divide; an example is the GESIS Panel¹⁸, where a mix of postal mail and Internet is used for data collection.

Finally, we investigated the effect of mode systems on point estimates answering the practical question of survey practitioners, of what happens when we change our data collection procedures. To tease out the reasons for mode system effects, a series of carefully designed experiments is needed. These experiments should take both selection and measurement effects into account and not only use point estimates, but also rely on other indicators, such as indices for response tendencies (cf. Tourangeau, 2013).

17 We did not have sensitive items to use for our analysis. However, for religiosity, we find that the percentage online panel respondents who choose the answer “not religious” is much higher than for interviewer-administered surveys (29.44% with SE=0.021, CI=[25.56, 33.64] for GOPP and 7.72% with SE=0.007, CI=[6.54, 9.10] for ALLBUS 2010, 6.27% with SE=0.005, CI=[5.32, 7.38], $p < .001$ for GOPP vs. either ALLBUS). This could be indicative of social desirability.

18 www.gesis-panel.org

References

- American Association for Public Opinion Research. (2011). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 7th edition: AAPOR.
- Auspurg, K., Burton, J., Cullinane, C., Delavande, A., Fumagelli, L., Iacovou, M., Zafar, B. (2013). Understanding Society Innovation Panel Wave 5: Results from Methodological Experiments. In J. Burton (Ed.), *Understanding Society Working Paper Series*. Essex: Institute for Social and Economic Research, University of Essex.
- Bandilla, W., Kaczmirek, L., Blohm, M., & Neubarth, W. (2009). Coverage und Nonresponse-Effekte bei Online-Bevölkerungsumfragen [Coverage and nonresponse effects in online population-surveys]. In N. Jakob, H. Schoen & T. Zerback (Eds.), *Sozialforschung im Internet: Methodologie und Praxis der Online-Befragung* (pp. 129-143). Wiesbaden: VS Verlag.
- Bergmann, M. (2011). Stata-ado "ipfweight": IPF-Algorithm to create adjustment survey weights. from <http://fmwww.bc.edu/repec/bocode/i/ipfweight.ado>
- Biemer, P. P. (1988). Measuring Data Quality. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls & J. Waksberg (Eds.), *Telephone Survey Methodology* (pp. 273-283). New York: Wiley.
- Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality*. Hoboken, New Jersey: Wiley.
- Bosnjak, M., Haas, I., Galesic, M., Kaczmirek, L., Bandilla, W., & Couper, M. P. (2013). Sample composition discrepancies in different stages of a probability-based online panel. *Field Methods*. doi: 10.1177/1525822X12472951
- Callegaro, M., Villar, A., Yeager, D., & Krosnick, J. A. (2014). A critical review of studies investigating the quality of data obtained with online panels based on probability and nonprobability samples. In M. Callegaro, R. Baker, J. D. Bethlehem, A. S. Göritz, J. A. Krosnick & P. J. Lavrakas (Eds.), *Online panel research: A data quality perspective* (pp. 23-53). New York: Wiley.
- Couper, M. P. (2011). The future of modes of data collection. *Public Opinion Quarterly*, 75, 5, 889-908.
- Couper, M. P., Kapteyn, A., Schonlau, M., & Winter, J. (2007). Noncoverage and Nonresponse in an Internet Survey. *Social Science Research*, 36, 131-148.
- Dehming, W. E., & Stephan, F. F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, 11(4), 427-444.
- de Leeuw, E. D. (1992). *Data quality in mail, telephone, and face-to-face surveys*. Amsterdam: TT.Publikaties. <http://edithl.home.xs4all.nl/pubs/disseddl.pdf>
- de Leeuw, E. D. (2005). To Mix or Not to Mix Data Collection Modes in Surveys. *Journal of Official Statistics*, 21(2), 233-255.
- de Leeuw, E. D., & Hox, J. J. (2011). Internet surveys as a part of a mixed-mode design. In M. Das, P. Ester & L. Kaczmirek (Eds.), *Social and behavioral research and the Internet* (Vol. 45-76). New York: Routledge.
- Ferguson, C. J. (2009). An Effect Size Primer: A Guide for Clinicians and Researchers. *Professional Psychology: Research and Practice*, 40(5), 532-538.
- Gabler, S. (1994). Eine allgemeine Formel zur Anpassung an Randtabellen [A general formula for adjustment for contingency tables with known margins]. In S. Gabler, J. H. P.

- Hoffmeyer-Zlotnik & D. Krebs (Eds.), *Gewichtung in der Umfragepraxis [Weighting in the Survey Practice]* (pp. 88-105). Westdeutscher Verlag
- Gabler, S., & Ganninger, M. (2010). Gewichtung [Weighting]. In C. Wolf & H. Best (Eds.), *Handbuch der sozialwissenschaftlichen Datenanalyse [Handbook of the Data Analysis in the Social Sciences]* (pp. 143-164). Wiesbaden: VS Verlag.
- Gabler, S., & Häder, S. (2002). Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland [Telephone samples: Methodological innovations and applications in Germany]. Münster: Waxmann Verlag.
- Gabler, S., Häder, S., Lehnhoff, I. & Mardian, E. (2012). Weighting for unequal inclusion probabilities and nonresponse in dual frame telephone surveys. In S. Häder, M. Häder & M. Kühne (Eds.), *Telephone surveys in Europe: Research and practice* (pp. 147-167). Heidelberg: Springer.
- Groves, R. M. (1989). *Survey errors and survey costs*. New York: Wiley.
- Groves, R. M. & Couper, M. P. (1998). *Nonresponse in household interview surveys*. New York: Wiley.
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). New York: John Wiley.
- Hox, J. J., de Leeuw, E. D., & Zijlman, E. A. O. (2015). Measurement equivalence in mixed-mode surveys. *Frontiers in Psychology*, 6(87). doi: 10.3389/fpsyg.2015.00087
- Klausch, T., Hox, J. J., & Schouten, B. (2013). Measurement effects of survey mode on the equivalence of attitudinal rating scale questions. *Sociological Methods & Research* 42(3), 227-263.
- Klausch, T., Hox, J. J., & Schouten, B. (2014). *Evaluating mixed-mode redesign strategies against benchmark surveys: The case of the Crime Victimization Survey*. Paper presented at the GOR General Online Conference, 05-07.03.2014 in Cologne.
- Lee, S. (2006). An Evaluation of Nonresponse and Coverage Errors in a Prerecruited Probability Web Panel Survey. *Social Science Computer Review*, 24, 460-475.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, California: Sage.
- Little, R. J. A., & Wu, M. (1991). Models for contingency tables with known margins when target and sampled populations differ. *Journal of the American Statistical Association*, 86(413), 87-95.
- Lois, D. (2011). Kirchenmitgliedschaft und Kirchengangshäufigkeit im Zeitverlauf – Eine Trendanalyse unter Berücksichtigung von Ost-West-Unterschieden [Church membership and frequency of church attendance over time – A trend analysis taking into consideration differences in East and West Germany]. *Comparative Population Studies – Zeitschrift für Bevölkerungswissenschaft*, 36(1), 127-160.
- Lyberg, L., & Kasprzyk, D. (1991). Data collection methods and measurement error: An overview. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 237-257). New York: Wiley.
- Mohorko, A., de Leeuw, E., & Hox, J. (2013). *Journal of Official Statistics*, 29, 4, 609-622. <http://www.degruyter.com/view/j/jos.2013.29.issue-4/jos-2013-0042/jos-2013-0042.xml>
- Schonlau, M. (2012). Stata-ado package “hotdeckvar” – Single hotdeck imputation. from <http://www.schonlau.net/stata/>

-
- Schouten, B., Van den Brakel, J., Buelens, B., & Klausch, T. (2013). Disentangling mode-specific selection and measurement bias in social surveys. *Social Science Research*, 42, 1555-1570.
- Tourangeau, R. (2013). Confronting the challenges of household surveys by mixing modes. Keynote address at the 2013 Federal Committee on Statistical Methodology Research Conference, Nov 4, 2013, Washington DC. <http://www.copafs.org/UserFiles/file/fcsm/Tourangeau2013FCSMResearchKeynote.pdf>
- Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). *The science of web surveys*. Oxford: Oxford University Press.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.
- Wasmer, M., Scholz, E., Blohm, M., Walter, J., & Jutz, R. (2012). Konzeption und Durchführung der "Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften" (ALLBUS) 2010. [Concept and implementation of the German General Social Survey (ALLBUS) 2010] *GESIS Technical Report* (Vol. 2012/12). Mannheim: GESIS – Leibniz-Institute for the Social Sciences.

Appendix

Table A1 Distributions of the demographic variables used for weighting

Variable	GOPP Questionnaires 1-5 & 7			ALLBUS 2010 Internet users			ALLBUS 2012 Internet users		
	%	SE	CI	%	SE	CI	%	SE	CI
Male	53.1	.015	[50.1, 56.0]	51.5	.012	[49.3, 53.8]	51.6	.010	[46.4, 50.3]
Age groups									
18-24	13.0	.010	[11.2, 15.1]	13.9	.008	[12.4, 15.6]	13.9	.007	[12.6, 15.3]
25-34	19.8	.012	[17.6, 22.3]	18.6	.009	[16.9, 20.4]	17.4	.008	[15.9, 18.9]
35-49	34.9	.014	[32.2, 37.8]	35.4	.011	[33.2, 37.6]	31.4 ^a	.009	[29.6, 33.3]
50-64	23.4	.013	[21.0, 26.0]	23.2	.010	[21.4, 25.2]	27.7 ^a	.009	[26.0, 29.5]
65+	8.8	.008	[7.3, 10.6]	8.9	.007	[7.6, 10.2]	9.6	.006	[8.5, 10.9]
Education									
low	12.5	.010	[10.7, 14.6]	21.4	.010	[19.6, 23.4]	20.9	.008	[19.4, 22.6]
medium	30.0	.014	[27.4, 32.7]	37.8	.011	[35.6, 40.0]	39.4	.010	[37.5, 41.3]
high	57.5	.015	[54.6, 60.4]	40.8	.011	[38.5, 43.0]	39.7	.010	[37.8, 41.6]

^a significant differences ($p < 0.01$) between ALLBUS 2010 and 2012. N for GOPP is 1114 respondents, which includes the respondents who completed at least one of the questionnaires 1 to 5, or 7. SE is short for standard error, CI is short for confidence interval.

Table A2 Distributions of demographic characteristics in the GOPP over questionnaires (percentages)

Variable/Q	1	2	3	4	5	7
Male	52.97 (016) [49.9, 56.1]	52.74 (017) [49.4, 56.1]	51.75 (018) [48.3, 55.2]	51.48 (018) [48.0, 55.0]	51.91 (018) [48.4, 55.5]	51.99 (019) [48.4, 55.6]
Age groups						
18–24	13.37 (011) [11.4, 15.6]	12.53 (011) [10.5, 15.0]	12.12 (012) [10.0, 14.6]	12.52 (012) [10.4, 15.0]	13.27 (012) [11.0, 15.9]	12.35 (012) [10.1, 14.9]
25–34	19.31 (012) [17.0, 21.9]	19.21 (014) [16.7, 22.0]	19.25 (014) [16.7, 22.1]	18.45 (014) [15.9, 21.3]	18.53 (014) [15.9, 21.5]	19.20 (015) [16.5, 22.2]
35–49	34.85 (015) [32.0, 37.8]	34.73 (016) [31.6, 38.0]	35.00 (017) [31.8, 38.3]	35.10 (017) [31.8, 38.5]	34.03 (017) [30.7, 37.5]	34.29 (018) [30.9, 37.8]
50–64	23.47 (013) [20.9, 26.2]	24.22 (015) [21.4, 27.2]	24.50 (015) [21.6, 27.6]	24.65 (015) [21.7, 27.8]	24.97 (017) [22.0, 28.2]	24.97 (016) [22.0, 28.2]
65+	9.00 (009) [7.4, 10.9]	9.31 (010) [7.5, 11.5]	9.13 (010) [7.3, 11.3]	9.28 (010) [7.4, 11.6]	9.20 (010) [7.3, 11.5]	9.19 (011) [7.3, 11.5]
Education						
low	12.18 (010) [10.3, 14.3]	11.81 (011) [9.8, 14.2]	10.25 (011) [8.3, 12.6]	10.06 (011) [8.1, 12.4]	10.25 (011) [8.3, 12.6]	10.01 (011) [8.0, 12.4]
medium	29.50 (014) [26.8, 32.4]	29.00 (016) [26.0, 32.2]	28.00 (016) [25.0, 31.2]	28.52 (016) [25.4, 31.8]	27.73 (016) [24.7, 31.0]	27.57 (017) [24.4, 30.9]
high	58.32 (016) [55.2, 61.3]	59.19 (017) [55.8, 62.5]	61.75 (017) [58.3, 65.1]	61.42 (017) [57.9, 64.8]	62.02 (018) [58.5, 65.4]	62.41 (018) [58.8, 65.9]
missing on age	2.08	1.55	1.88	1.68	1.45	1.37
missing on education	8.02	3.34	3.88	3.75	3.82	3.71
missing design weight	7.13	1.91	2.75	2.84	2.50	2.47
N	1010	838	800	775	761	729

Note: standard errors in parentheses, 95%-confidence intervals in brackets. The data are not weighted. Q is short for GOPP questionnaire, SE is short for standard error, CI is short for confidence interval.

Table A3 Post-stratification weights obtained with IPF-Weighting

Questionnaire	1	2	3	4	5	7
N	1010	838	800	775	761	729
Min	0.585	0.539	0.520	0.510	0.507	0.498
Max	2.299	2.367	2.445	3.268	2.956	3.086
1% percentile	0.585	0.539	0.520	0.510	0.507	0.498
99% percentile	2.044	2.137	2.907	2.514	2.528	2.558

Calculation of the effect sizes:

Formulas used for calculating the unweighted effect sizes (Lipsey & Wilson, 2001, p. 172 ff.):

The standardized mean difference effect size (ES_{sm}):

$$ES_{sm} = \frac{\bar{X}_1 - \bar{X}_2}{s_{pooled}}, \quad s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

The approximations based on dichotomous data: $ES_{sm} = \arcsine(p_1) - \arcsine(p_2)$, where p is the proportion for each group (Lipsey & Wilson, 2001, p. 198-200).

Formulas used for calculating the mean effect sizes:

$$\overline{ES} = \frac{\sum_{i=1}^k (w_i \times ES_i)}{\sum_{i=1}^k w_i}; \quad w_i = \frac{1}{SE_i^2}$$

where w is the inverse variance weight, ES is the unweighted effect size (d), and SE is the standard error of the difference between the effect sizes (Lipsey & Wilson, 2001, p. 114).

For the calculation of the effect sizes the effect size calculator at http://www.campbellcollaboration.org/resources/effect_size_input.php was used.

We used the unweighted sample sizes to calculate the mean effects.

Questions used in the analysis

Asterisk marks the questions where a card with answer options was offered to the respondents by interviewers in ALLBUS interviews. Question wordings are provided as translated by ALLBUS team, original German question wordings were identical among surveys unless otherwise noted.

German economy*

How would you generally rate the current economic situation in Germany?

Very good

Good

Partly good/partly bad

Bad

Very bad

Own financial situation*

And your own current financial situation?

Very good

Good

Partly good/partly bad

Bad

Very bad

German economy in 1 year*

What do you think the economic situation in Germany will be like in one year?

Considerably better than today

Somewhat better than today

The same

Somewhat worse than today

Considerably worse than today

Own financial situation in 1 year*

And what will your own financial situation be like in one year?

Considerably better than today

Somewhat better than today

The same

Somewhat worse than today

Considerably worse than today

General health*

A question about your health: How would you describe your health in general?

Very good

Good

Satisfactory

Poor

Bad

Employment status*

And now let's continue with employment and your occupation. Which of the categories on the card applies to you?

Full time employment

Part ("half") time employment

Less than part ("half") time employment

Not working

If there are difficulties referring to the classification, here are some hints for you:

Trainees are considered employees in a regular occupation.

Family members assisting in a family business who are full- time or part- time ("halftime") employees in the business of a household or a family member, without having a formal contract, are also considered employees in a regular occupation (either full- time or part- time).

"Employed less than part- time" are persons who are gainfully employed while, at the same time, one of the following applies:

Attend a full- time school (pupils and students),

Are registered as unemployed or

Draw a retirement benefit / pension as a result of previous employment.

Persons on maternity / parental leave or on another type of leave of absence are not considered employees in a regular occupation.

Marital status*

What is your marital status? Are you...

Married and living with your spouse

Married and living apart

Widowed

Divorced

Never married

Civil partnership, living together

Civil partnership, living apart

Registered partner deceased

Civil partnership dissolved

Note that in GOPP this question involved a filter: "civil partnership" category was added to the first five answer options and followed by the four civil partnership answer choices in case a respondent selected the category "civil partnership".

Religiosity*

Would you describe yourself as tending to be religious or tending to be not religious?

(1) “religious”

(10) “not religious”

Note: 10pt-scale, in ALLBUS end labels “religious”, “not religious” and random letters for scale points, in GOPP end labels only.

Confession*

May I ask what religious confession you belong to?

The Roman Catholic church

The German Protestant church (excluding free churches)

A Protestant free church

Another Christian denomination

Another non-Christian religion

No religious affiliation

Note: In GOPP “May we ask what religious confession you belong to?”

Frequency of church attendance

As a rule, how often do you go to church?

More than once a week

Once a week

Between one and three times a month

Several times a year

Less

Never

Note: in ALLBUS 2012 the question is split into two: “How often go to church?” for respondents who belong to a Christian religious denomination and “How often go you go to church or to mosque, synagogue or other chapel?”

Born in Germany

Were you born within the current borders of Germany?

Yes

No

Citizenship

What citizenship do you have? If you have several citizenships, please name all of them.

Germany

Greece

Italy

Former Yugoslavia:

Bosnia and Herzegovina,

Serbia,

Croatia,

Macedonia,

Slovenia

Poland

Turkey

Other country, please enter _____

None, stateless

Self-assessed social class

There is a lot of talk about social class these days. What class would you describe yourself as belonging to?

Lower class

Working class

Middle class

Upper middle class

Upper class

Ownership of dwelling*

The next question deals with the accommodation you/your family live in. Please tell me which of the categories on the card applies to you/your family.

Sublet

In an official/company flat

In subsidized municipal housing

In a rented flat (not subsidized housing)

In a rented house (detached/semi-detached)

In a flat owned by you or your family

In a house owned by you or your family

Other type of accommodation, please enter _____

Attitudes

I'm going to read you some statements now. Please tell me after each one whether you have the same or a different opinion.

No matter what some people say, life for ordinary people is getting worse rather than better.

With the future looking as it does, it's almost irresponsible to bring children into the world.

Most politicians are not really interested at all in the problems of ordinary people.

Most people don't really care in the slightest what happens to others.

Have the same opinion

Have a different opinion

Note: question wording for GOPP: "Please indicate for each statement whether you have the same or a different opinion."

Left-right orientation*

Many people use the terms "left" and "right" when they want to describe different political views. Here we have a scale which runs from left to right. Thinking of your own political views, where would you place these on this scale?

(1) "left"

(10) "right"

Note: 10pt-scale, in ALLBUS end labels "left", "right" and random letters for scale points, in GOPP end labels only. The text "Here we have a scale which runs from left to right." omitted in GOPP.

Table A4 Comparison of the GESIS Online Panel Pilot to the subsamples of Internet users from two ALLBUS surveys: attitudinal and factual questions with post-stratification, no imputation of variables education and age

Variable	Q	GOPP			ALLBUS 2010 Internet users			ALLBUS 2012 Internet users		
		Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
<i>Attitudinal:</i>										
German economy	1	3.788	0.121	[3.551, 4.025]	3.086***	0.019	[3.048, 3.124]	3.351***	0.018	[3.317, 3.385]
German economy in 1 year	1	2.427	0.242	[1.952, 2.902]	3.077**	0.021	[3.036, 3.117]	2.667	0.016	[2.636, 2.698]
Self-rated financial situation	1	3.210	0.120	[2.974, 3.445]	3.471*	0.020	[3.433, 3.510]	3.593**	0.017	[3.560, 3.626]
Self-rated financial situation in 1 year	1	3.112	0.066	[2.982, 3.241]	3.182	0.017	[3.148, 3.216]	3.157	0.015	[3.128, 3.186]
Health status	1	3.874	0.076	[3.725, 4.024]	3.823	0.023	[3.778, 3.868]	3.748	0.020	[3.709, 3.787]
Religiosity	4	5.734	0.145	[4.334, 7.134]	5.974	0.073	[5.832, 6.117]	6.341	0.063	[6.218, 6.466]
Self-assessed social class	5									
lower class		3.60%	0.011	[1.96, 6.51]	2.27%	0.003	[1.67, 3.06]	1.55%	0.003	[1.13, 2.13]
working class		26.18%	0.025	[21.56, 31.39]	20.23%*	0.010	[18.42, 22.17]	22.21%	0.009	[20.55, 23.97]
middle class		58.11%	0.025	[53.09, 62.96]	63.38%	0.012	[61.08, 65.63]	63.24%	0.010	[61.21, 65.22]
upper middle class		11.65%	0.012	[9.50, 14.21]	13.80%	0.008	[12.22, 15.55]	12.36%	0.007	[11.03, 13.83]
upper class		0.46%	0.002	[0.18, 1.20]	0.32%	0.001	[0.14, 0.73]	0.64%	0.002	[0.38, 1.09]
Life is getting worse	7	80.17%	0.029	[73.90, 85.24]	82.00%	0.009	[80.10, 83.74]	75.58%	0.009	[73.70, 77.36]
Irresponsible to have children	7	37.78%	0.034	[31.30, 44.73]	37.35%	0.012	[35.10, 39.65]	31.06%	0.010	[29.17, 33.01]
Politicians not interested	7	81.14%	0.028	[75.03, 86.03]	77.42%	0.010	[75.37, 79.35]	73.56%*	0.010	[71.64, 75.40]

Table A4 continued

Variable	Q	GOPP			ALLBUS 2010 Internet users			ALLBUS 2012 Internet users		
		Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
People don't care about others	7	62.74%	0.034	[55.82, 69.17]	71.72%*	0.011	[69.55, 73.79]	68.63%	0.010	[66.64, 70.55]
Left-right orientation	7	5.072	0.112	[4.852, 5.291]	5.143*	0.041	[5.063, 5.222]	4.946	0.036	[4.876, 5.016]
<i>Factual:</i>										
Working for pay	2	83.71%	0.016	[80.30, 86.62]	72.41%***	0.011	[70.26, 74.46]	73.60%***	0.009	[71.75, 75.37]
Legal marital status	3	49.70%	0.024	[44.92, 54.49]	56.89%***	0.012	[54.56, 59.19]	56.23%*	0.011	[54.16, 58.28]
Confession	4	85.53%	0.082	[61.73, 95.59]	70.57%	0.011	[68.47, 72.59]	71.29%	0.009	[69.46, 73.04]
Frequency of church attendance	4	2.599	0.227	[2.153, 3.044]	2.301	0.029	[2.244, 2.359]	2.286	0.026	[2.234, 2.337]
Born in Germany	4	97.19%	0.017	[91.03, 99.16]	87.44%***	0.008	[85.76, 88.94]	86.25%***	0.008	[84.69, 87.67]
German citizenship	4	99.07%	0.006	[96.53, 99.76]	94.53%***	0.006	[93.32, 95.52]	93.50%***	0.006	[92.33, 94.50]
Owner of dwelling	5	45.31%	0.024	[40.59, 50.13]	56.64%***	0.012	[54.30, 58.95]	58.06%***	0.010	[56.99, 60.10]

Note: * $p < .05$, ** $p < .01$, *** $p < .001$. Differences between ALLBUS 2010 and 2012 in italic bold, Q short for GOPP questionnaire, SE short for standard error, CI short for confidence interval. N for GOPP (started): Q1=1010; Q2=838; Q3=800; Q4=775; Q5=761; Q7=729. Weighted data: ALLBUS 2010 design, GOPP & ALLBUS 2012 design and post-stratification. Significance tests are pairwise t-tests for continuous/ordinal variables or tests of proportions.

Datasets used in preparing the article:

1 **GESIS Online Panel Pilot (GOPP)**

Fieldwork: February 2011–May 2012

Data Collection agency: GESIS – Leibniz Institute for the Social Sciences
Archive Link to demographic variables dataset: <http://info1.gesis.org/dbksearch19/SDesc2.asp?no=5590&tab=0&ll=10¬abs=&af=&nf=1&search=&search2=&db=E>

DOI: 10.4232/1.11577

Archive Link to the first online questionnaire:

<http://info1.gesis.org/dbksearch19/SDesc2.asp?no=5582&tab=0&ll=10¬abs=&af=&nf=1&search=&search2=&db=E>

DOI: 10.4232/1.11570 (includes three datasets for three study parts).

Description: The GESIS Online Panel Pilot Study consists of an initial CATI-recruitment interview (probability based sample, including landline and mobile telephone numbers), followed by eight monthly online surveys. Respondents to the CATI-recruitment interview (ZA5581), who had agreed to participate in the online panel by providing their email addresses, received an email invitation with a link to the first online survey. In the following, the respondents received such invitations for participation every month for the total duration of eight months. The overall concept for the online part was to have a leading topic for each monthly questionnaire and to include some demographic questions in every wave in order not to burden respondents with unpleasant factual questions. The most important demographic information had been collected during the recruitment interview. (For participants of short interviews the demographic information, which was supposed to be collected on telephone during the recruitment interview, was collected online at a later time.) An integrated dataset of the demographic information of all waves (ZA5590) is provided. The average duration of each monthly questionnaire was conceptualized to be around fifteen minutes. Generally respondents have one month for filling out a questionnaire, but they also had the possibility to do so at a later stage. The last date for filling out the questionnaires for the first study was 31 December 2011, for the second and the third study it was 31 May 2012.

The design has both longitudinal and cross-sectional aspects: monthly waves are designed as cross-sections with various topics; the longitudinal component includes survey evaluation questions, which are asked at the end of every monthly questionnaire. Most of the questions used in the online questionnaires were originally conducted within other German and international surveys. The reasons for replicating the questions were twofold: first,

in the course of the project the feasibility of asking questions via the Internet which are of interest for the social sciences were assessed; second, for the purposes of data quality assessment, comparisons with external benchmarks were planned as part of the project. For these reasons, most questions stem from the German General Social Survey (ALLBUS) 2010 and the German version of the European Social Survey 2010, large cross-sectional face-to-face surveys conducted by GESIS. Some other surveys, from which questions were implemented in the course of the panel, include questions of the German Socio-economic Panel (GSOEP 2008), the Eurobarometer, and the International Social Survey Programme (ISSP). Borrowed questions were generally implemented using the original question wording. In some cases these were tailored to the self-administered mode.

The main topics of the monthly questionnaires were: Wave 1: Multitopic introductory wave (ZA5582), Wave 2: Education and employment (ZA5583), Wave 3: Family life (ZA5584), Wave 4: Religion and values (ZA5585), Wave 5: Ecology (ZA5586), Wave 6: Social networks (ZA5587), Wave 7: Politics (ZA5588), Wave 8: Multitopic (focus on personality) (ZA5589).

Link to the study description:

<http://info1.gesis.org/dbksearch19/download.asp?db=E&id=48472>

2 The German General Social Survey (ALLBUS) 2010

The following information about the study was taken from the webpages of the study.

Fieldwork: May 2010–November 2010

Data Collection agency: TNS Infratest Sozialforschung, Munich

Archive Link: <http://www.gesis.org/en/allbus/data-access/> and http://info1.gesis.org/dbksearch13/SDesc.asp?nf=1&search=ALLBUS&field=TI&DB=E&sort=dbk_ext.SN+DESC&maxRec=100&ll=10&tab=0&choose_study_number=ZA4612

DOI: 10.4232/1.10760

Description: The German General Social Survey (ALLBUS) is a biennial survey that has been conducted since 1980 on the attitudes, behaviour, and social structure of persons resident in Germany. A representative cross-section of the population is questioned using face-to-face interviews. As a service to social scientific research and teaching, ALLBUS data are disseminated to all interested persons and institutions as soon as the data and documentation have been prepared. ALLBUS is a substantively rich and methodologically sophisticated database which can be used for a variety of analytical purposes:

- to describe and analyse attitudes, behaviour and social structure of Germans with up-to-date cross-sectional data,
- for longitudinal analysis of German society (ALLBUS time series, replicated questions from other survey studies),
- for international comparative analysis (ALLBUS includes questions also asked in the American General Social Survey (GSS) and in the ISSP),
- to examine methodical issues (e.g., checking for context effects, question effects, nonresponse).

Link to the study description: <http://www.gesis.org/en/allbus/general-information/>; see also <http://www.gesis.org/en/allbus/study-profiles/2010/>

3 **The German General Social Survey (ALLBUS) 2012**

The following information about the study was taken from the webpages of the study.

Fieldwork: April 2012–September 2012

Data Collection agency: TNS Infratest Sozialforschung, Munich

Archive Link: <http://www.gesis.org/en/allbus/data-access/> and http://info1.gesis.org/dbksearch13/SDesc.asp?nf=1&search=ALLBUS&field=TI&DB=E&sort=dbk_ext.SN+DESC&maxRec=100&ll=10&tab=0 choose study number “ZA4614”

DOI: 10.4232/1.11753

Link to the study description: <http://www.gesis.org/en/allbus/study-profiles/2012/>

Table A5 Variable labels, value labels and recoding procedures for variables used in the analyses

Variable	GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
	Variable label	Value label	Variable label	Value label	Variable label	Value label	
Current economic situation in Germany	h1ulgereco h2ulgereco h3ulgereco	1 Very good 2 Good 3 Partly good/partly bad 4 Bad 5 Very bad . Missing	v79	1 Very good 2 Good 3 Partly good/partly bad 4 Bad 5 Very bad 8 Don't know 9 Missing (No answer)	v118	1 Very good 2 Good 3 Partly good/partly bad 4 Bad 5 Very bad 8 Don't know 9 Missing (No answer)	5 Very good 4 Good 3 Partly good/partly bad 2 Bad 1 Very bad . Missing (ALLBUS 8, 9)
Own current financial situation	h1ulsifeco h2ulsifeco h3ulsifeco	1 Very good 2 Good 3 Partly good/partly bad 4 Bad 5 Very bad . Missing	v80	1 Very good 2 Good 3 Partly good/partly bad 4 Bad 5 Very bad 8 Don't know 9 Missing (No answer)	v119	1 Very good 2 Good 3 Partly good/partly bad 4 Bad 5 Very bad 8 Don't know 9 Missing (No answer)	5 Very good 4 Good 3 Partly good/partly bad 2 Bad 1 Very bad . Missing (ALLBUS 8, 9)
Economic situation in Germany in one year	h1ulgereco1 h2ulgereco1 h3ulgereco1	1 Considerably better than today 2 Somewhat better than today 3 The same 4 Somewhat worse than today 5 Considerably worse than today . Missing	v81	1 Considerably better than today 2 Somewhat better than today 3 The same 4 Somewhat worse than today 5 Considerably worse than today 8 Don't know 9 Missing (No answer)	v120	1 Considerably better than today 2 Somewhat better than today 3 The same 4 Somewhat worse than today 5 Considerably worse than today 8 Don't know 9 Missing (No answer)	5 Considerably better than today 4 Somewhat better than today 3 The same 2 Somewhat worse than today 1 Considerably worse than today . Missing (ALLBUS 8, 9)

GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
Variable	Variable label	Variable label	Value label	Variable label	Value label	
Own economic situation in one year	h1u1sfecol	v82	1 Considerably better than today	v121	1 Considerably better than today	5 Considerably better than today
	h2u1sfecol		2 Somewhat better than today		2 Somewhat better than today	4 Somewhat better than today
	h3u1sfecol		3 The same		3 The same	3 The same
			4 Somewhat worse than today		4 Somewhat worse than today	2 Somewhat worse than today
			5 Considerably worse than today		5 Considerably worse than today	1 Considerably worse than today
		. Missing	8 Don't know		8 Don't know	. Missing (ALLBUS 8, 9)
		9 Missing (No answer)	9 Missing (No answer)		9 Missing (No answer)	
General health	h1u1health	v398	1 Very good	v268	1 Very good	5 Very good
	h2u1health		2 Good		2 Good	4 Good
	h3u1health		3 Satisfactory		3 Satisfactory	3 Satisfactory
			4 Poor		4 Poor	2 Poor
			5 Bad		5 Bad	1 Bad
		. Missing	9 No answer		9 No answer	. Missing (ALLBUS 9)
Employment status	h1u2employed	v340	1 Full time employment	v245	1 Full time employment	1 Working for pay (1, 2, 3)
	h2u2employed		2 Part ("half") time employment		2 Part ("half") time employment	0 Not working for pay (4)
	h3u2employed		3 Less than part ("half") time employment		3 Less than part ("half") time employment	. Missing (ALLBUS 9)
			4 Not working		4 Not working	
			. Missing	9 Missing (No answer)		9 Missing (No answer)

Variable	GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
	Variable label	Value label	Variable label	Value label	Variable label	Value label	
Marital status	h1u3famstdu3 h2u3famstdu3 h3u3famstdu3	1 Married and living with a spouse 2 Married and living apart 3 Widowed 4 Divorced 5 Never married 6 Civil partnership . Missing r Missing (Refusal to provide answer)	v405	1 Married and living with your spouse 2 Married and living apart 3 Widowed 4 Divorced 5 Never married 6 Civil partnership, living together	v274	1 Married and living with your spouse 2 Married and living apart 3 Widowed 4 Divorced 5 Never married 6 Civil partnership, living together 99 Missing (No answer)	1 Married (ALLBUS 1, 2, 6; GOPP variable famstdu3 1, 2 & variable elpstd 1, 2) 0 Not married (ALLBUS 3, 4, 5; GOPP variable famstdu3 3, 4, 5, 6 & variable elpstd 3, 4) . Missing (ALLBUS 99)
Religiosity	h1u4relig h2u4relig h3u4relig	1 "religious" (10) "not religious" . Missing	v729	1 "religious" (10) "not religious" 99 Missing (No answer)	v134	1 "religious" (10) "not religious" 99 Missing (No answer)	(10) religious (1) not religious . Missing (ALLBUS 99)

		GOPP		ALLBUS 2010		ALLBUS 2012			
Variable	Variable label	Value label	Variable label	Value label	Variable label	Value label	Variable label	Value label	How the variable was recoded for analysis
Confession	h1u4religcomm1 h2u4religcomm1 h3u4religcomm1	1 The Roman Catholic church 2 The German Protestant church (excluding free churches) 3 A Protestant free church 4 Another Christian denomination 5 Another non-Christian religion 6 No religious affiliation . Missing	v730	1 The Roman Catholic church 2 The German Protestant church (excluding free churches) 3 A Protestant free church 4 Another Christian denomination 5 Another non-Christian religion 6 No religious affiliation 7 Missing (Refused to answer) 9 Missing (No answer)	v188	1 The Roman Catholic church 2 The German Protestant church (excluding free churches) 3 A Protestant free church 4 Another Christian denomination 5 Another non-Christian religion 6 No religious affiliation 7 Missing (Refused to answer) 9 Missing (No answer)			1 Confession named (1, 2, 3, 4, 5) 0 Confession not named (6) . Missing (ALLBUS 7, 9)
Frequency of church attendance	h1u4church h2u4church h3u4church	1 More than once a week 2 Once a week 3 Between one and three times a month 4 Several times a year 5 Less 6 Never . Missing	v732	1 More than once a week 2 Once a week 3 Between one and three times a month 4 Several times a year 5 Less 6 Never 9 Missing (No answer)	v269	0 Not applicable 1 More than once a week 2 Once a week 3 Between one and three times a month 4 Several times a year 5 Less 6 Never 9 Missing (No answer)			6 More than once a week 5 Once a week 4 Between one and three times a month 3 Several times a year 2 Less 1 Never . Missing (ALLBUS 9)
Frequency of attendance of a temple	n/a	n/a	n/a	n/a	v270	0 Not applicable 1 More than once a week 2 Once a week			6 More than once a week 5 Once a week 4 Between one and three times a month

GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
Variable	Variable label	Value label	Vari-able label	Value label	Vari-able label	
Born in Germany	h1u4birth_area h2u4birth_area h3u4birth_area	1 Yes 2 No	v303	1 Yes 2 No	v222	4 Between one and three times a month 3 Several times a year 2 Less 1 Never . Missing (ALLBUS 9)
Citizenship	h1u4citizenship1 h2u4citizenship1 h3u4citizenship1	0 No German citizenship 1 Has German citizenship . Missing	v6	1 Yes 2 Yes, along with other citizenship 3 No 4 Stateless 9 Missing (No answer)	v6	1 Yes (ALLBUS 1, 2) 0 No (ALLBUS 3, 4) . Missing (ALLBUS 9)
Self-assessed social class	h1u5s1fses h2u5s1fses h3u5s1fses	1 Upper class 2 Upper middle class 3 Middle class 4 Working class 5 Lower class	v33	1 Lower class 2 Working class 3 Middle class 4 Upper middle class 5 Upper class 6 None of the above 7 Missing (Refused to answer) 8 Don't know 9 Missing (No answer)	v17	1 Lower class (GOPP 5) 2 Working class (GOPP 4) 3 Middle class 4 Upper middle class (GOPP 2) 5 Upper class (GOPP 1) . Missing (ALLBUS 6, 7, 8, 9)

Variable	GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
	Variable label	Value label	Variable label	Value label	Variable label	Value label	
Ownership of dwelling	h1u5owner h2u5owner h3u5owner	1 Sublet 2 In an official /company flat 3 In subsidized municipal housing 4 In a rented flat (not subsidized housing) 5 In a rented house (detached/semi-detached) 6 In a flat owned by you or your family 7 In a house owned by you or your family 8 Other type of accommodation, please enter 9 Missing	v714	1 Sublet 2 In an official/company flat 3 In subsidized municipal housing 4 In a rented flat (not subsidized housing) 5 In a rented house (detached/semi-detached) 6 In a flat owned by you or your family 7 In a house owned by you or your family 8 Other type of accommodation, please enter 9 Missing (No answer)	v501	1 Sublet 2 In an official/company flat 3 In subsidized municipal housing 4 In a rented flat (not subsidized housing) 5 In a rented house (detached/semi-detached) 6 In a flat owned by you or your family 7 In a house owned by you or your family 8 Other type of accommodation, please enter 9 Missing (No answer)	1 Yes (6, 7) 0 No (1, 2, 3, 4, 5, 8) . Missing (ALLBUS 99)
No matter what some people say, life for ordinary people is getting worse rather than better.	h1u7attitude1 h2u7attitude1 h3u7attitude1	1 Have the same opinion 2 Have a different opinion 3 Missing	v61	1 Have the same opinion 2 Have a different opinion 3 Don't know 4 Missing (No answer)	v32	1 Have the same opinion 2 Have a different opinion 3 Don't know 4 Missing (No answer)	1 Yes (ALLBUS 1) 0 No (ALLBUS 2) . Missing (ALLBUS 8, 9)

Variable	GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
	Variable label	Value label	Variable label	Value label	Variable label	Value label	
With the future looking as it does, it's almost irresponsible to bring children into the world.	h1u7attitude2 h2u7attitude2 h3u7attitude2	1 Have the same opinion 2 Have a different opinion . Missing	V62	1 Have the same opinion 2 Have a different opinion 8 Don't know 9 Missing (No answer)	v33	1 Have the same opinion 2 Have a different opinion 8 Don't know 9 Missing (No answer)	1 Yes (ALLBUS 1) 0 No (ALLBUS 2) . Missing (ALLBUS 8, 9)
Most politicians are not really interested at all in the problems of ordinary people.	h1u7attitude3 h2u7attitude3 h3u7attitude3	1 Have the same opinion 2 Have a different opinion . Missing	V63	1 Have the same opinion 2 Have a different opinion 8 Don't know 9 Missing (No answer)	v34	1 Have the same opinion 2 Have a different opinion 8 Don't know 9 Missing (No answer)	1 Yes (ALLBUS 1) 0 No (ALLBUS 2) . Missing (ALLBUS 8, 9)
Most people don't really care in the slightest what happens to others.	h1u7attitude4 h2u7attitude4 h3u7attitude4	1 Have the same opinion 2 Have a different opinion . Missing	V64	1 Have the same opinion 2 Have a different opinion 8 Don't know 9 Missing (No answer)	v35	1 Have the same opinion 2 Have a different opinion 8 Don't know 9 Missing (No answer)	1 Yes (ALLBUS 1) 0 No (ALLBUS 2) . Missing (ALLBUS 8, 9)
Left-right orientation	h1u7leftright h2u7leftright h3u7leftright	(1) "left" (10) "right" . Missing	v78	(1) "left" (10) "right" 99 Missing (No answer)	v101	(1) "left" (10) "right" 99 Missing (No answer)	. Missing (ALLBUS 99)

Variable	GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
	Variable label	Value label	Variable label	Value label	Variable label	Value label	
Sex of respondent	sex_cati sex_onl	1 Male 2 Female . Missing	v298	1 Male 2 Female	v217	1 Male 2 Female	1 Male 0 Female
Age of respondent/ Year of birth	ybirth_cati age_onl	. Missing . r Missing (Refusal to provide answer)	v301	999 Missing	v220	999 Missing	GOPP: 2011-ybirth_cati
Education: highest school leaving certificate	graduation_cati_hl	1 Still at school (general education) 2 Still at school (vocational education) 3 Finished school without school leaving certificate 4 Lowest formal qualification of 5 Germany's tripartite secondary school system, after 8 or 9 years of schooling 5 Intermediary secondary qualification, after 10 years of schooling 6 Secondary qualification from Eastern Germany (Polytechnic secondary school)	v327	1 Finished school without school leaving certificate 2 Lowest formal qualification of Germany's tripartite secondary school system, after 8 or 9 years of schooling 3 Intermediary secondary qualification, after 10 years of schooling 4 Certificate fulfilling entrance requirements to study at a polytechnical college 5 Higher qualification, entitling holders to study at a university 6 Other school leaving certificate 7 Still at school 99 Missing (No answer)	v230	1 Finished school without school leaving certificate 2 Lowest formal qualification of Germany's tripartite secondary school system, after 8 or 9 years of schooling 3 Intermediary secondary qualification, after 10 years of schooling 4 Certificate fulfilling entrance requirements to study at a polytechnical college 5 Higher qualification, entitling holders to study at a university 6 Other school leaving certificate 7 Still at school 99 Missing (No answer)	1 Low (ALLBUS 1,2) 2 Middle (ALLBUS 3) 3 High (ALLBUS 4, 5) . Missing (ALLBUS 6, 7, 99)

GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
Variable	Variable label	Variable label	Value label	Variable label	Value label	
	Value label					
	7 Certificate fulfilling entrance requirements to study at a polytechnical college 8 Higher qualification, entitling holders to study at a university 9 Other school leaving certificate . Missing .r. Missing (Refusal to provide answer)					
graduation_cati_h2	1 Still at school 2 Finished school without school leaving certificate 3 Lowest formal qualification of Germany's tripartite secondary school system, after 8 or 9 years of schooling 4 Intermediary secondary qualification, after 10 years of schooling 5 Secondary qualification from Eastern Germany (Polytechnic secondary school, 8 or 9 years)					

GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
Variable	Variable label	Value label	Variable label	Value label	Variable label	
		6 Secondary qualification from Eastern Germany (Polytechnic secondary school, 10 years) 7 Certificate fulfilling entrance requirements to study at a polytechnical college 8 Higher qualification, entitling holders to study at a university 9 Higher qualification, entitling holders to study at a university obtained later 10 Other school leaving certificate . Missing .r Missing (Refusal to provide answer)				
Person-related East-West-weight	n/a	n/a	v977	v743		
Internet use for private purposes	n/a	n/a	v772	v530	1 Yes, I use the Internet for private purposes 2 No, I don't use the Internet	1 Yes, I use the Internet for private purposes 2 No, I don't use the Internet

GOPP		ALLBUS 2010		ALLBUS 2012		How the variable was recoded for analysis
Variable	Variable label	Value label	Variable label	Value label	Variable label	
Design weight	pweightcomb		n/a	n/a	n/a	n/a
Study part (CATI)	cati	1 h1 (HS1) 2 h2 (HS2) 3 h3 (HS3)	n/a	n/a	n/a	n/a
Study part (online)	study	1 h1 (HS1) 2 h2 (HS2) 3 h3 (HS3) . Missing	n/a	n/a	n/a	n/a
Recruitment question at the end of the CATI interview	panelquestion	1 Agreed to participate, provided Email 2 Respondent in doubt (homepage, call later) 3 No Email address 4 Not interested, don't want to participate . Missing	n/a	n/a	n/a	n/a
discode	Disposition code for the first online questionnaire	12 Active (not started) 22 Break-off 31 Completed without pause 32 Completed with pause . Missing	n/a	n/a	n/a	n/a
pseudonym	Respondent ID (Online)		n/a	n/a	n/a	n/a

