
gesis

Leibniz-Institut
für Sozialwissenschaften

German Longitudinal
Election Study



Short-term Campaign Panel (GLES 2017)

Wave 1 - 9

ZA6804, Version 7.0.0

Study description

Contents

Preliminary notes.....	2
1 Study characteristics.....	3
2 Conception	4
3 Sampling.....	8
4 Data collection	14
5 Variables in the dataset	28
6 Data preparation.....	48
7 Data quality.....	51
8 Matching of other datasets	57
9 Version history	58
References	63
Links	64
Picture credits.....	64

Preliminary notes

Please note: Working with GLES data

This publication and the corresponding dataset are data of the German Longitudinal Election Study (GLES), released by GESIS in cooperation with the German Society of Electoral Research (Deutsche Gesellschaft für Wahlforschung, DGfW). Despite thorough controlling and statistical processing of the data, GESIS and the DGfW cannot guarantee that this release will satisfy all demands. Errors are documented in the data catalogue's errata list (www.gesis.org/dbk).

If you discover an error while working with GLES data, we would highly appreciate you informing us via e-mail (gles@gesis.org). Please send us the description of the error, the study number (ZA-number), as well as the version of the dataset you are using.

We recommend using the latest version of GLES data at all times. They can be downloaded via the GESIS data catalogue. Links to the direct download can also be found on the GESIS website (www.gesis.org/gles).

Announcement of publication with GLES data

To gain an overview of the actual use of the data, we kindly request users of GLES data to inform us about publications that utilize those data (bibliographic notice, study no. of the used dataset). Publications which are completely or partially based on GLES data will be listed in the official bibliography of GLES. In case of limited access to the publication (e.g. conference papers), we would highly appreciate it if you sent us a PDF-file or a print copy of your publication.

Contact

GESIS – Leibniz-Institute for Social Science

Postfach 122155

68072 Mannheim

E-mail: gles@gesis.org

Citation of GLES data

Please include the following citation in your publication with GLES data:

Roßteutscher, Sigrid; Schmitt-Beck, Rüdiger; Schoen, Harald; Weißels, Bernhard; Wolf, Christof; Gärtner, Lea; Preißinger, Maria; Kratz, Agatha; Wuttke, Alexander (2019): Short-term Campaign Panel (GLES 2017). GESIS Data Archive, Cologne: ZA6804 Datafile Version 7.0.0, doi: 10.4232/1.13323.

1 Study characteristics

1.1. Study no.

ZA6804

1.2. Title

German Longitudinal Election Study, Component 3: Short-term Campaign Panel

1.3. Version

7.0.0, 24.07.2019, doi: 10.4232/1.13323

1.4. Date of collection

Wave 1 (10/06/2016 – 11/10/2016)

Wave 2 (02/16/2017 – 03/03/2017)

Wave 3 (05/11/2017 – 05/23/2017)

Wave 4 (07/06/2017 – 07/17/2017)

Wave 5 (08/17/2017 – 08/28/2017)

Wave 6 (09/04/2017 – 09/13/2017)

Wave 7 (09/18/2017 – 09/23/2017)

Wave 8 (09/27/2017 – 10/09/2017)

Wave 9 (03/15/2018 – 03/26/2018)

Profile wave of the first refreshment sample A1 (07/20/2017 – 08/09/2017)

1.5. Principal investigators

Prof. Dr. Sigrid Roßteutscher (Goethe University Frankfurt am Main)

Prof. Dr. Rüdiger Schmitt-Beck (University of Mannheim)

Prof. Dr. Harald Schoen (University of Mannheim)

Prof. Dr. Bernhard Weißels (Social Science Research Center Berlin)

Prof. Dr. Christof Wolf (GESIS – Leibniz-Institute for the Social Sciences)

1.6. Funding agency

German Research Foundation (Deutsche Forschungsgemeinschaft e. V.)

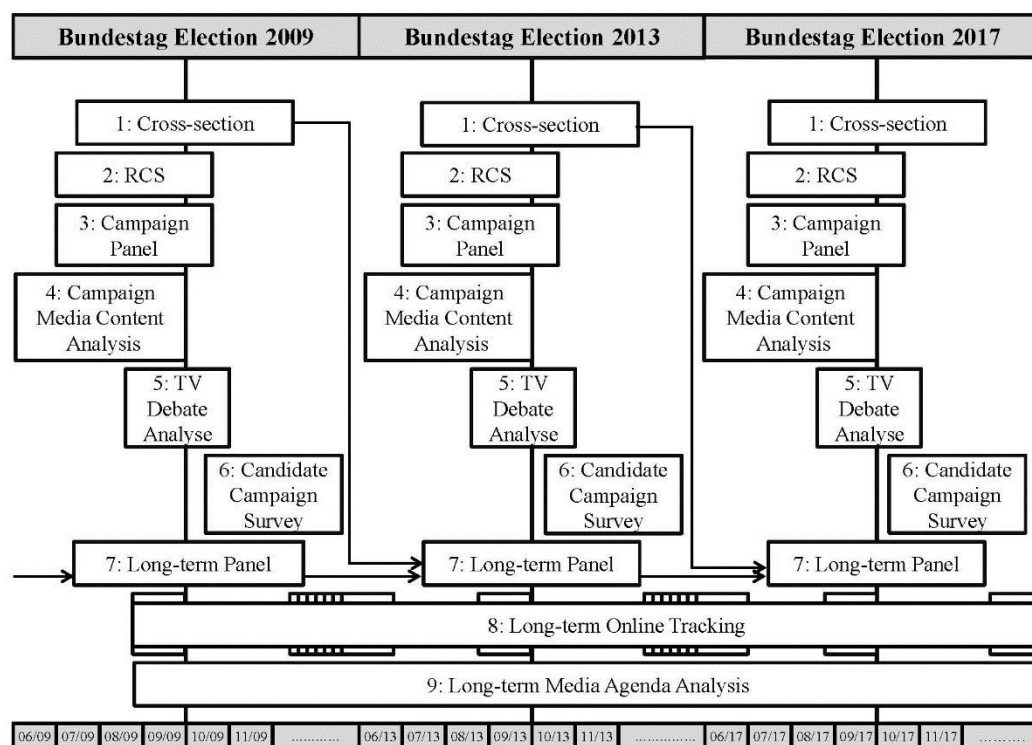
2 Conception

2.1. Background and objective

The German Longitudinal Election Study (GLES) is the largest and most ambitious election study held so far in Germany. GLES started with the 2009 federal election as a long-term project funded by the German Research Foundation (Deutsche Forschungsgemeinschaft) and was continued with the 2013 and 2017 federal elections. Since 2018 the GLES is held by GESIS as institutionalized election study in close cooperation with the German Society for Electoral Studies (Deutsche Gesellschaft für Wahlforschung).

The GLES allows the analysis of electoral behaviour in both cross-sectional and longitudinal perspectives as well as with regard to short-term dynamics during the election campaign (Schmitt-Beck et al. 2010).

Figure 1: The design of the German Longitudinal Election Study (GLES)



The Campaign Panel of the German Longitudinal Election Study (Component 3) allows for an analysis of changes in individuals' political attitudes and political behaviour over the course of an election campaign. To this end, a steady selection of citizens is interviewed at short intervals during the election campaign and immediately after the election. In contrast to its predecessors in 2009 and 2013, the 2017 Campaign Panel has started about 12 months before the federal election. This makes it possible to measure political attitudes and behaviour even in periods outside the election campaign and compare them with those in the election campaign phase.

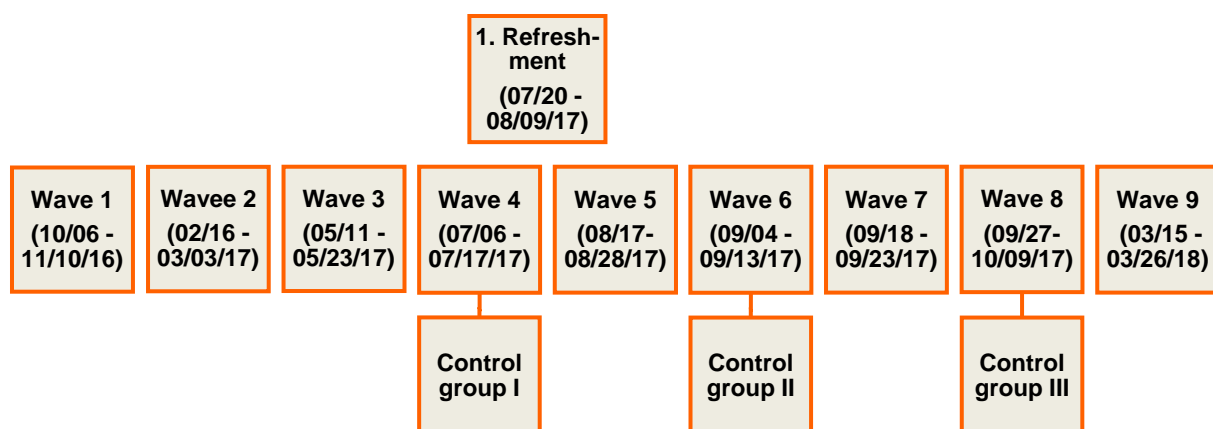
2.2. Study design

A total of nine waves were carried out: The first two surveys, in October 2016 and February 2017, were designed to survey political attitudes and behaviour before the start of the election campaign. Between May and September 2017, five waves were conducted at short intervals so that possible election campaign effects on respondents' attitudes could be tracked. Immediately after the 2017 federal election, another survey was carried out. The ninth wave was conducted after the formation of the government was completed in March 2018.

In addition, control groups are implemented into the 2017 Campaign Panel design, i.e. three independent samples were drawn simultaneously to the fourth, sixth and eighth interview of the 2017 Campaign Panel. The samples included round about 1,000 people who were interviewed with an almost identical¹ questionnaire (see Figure 1). By comparing panel participants and control group participants, it can be analysed whether and to which extent the repeated interviewing of the same participants is influential to the measuring of political attitudes and behaviour patterns. The present study description only refers to the nine waves of the Campaign Panel. The documentation and the data of the control groups were published separately under the ZA-numbers 6805, 6806 and 6807.

Between the fourth and fifth wave, the Campaign Panel was refreshed with new participants. In a separate profile wave, people with the combinations of characteristics for which the highest number of panel dropouts had been recorded up to that point were purposefully invited. From the fifth wave onwards, participants with a completed interview in this profile wave participated on a regular basis in the Campaign Panel. It is planned to interview the participants of the Campaign Panel once or twice a year starting in autumn 2018. Since further refreshment samples will probably be added in the future, we will refer to this group in the following as "1. Refreshment sample".

Figure 2: Waves and control groups in the 2017 Campaign Panel



¹ The questionnaire of the control groups was slightly longer than the questionnaire of the respective panel wave because additional socio-demographic characteristics were collected. For the panel respondents, these characteristics were already known from previous waves.

2.3. Content of the study

In the first two waves, socio-demographics and psychological concepts were measured, which can be assumed to be largely stable over time. The "core questionnaire" includes those questions which are requested in unchanged/original form in every single interview, no matter in which wave it takes place. These include questions on electoral and communication behaviour, evaluations of political parties and politicians or issues perceived as subjectively important. During the election campaign, the individual waves also focus on specific content. The profile wave of the refreshment sample was similar in structure to the first wave and contained questions on socio-demographics and psychological constructs in addition to the core questionnaire. Table 1 gives an overview of the thematic focus of the surveys already carried out.

Table 1: Overview of the nine panel waves' main focus

Contents	Wave 1 (10/06 - 11/10/16)	Wave 2 (02/16 - 03/03/17)	Wave 3 (05/11 - 05/23/17)	Profile wave of the refreshment sample (07/20 - 08/09/17)	Wave 4 (07/06 - 07/17/17)	Wave 5 (08/17 - 08/28/17)	Wave 6 (09/04 - 09/13/17)	Wave 7 (09/18 - 09/23/17)	Wave 8 (09/27 - 10/09/17)	Wave 9 (03/15 - 03/26/18)
Core questionnaire	X	X	X	X	X	X	X	X	X	X
Voting behaviour, retrospective	X								X	X
Ideology, parties		X			X			X		
Ideology, candidates			X				X			
Issues, ego	X	X	X	X	X		X	X	X	X
Issues, parties		X			X			X		
Issues, candidates			X				X			
Candidates			X		X	X	X			
Strategic voting			X			X	X	X	X	X
Economy	X		X			X	X		X	
State election in Schleswig-Holstein			X							
State election in NRW			X							
State election in Niedersachsen						X	X	X	X	
TV debate							X	[X]		
Coalition talks										X
Psychological concepts	X			X						
Socio-demographics, ego and partner	X			X						
Socio-demographics, migrant background		X	[X]							

Explanation: "X" stands for the respective survey wave. [X] implies that only participants of the Campaign Panel who did not take part in earlier waves where this issue had been monitored/administered received these module questions (see also section 6.7.).

3 Sampling

3.1. Geographic coverage

Germany (DE)

3.2. Universe

The target population of the Campaign Panel comprises all German citizens who were eligible to vote in the 2017 election of the German Bundestag. As the present study is an online questionnaire, it was not possible to draw a random sample. Therefore, the sampling population only includes members of the Online Access Panel of the Respondi AG and GapFish GmbH who were entitled to vote. In 2016, the Respondi panel included 65,000-70,000 active German users. According to the definition of Respondi, active users have completed the master data questionnaire after (double-opt-in) registration² and have participated in at least one survey within the last three months.

Since Respondi had difficulties in achieving certain quota targets, participants in the online access panel GapFish GmbH were also invited. The GapFish panel comprises approximately 113,000 respondents whose master data is not older than one year and who took part in at least one survey last year. Like Respondi, GapFish has high-quality standards in recruitment, invitation procedures, incentive models, and panel maintenance, so that the participants of both panels should not differ in quality.

Table 2: Distribution of the socio-demographic characteristics in the Respondi and GapFish access-panel (percentage share)

	Respondi	GapFish
Gender		
Female	62	55
Male	38	45
Education		
Low (i.e., no graduation or graduation after 8 or 9 years of schooling) ("Hauptschulabschluss", "Volksschulabschluss")	19	16
Intermediate (i.e., secondary qualification, graduation after 10 years of schooling) ("Realschulabschluss")	39	34
High (i.e. Abitur, advanced technical college certificate, studies)	42	48
Age group		
18-29 years	30	34
30-39 years	21	22
40-49 years	18	17
50-59 years	18	16
60 years and older	13	10

100% missing: no answer.

² Double-opt-in means that the users receive an e-mail which invites them to confirm their membership. Only after successful conclusion of these steps, the user is asked to fill in the master data questionnaire.

The members of the Online Access Panel are recruited by Respondi in various ways. To this end, Respondi approaches people under the brand name "Mingle". Most approaches are made online, but to a lesser extent, also offline. Members are predominantly recruited via surveys in topic-specific portals, in forums and communities. The motive of the recruitment is that people are enabled to express their personal opinion, not the prospect of being financially remunerated. GapFish also recruits the majority of its panel members online, using search engine marketing, e-mail campaigns, social networking or website partnerships, for example.

Table 3: Ways of recruitment at Respondi

	Percentage
Mingle Trend Blog	2.1
Via online advertisements (of that 50% advertised topic-related surveys; 50% direct publicity for the panel, including affiliate marketing)	38
Self-entries	11
Facebook fan page	35
Search engines	10
Recommendations (tell-a-friend advertising)	0.9
Cooperations	2.9
Acquisition by telephone (CATI)	0.1

Respondi states that it operates an effective quality management system. The response behaviour of the Respondi panel members is continuously evaluated via an internal evaluation system. Reasons for deactivating members from the database include, for example, a lack of participation over a period of three months (a prerequisite for Respondi is that members have received at least ten invitations) as well as double registrations and deliberate misrepresentations in several surveys.

Quality management and a moderate frequency of invitations should help to avoid undesirable effects such as panel distortions (panel attrition) or professionalization of members. An average member stays eleven months in the Respondi panel and 24 months in the GapFish panel. Within one year, approximately 40 percent of panel members are removed from the database due to quality control measures and panel attrition (10 percent per month for GapFish).

The average participation rate of the members in a survey with five field days is about 50 percent. The participation rate is calculated by Respondi from the number of started interviews, which includes completed interviews, screen-outs (exclusion when the respondents were chosen by topic), quota-fulls (exclusion because of fulfilled quotas in quota sampling) and interviews that were interrupted. The number of started and interrupted interviews is set in proportion to the number of all forwarded invitations.

The members of the access panel are rewarded by Respondi because of their participation in surveys. On a normal basis, they receive five "mingle points" per minute, which is equivalent to €0.05 (in the 2017 Campaign Panel, the participants received a higher amount, see section 4.5.). As soon as the participants have collected 1500 mingle points, the amount can be received in vouchers or, it can be donated. Starting from 2000 mingle points, these can also be paid out in cash. For panel care, additional raffles are conducted periodically amongst the members. GapFish has a higher standard incentive of €0.10 per minute. Panel members can also collect credits by participating in surveys, which can be paid out or donated, starting from an amount of €10.

3.3. Selection method and quota

The 2017 Campaign Panel currently comprises three survey groups. In the dataset, these survey groups can be identified by the variable "gruppe".

- 1.) The first group consists of a sample comprising 15,802 Respondi panellists who were selected for the first wave in autumn 2016 using quota sampling and for whom a completed interview in that wave is available. In order to meet the quota requirements, some cells were filled with respondents from the partner panel GapFish. A total of 1,257 GapFish respondents participated. Quota sampling of the first subgroup is based on the categories gender, age (in five age groups: 18-29, 30-39, 40-49, 50-59, 60 and older) and education (three categories: low: leaving school without graduation, leaving school after 8 or 9 years of schooling („Haupt"- or „Volksschule"); intermediate: secondary qualification, leaving school after 10 years of schooling ("mittlere Reife"); high: Abitur, advanced technical college certificate). In order to obtain a mostly heterogeneous sample, every one of the 30 possible combinations of the three categories should be represented with a percentage of about 3.33%. Minor deviations from those predefined quotas were accepted as Respondi could not always guarantee for sufficient representation of particular combinations, even with the addition of respondents from the GapFish panel. For instance, it was very challenging to motivate younger males with low educational status to participate (Table 4).
- 2.) The second group consists of 4,608 people who had already participated in at least three waves of the 2013 Campaign Panel. A total of 2,725 of these 4,608 people participated with a completed interview in at least one of the nine waves of the 2017 Campaign Panel. Of these people, 2,277 already completed the first wave; 448 did not join until a later wave and completed at least one interview. No quota was set for the re-invitation of respondents from 2013 to participate in the 2017 Campaign Panel. In contrast to group 1, incomplete interviews from this group in wave 1 were also taken into account. The 2013 Campaign Panel dataset and the interim surveys carried out in 2014 and 2015 can be merged with the 2017 Campaign Panel via the serial number of the respondents (see section 8.2.).
- 3.) The third group consists of 3,960 people who were selected from the Respondi panel between July and August 2017 for a refreshment sample by means of quota sampling. In a separate profile wave, information on stable socio-demographic characteristics and psychological concepts was collected for these people. Of these 3,960 people, only those for whom a completed interview in the profile wave was available were considered.³ From the fifth wave onwards, the refreshment sample participated on a regular basis in the Campaign Panel. When recruiting the refreshment sample, people with the combinations of characteristics for which the highest number of panel dropouts had been recorded up to that point were purposefully invited. In other words, an effort was made to recruit as many low-educated and young respondents as possible (see section 4.6.).

25 respondents of the 2013 Campaign Panel who did not meet the criterion for group 2 (with less than three participations in 2013), were randomly drawn into the quoted sample of new respondents, i.e. group 1, in 2016. Similarly, 10 respondents who had participated in less than three waves in 2013 were drawn into the refreshment sample (group 3) in the summer of 2017. These cases are not included in the "group" variable as "2013 re-contact sample" to reflect the fact that they were not treated as members of the re-contact sample in their first wave of the 2017 Campaign Panel: for these individuals, there was no comparison with the verification data from 2013 at the beginning of the

³ Based on the experiences in wave 1, questions about the Schwartz values were asked at the very end of the questionnaire in the profile wave of the refreshment sample. Moreover, interviews of respondents who interrupted the survey during the Schwartz values (58 observations) were also rated as "complete".

interview (see section 4.7.). However, these cases can be referred to the dataset of the 2017 Campaign Panel via the variable "lfdn13" (see section 8.2.).

Due to an inadequacy in the survey software used, some respondents were able to take part in the survey more than once in the first wave and in the profile wave of the refreshment sample. Only the first interview of each person was considered. In the case mentioned above, numbers do not include such duplicates.

Table 4: Quoting in the 2017 Campaign Panel

Quotation features	Target distribution in percent	Actual distribution* in percent	
		N=15,802 (without participants of the 2013 CP)	N=18,079 (with participants of the 2013 CP)
Gender			
Female	50.0	46.1	46.8
Male	50.0	53.9	53.2
Education (without pupils)			
Low (i.e. no graduation or graduation after 8 or 9 years of schooling) ("Hauptschule", "Volksschule")	33.3	27.5	27.2
Intermediate (i.e. secondary qualification) ("Realschule", "mittlere Reife")	33.4	35.9	36.1
High (i.e. Abitur, advanced technical college certificate, studies)	33.3	36.6	36.7
Age group			
17-29 years	20.0	20.0	18.2
30-39 years	20.0	19.1	18.4
40-49 years	20.0	19.7	19.5
50-59 years	20.0	20.9	21.9
60 years and older	20.0	20.3	22.0

* Only completed interviews in the first wave considered. Of a total of 18,128 interviews, 49 are incomplete (only participants from the 2013 Campaign Panel), which explains the case number of 18,079 in the last column.

From the second wave onwards, the problem of panel attrition, which is typical for panel surveys, arose: Not all persons from the first survey participated in the subsequent waves (see also section 4.8.). As can be seen in Table 5, panel attrition between wave 1 and the following waves influenced the distribution of all three quota characteristics. The share of the low-educated decreased by five percentage points up to wave 9, mainly in favour of the share of the high-educated. The highest systematic non-response rate can be registered in the five age groups: While the share of the youngest age group decreased by more than eight percentage points, the share of the oldest group increased by more than five percentage points.

The extent of the systematic bias appears even greater if one considers the common distribution of quota characteristics (not shown in the table). The 17-29-year-old men and women with a low level of education would be most affected by the panel failure: In this group, only 15 percent of respondents from the first wave took part in the ninth wave. By contrast, other groups were comparatively less affected by panel attrition: the participation rate of women over 60 and men with intermediate

education was approximately two-thirds. The low re-interview rate affected above all those cells that were already relatively low occupied in wave 1.

Table 5: Distribution of gender, education and age, separated by wave, in percent (with respondents of 2013 CP who participated in W1, without incomplete interviews)

Wave		W1	W2	W3	W4	W5	W6	W7	W8	W9
Gender	Male	46.8	48.0	48.9	49.1	49.2	49.9	50.2	49.5	50.0
	Female	53.2	52.0	51.2	50.9	50.8	50.1	49.8	50.5	50.0
Education	Low	27.2	24.5	23.2	23.0	23.3	22.9	22.7	22.6	22.2
	Intermediate	36.1	37.0	37.3	37.2	37.0	36.9	36.6	36.8	37.2
	High	36.7	38.6	39.6	39.8	39.8	40.2	40.8	40.7	40.6
Age group	17-29 years	18.2	14.2	12.3	11.6	11.5	10.7	10.2	10.8	9.9
	30-39 years	18.4	17.0	16.8	16.7	16.5	16.0	16.0	16.2	16.0
	40-49 years	19.5	20.2	20.6	20.5	20.4	20.6	20.8	20.6	20.9
	50-59 years	21.9	23.9	24.6	25.0	25.2	25.6	25.6	25.4	26.0
	60 years >	22.0	24.8	25.8	26.3	26.4	27.2	27.4	26.9	27.2
N		18,079	13,091	11,249	10,760	10,551	10,132	9,346	10,178	9,472

As already mentioned, when recruiting the refreshment sample, targeted persons with the combinations of characteristics were invited, who up to this point had recorded the highest number of panel withdrawals. In other words, attempts were made to recruit as many low-educated and young respondents as possible (see also section 4.6.). As Table 6 shows, these objectives could only be achieved to a limited extent. The dropout pattern for the refreshment sample shows great parallels with that of the rest of the Campaign Panel.

Table 6: Distribution of gender, education and age in the refreshment sample, separated by wave, in percent (without incomplete interviews)

Wave		Profile wave	W5	W6	W7	W8	W9
Gender	Male	53.0	51.3	51.0	49.4	50.3	50.6
	Female	48.0	48.7	49.0	50.6	49.7	49.4
Education	Low	26.2	24.6	23.4	23.5	23.5	22.2
	Intermediate	36.9	37.1	36.4	36.1	36.2	35.5
	High	37.0	38.3	40.2	40.4	40.3	42.3
Age group	17-29 years	24.3	22.8	21.3	20.3	20.2	18.5
	30-39 years	20.1	19.5	19.1	19.8	19.3	19.1
	40-49 years	18.3	18.9	19.1	19.0	19.4	20.4
	50-59 years	19.1	19.7	20.1	20.5	21.1	21.0
	60 years >	18.3	19.0	20.3	20.5	20.0	20.9
N		3,960	3,128	2,900	2,688	2,822	2,335

3.4. About the handling of lacking representativeness in the Campaign Panel

The participants of the Campaign Panel are not representative of the electorate in Germany. An above-average number of young, educated and Internet-savvy people are represented in online access panels. These groups differ in their attitudes and behaviour from the population as a whole. Therefore a generalisation of *proportions*, which are calculated based on Campaign Panel data, is not allowed.

The decisive advantage of panel surveys lies in their potential to analyse causal mechanisms.⁴ It cannot be ruled out that a lack of representativeness also causes distortions to a certain extent, even when using panel causal analyses. However, it can be assumed that these distortions are considerably lower than in causal analyses with representative cross-sectional data whose effect estimates are affected to a considerable extent by the problem of "unobserved heterogeneity" (also called "omitted variable bias") (cf. Green et al. 2001; Brüderl 2010). Panel data and so-called "fixed-effects" panel models can largely eliminate unobserved heterogeneity (cf. Brüderl 2010). For causal analyses, fixed-effect regressions based on non-representative panel data usually provide less distorted effect estimates than ordinary cross-sectional regression methods based on representative data.

⁴ This paragraph is strongly oriented towards the work of Plischke (2014: 198-199).

4 Data collection

4.1. Mode of data collection

Online Panel Survey with a standardized questionnaire

4.2. Data collector

University of Mannheim, Chair of Political Science, Political Psychology

4.3. Data collection software

The programming of the questionnaire and the data collection were carried out using the QuestBack software.

4.4. Testing of the programmed questionnaire

The programmed questionnaire was intensively checked. Any problems or programming errors found in the tests were corrected. Subsequently, the questionnaire was released in the final version of the survey.

4.5. Incentive

In order to achieve the highest possible willingness to be re-questioned, the participants of the Campaign Panel received a little more remuneration than in usual Respondi surveys. In consultation with Respondi, a basic remuneration of 150 Mingle Points (€1.50) was agreed upon, whereas the normal remuneration for a 20-minute interview is 100 Mingle Points. Since the respondents needed considerably longer than 20 minutes on average for the first wave, a one-off payment of €2.00 instead of €1.50 was made in the second wave (see Table 7). Due to the longer duration of the survey, €2.00 was also paid out for participation in the profile wave of the refreshment sample.

Furthermore, the participants were told that frequent participation in the Campaign Panel would enable them to receive additional Mingle Points. The exact amount of the additional incentive was left open until wave 2. From wave 3 onwards, respondents were informed of the amount of the additional incentive. For the payment of an additional incentive, participants who were surveyed for the first time in 2016 or 2013 needed at least five completed surveys. Since the refreshment sample only took part from the fifth wave onwards, they were assured a higher incentive if they took part in at least two waves (see Table 8).

From the seventh wave onwards, respondents were also shown on the welcome page of the survey how many surveys they had already taken. This information was not shown to participants who had no chance of receiving an additional incentive due to their completed surveys. The additional incentive was paid out at the end of the ninth wave.

Table 7: Mingle Points for participation in single waves

Participation in ...	Respondi	GapFish
Wave 1	€1.50	€2.00
Wave 2	€2.00	€2.50
Wave 3	€1.50	€2.00
Wave 4	€1.50	€2.00
Profile wave of the refreshment sample	€2.00	
Wave 5	€1.50	€2.00
Wave 6	€1.50	€2.00
Wave 7	€1.50	€2.00
Wave 8	€1.50	€2.00
Wave 9	€1.50	€2.00

Table 8: Additional Mingle Points

Participation in ... waves	Participation since 2016/2013	Refreshment sample
2		€1.20
3		€1.70
4		€2.20
5	€1.20	€2.50
6	€1.70	€3.50
7	€2.20	
8	€2.50	
9	€3.50	

4.6. Invitations and und reminders

The procedure for invitations and reminders differs between waves in which new participants were recruited (wave 1 and the profile wave for the refreshment sample) and the following waves.

At the beginning of the first wave, the invitations and reminders for the "new" respondents and the respondents from the 2013 Campaign Panel were controlled separately. After a soft launch, about 2,000-3,000 active respondents of the Respondi panel were invited every day. In order to meet predefined quotas, groups with a low probability to participate were contacted first (i.e. younger persons, people with lower educational status); the invitations for the other groups were sent out a few days later. Depending on requirements or if the quota had not yet been met, participants who had not yet started the survey were reminded several days later and were once again invited to participate (see Table 9). From 21 October 2016, participants from the GapFish panel were invited to fill difficult quota cells. The respondents who had already participated in 2013 were contacted in the first field week with about 200 invitations per day. On 13 October 2016, all remaining respondents were invited in one block.

The refreshment sample was only recruited from the Respondi panel. After the end of the fourth wave, certain groups of people showed a higher drop-out rate than others (see Table 5). Especially among young and low-educated men and women, a lower probability of participation was measurable. The quota targets for the refreshment sample were therefore based on the probability of dropouts in the previous waves of the Campaign Panel. After a successful soft launch with 500 completed surveys, about 4,000-5,000 active respondents of the Respondi panel were invited daily. In particular, young and low-educated people were invited more frequently, and reminder e-mails were sent more frequently to remind them of the survey (see Table 10). From the fifth wave onwards, the refreshment sample was interviewed together with the other participants in the election campaign panel. Only the payout mode of the additional incentive differs between the refreshment sample and the previous respondents (cf. Table 8).

Oct 28	Partic. CP 2013 1-8 (R)	5										
Nov 03	Respondi new 21 (E)	493	264	229	217	155	121					493
Nov 03	Respondi new 1-8, 12-16 (R)	4,159										
Nov 03	GapFish new 17-20 (R)	12,790	8,677	4,113	5,173	5,596	2,021					12,790
Nov 03	Partic. CP 2013 1-8 (R)	1,392										
Total invitations		77,457										

E=Invitation, R=Reminder, A=Reminder for interrupters.

Table 10: Recruitment of participants in profile wave of the refreshment sample

Date	Group	Total	Gender		Age					Education		
			male	female	18-29	30-39	40-49	50-59	60+	low	inter- medi- ate	high
Jul 20	Respondi new 1(E)	500	250	250	200	110	75	70	45	80	190	230
Jul 20	Respondi new 2 (E)	5,000	2,500	2,500	2,000	1,050	750	700	500	1,000	1,750	2,250
Jul 21	Respondi new 3 (E)	4,985	2,493	2,492	2,000	1,100	750	697	438	735	1,950	2,300
Jul 22	Respondi new 4 (E)	4,972	1,344	3,628	2,000	1,050	799	688	435	673	1,950	2,349
Jul 23	Respondi new 5 (E)	4,838	1,342	3,496	2,000	1,050	736	687	365	708	1,950	2,180
Jul 24	Respondi new 6 (E)	4,745	1,293	3,452	2,000	1,100	736	630	279	709	1,900	2,136
Jul 24	Respondi new 1-3 (R)	1,540									1,540	
Jul 25	Respondi new 7 (E)	4,403	1,172	3,231	1,881	1,093	736	440	253	685	1,772	1,946
Jul 25	Respondi new 4 (R)	577									577	
Jul 26	Respondi new 8 (E)	3,985	1,032	2,953	1,751	971	537	459	267	561	1,614	1,810
Jul 26	Respondi new 9 (E)	1,115	602	513	32	312	210	358	203	1,115		
Jul 26	Respondi new 10 (E)	27	27		27						27	
Jul 26	Respondi new 11 (E)	328	328		306				22		328	
Jul 26	Respondi new 5 (R)	590									590	
Jul 27	Respondi new 6 (R)	615									615	
Jul 27	Respondi new 12 (E)	3,162	537	2,625	7	746	1,046	1,246	117	171	2,991	
Jul 28	Respondi new 1-7 (R)	3,368									3,368	
Jul 28	Respondi new 1-7 (A)	778										
Jul 29	Respondi new 8-11 (R)	4,345										
Jul 29	Respondi new 8-11 (A)	120										
Jul 30	Respondi new 12 (R)	2,523										
Jul 30	Respondi new 12 (A)	53										
Aug 04	Respondi new 1-12 (R)	27,259										
Aug 04	Respondi new 1-12 (A)	964										
Total invitations		38,060										

E=Invitation, R=Reminder, A=Reminder for interrupters.

From the second wave onwards, only those persons were re-invited who had fully completed the first survey (wave 1 drop-outs were not invited again). In Table 9 and 10, only complete interviews in wave 1 and the profile wave of the refreshment sample were considered. All participants were invited on the same day but evenly distributed throughout the day to avoid server overload. In addition to an e-mail reminder for people who did not fill out the questionnaire yet, the people who began the questionnaire

but did not complete it also received a reminder (see Table 11). Persons who made false statements during the verification process or tried to participate with a smartphone were reset and reminded at most once during each field time (see section 4.7.). In the second wave, respondents who were screened out for using a smartphone were reset up to three points in time before each reminder. However, the resetting process was very time-consuming and error-prone, so that from the third wave onwards all screened persons were reset only once, immediately before the final reminder.

Whilst formulating invitation and reminding e-mails, importance was attached to highlight the special characteristics of the study. The survey was presented to the panellists as "GLES Study", and in every invitation, the subject line included the study title in order to ensure recognition (i.e. subject line: "mingle - GLES Study part 2"). In agreement with Respondi and GapFish, the standard text in e-mails was changed, and the GLES logo was included. The layout and the exact wording of the invitations and the homepage of every wave can be reviewed in the documents "Screen Views" which can be downloaded from GESIS Data Catalogue.

Table 11: Invitations and reminders

	W2	W3	W4	Wa1	W5	W6	W7	W8	W9
Invitations	Feb 16	May 11*	Jul 06	Jul 20	Aug 17	Sep 04	Sep 18	Sep 27	Mar 15
Reminder & Reminder for interrupters	Feb 23	May 14	Jul 09	Jul 24	Aug 19	Sep 06	Sep 20	Sep 29	Mar 17
Reminder & Reminder for interrupters	Feb 26**	May 18	Jul 12	Jul 29	Aug 22	Sep 08		Oct 02	Mar 19
Final reminder & Reminder for interrupters	Mar 01	May 21	Jul 15	Aug 04	Aug 25	Sep 11	Sep 22	Oct 04	Mar 22
Resetting the screen-outs	Feb 23 Feb 26 Mar 01	May 19	Jul 12		Aug 25	Sep 11	Sep 22	Oct 04	Mar 22
End of fieldwork	Mar 03	May 23	Jul 17	Aug 09	Aug 28	Sep 13	Sep 23	Oct 09	Mar 26

*Respondents from North Rhine-Westphalia were only invited on 15 May and received only two reminders due to the shorter field time (see section 4.8.). **Respondents who did not get access to the survey due to a programming error in their verification data received an additional reminder on 28 February (see section 4.7.).

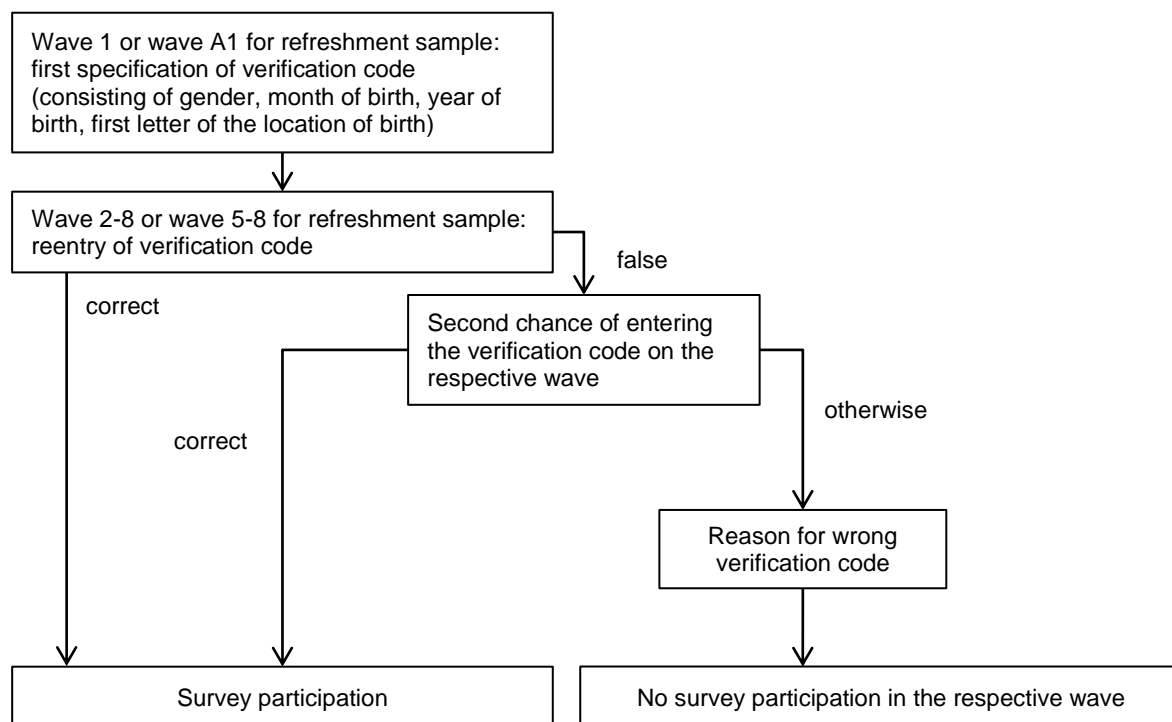
4.7. Participant verification

When carrying out online panel surveys, it is important to make sure that only the target groups take part in the survey, as web surveys allow a high degree of anonymity. For instance, household members who have access to the e-mail account of the target person can participate instead. In order to prevent incorrect participation, the invitational e-mail and the homepage included a request to abstain from participating in the survey if one did not participate in the survey before.

Plus, participant verification was implemented before starting the questionnaire of every wave. This included requesting four stable features of every participant: gender, the year of birth, the month of birth and the first letter of the place of birth. From the second wave on, the statements given in the second wave were checked to be correspondent to the information given in the first wave (see Figure 3). The participants were only transmitted without detours to the questionnaire when the statements of wave 1 and wave 2 were in full compliance. If at least one feature proved to be false, the participants were given the opportunity to enter their four features a second time. When the data did not correspond to the data given before again, the participants were told that they could not take part in

the survey. They were informed that the problem was to be solved and they would be re-invited later. Before the participants were rejected, they were able to give a reason why the verification code did not correspond to the statements given in the first wave (see "Reason for wrong verification code" in Figure 3).

Figure 3: Verification process



The target persons were re-invited to the survey in the course of reminder e-mails for interrupters and could try again to enter their correct verification code. This possibility was only offered once to each respondent with the wrong verification information. Of those respondents who had already participated in the 2013 Campaign Panel, information on the four characteristics was already available. For this reason, the verification information of this group of participants was compared with their data in the 2013 Campaign Panel from the very first wave onwards.

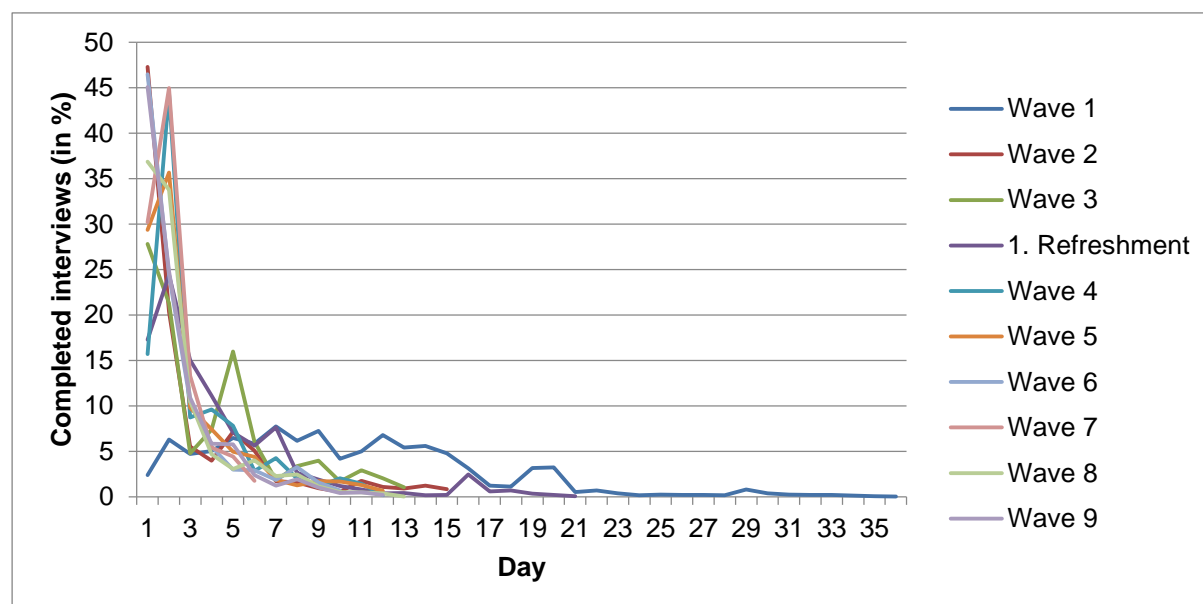
For 107 people who had already participated in the 2013 Campaign Panel, wave 2 had errors in the data used to cross-check in the verification process. These respondents had made errors in their first attempt at verification in the first wave, corrected them, and got access to the first survey in the second attempt. In the data preparation process of wave 1, however, its first input was accidentally saved as "correct" and stored as verification data in the online programming for wave 2. The error was noticed and corrected on the 12th field day. At this point, 56 of these respondents had reached the second survey despite errors in their verification characteristics because they had repeated the error from the first wave. A separate reminder e-mail was sent to 39 respondents who had unsuccessfully attempted to participate in the survey with their correct verification information (see wave 2 screenshots). The 56 persons who came to the survey despite an error in their verification characteristics can be identified by the variable "kp2_info1".

4.8. Distribution of participation over field time

In the first wave as well as in the profile wave of the refreshment sample, the field time of five and three weeks respectively was significantly longer than in the other wave, as not all invitations were sent on the same day. This invitation procedure intended to guarantee that the quota target was reached. The distribution of the completed interviews is spread on the whole length of the field time during wave 1 and the profile wave of the refreshment sample. From the second wave on, all participants were invited on the same day. This entailed a strong concentration of accesses to the survey on the first two days of the field time. On the third and fourth field day, respectively, in most waves, a light increase of participation can be noticed, which is due to reminder e-mails.

The distribution of the completed surveys differs slightly in the third wave, as the respondents from North Rhine-Westphalia were invited only on the fifth regular day of the field time. This ensured that all respondents from North Rhine-Westphalia could only take part in the survey after the state elections on 14 May 2017. Accordingly, these respondents received reminder e-mails in a slightly offset rhythm.

Figure 4: Distribution of the finished interviews during the field time (in percent of the number of all finished interviews per wave)



4.9. Participation rates

Systematic unit non-response is a source of error in surveys which can substantially reduce data quality.⁵ When the collected variables are correlated with the panellists' likelihood to participate, a distortion of the survey results ("non-response bias") is the consequence. The specification of response rates attempts to provide an estimation of how strongly a survey is affected by possible systematic non-participation. The information about standardized response rates and related indicators for the panellists' participation behaviour ensures the comparability of these patterns beyond other studies. For the GLES Campaign Panel, indicators of the participation behaviour are stated according to the standards of the American Association for Public Opinion Research, which were

⁵ This section is strongly oriented towards the work about participation rates in the technical report of the Campaign Panel 2009 (see Steinbrecher et al. 2013: 25f., as well as Blumenstiel/Gummer 2012).

published in 2011 and according to recommendations on the calculation of participation rates by Callegaro and DiSogra (2008).

The interpretation of response rates in online surveys is in many cases challenging as there are usually no random processes of sampling. When questioning participants from self-recruiting online access panels, like Respondi's or GapFish's online access panel, it is not reasonable to refer to response rates (see The American Association for Public Opinion Research 2011). Therefore, AAPOR recommends stating a 'participation rate' when interviewing participants from self-recruiting online panels. The participation rate is calculated as a share of the utilizable interviews of the absolute number of all forwarded personal invitations to take part in the survey.

To be able to calculate the participation rates, one has to distinguish sufficient and insufficient interviews. An interview is classified as sufficient when the questionnaire was fully completed by an eligible participant. This includes complete (I) and partial (P) interviews which occur when participants do not answer all questions, but still complete the questionnaire till the end. By the use of a screening question at the beginning of the first wave interview, only those persons were registered who were eligible to vote in Germany and met certain features (gender, age, education). The other people were rejected when the fixed quota was fulfilled (NE) and are not included in the calculation of the participation rates.

The insufficient interviews are defined by interrupted interviews or by no interview at all. Participants who dropped out of the survey before the interview was finished and did not continue at a later point in time are included in the category of break off (RBreakoff). From the second wave on, there were two different kinds of drop-outs: Participants can exit the survey prematurely before or after entering the verification code. While eligibility is guaranteed after successful verification, it is uncertain when the participant cancels the interview before verification. Eligibility is yet relevant for the calculation of a break off rate in order to make it as comparable as possible. Therefore, these two types should not be used undifferentiated. Through the panel design, the suitability has already been identified in the first wave. Furthermore, a not eligible person can only be included if an unknown third person has access to the Respondi account as well. Since participants who discontinue the survey before the verification process cannot be classified beyond doubt as not eligible, it was decided as a consequence to calculate two break-off rates; one rate which only includes the discontinuation of participants who are eligible (R1Breakoff) and a second rate which additionally comprises the discontinuations of participants with unknown status of eligibility (R2Breakoff).

About Respondi members who did not respond to the first wave invitation and thereby gave no interview, one cannot state clearly whether they are eligible or not. As a consequence, these participants must be defined as vaguely eligible (UH). Therefore, they are only included in the calculation of the participation rates of the first waves, as eligible participants for the following waves had to fully complete the first wave interview. If those eligible Respondi members did not react to invitations for the following waves, they are counted as no contact (NC).

The categories are finally supplemented by the others with unclear eligibility. This includes the participants who were excluded from the survey due to an incorrectly entered verification code. According to usual AAPOR recommendations, these cases would be marked as screened out and would not be included in the calculation of the participation rate as they would be considered as non-eligible. As the fundamental eligibility has already been determined in the first wave and therefore the addressee of the invitation is in any case eligible since wave 2, a verification failure must be due to an unknown third party who used the Respondi account as well or without the approval of the owner. Since participants who entered a wrong verification code were invited several times again in order to enter a correct code (see section 4.7.), it can be assumed that the addressed participant did not come into contact with the survey when the verification failed. If those participants now had been screened out, one would assume the participants had each changed in a way which did not make them eligible anymore. This is not the case. Every eligible participant who did not quit the Respondi panel received a fresh invitation for every wave. Consequently, a wrong verification code does not mean the wrong

person was addressed, but that the eligible Respondi member simply did not react to the invitation and would be assigned to the category "no contact". This distinction is highly important for the comparability of the participation rate. A classification as "non-eligible" would lead to variations of the reference value "all forwarded invitations" although the number of eligible peoples' invitations would remain the same. Therefore, interviews which did not take place because of the verification process should also be included in the calculation. As it is not clear why the verification failed, those interviews are referred to like others with unclear eligibility (UO). The calculation of the participation rates is as follows:

$$\text{Participation Rate (PR)} = \frac{I + P}{(I + P) + (R1_{\text{Breakoff}} + R2_{\text{Breakoff}} + NC) + (UH + UO)}$$

Another informative indicator is the share of interrupted interviews in the online survey. For this, the breakoff rate (Callegaro/DiSogra 2008) is calculated. As indicated above, one differentiates between a break-off before and after the verification; therefore, two break-off rates are calculated. R1Breakoff includes all break-offs which took place after the verification process. R2Breakoff, on the other hand, covers all drop-outs who quit the survey before the verification. I and P form the number of sufficient interviews. The breakoff rates, in other words, reflect the share of the interview break-offs of all begun interviews.

$$\begin{aligned} \text{Breakoff Rate 1 (BR1)} &= \frac{R1_{\text{Breakoff}}}{(I + P) + (R1_{\text{Breakoff}})} \\ \text{Breakoff Rate 2 (BR2)} &= \frac{(R1_{\text{Breakoff}} + R2_{\text{Breakoff}})}{(I + P) + (R1_{\text{Breakoff}} + R2_{\text{Breakoff}})} \end{aligned}$$

The gross sample for the first wave of the Campaign Panel was drawn from the Respondi and GapFish panel respectively. The chosen panellists were invited to participate in the online access panels. The gross sample for the following wave of the Campaign Panel consists of the panellists who completed the first panel wave as well as all respondents from the 2013 Campaign Panel who took the first wave (with interrupters after a correct verification). They thus establish the basis of calculation of the participation rate and the Breakoff 1 and 2 rates. Table 12 provides an overview of the rates per wave. Table 13 shows the rates for the refreshment sample.

Table 12: Participation rate statistics of the Campaign Panel*

Code	Description	W1	W2*	W3*	W4*	W5*	W6*	W7*	W8*	W9*
I + P	Complete and partially complete interviews	18,079	13,114	11,271	10,782	10,570	10,145	9,359	10,192	9,491
R1_{Breakoff}	Break-off after verification process	1,547	246	175	210	176	160	141	148	163
R2_{Breakoff}	Break-off before verification process	/	96	93	168	158	71	64	62	96
NC	No Contact	/	4,322	6,327	6,755	7,074	7,630	8,481	7,628	8,221
UH	Ambiguous eligibility	52,463	/	/	/	/	/	/	/	/
UO	Others with ambiguous eligibility	56	350	262	213	150	122	83	98	157
NE	Met quota or screened out in wave 1	5,312	/	/	/	/	/	/	/	/
Number of interview invitations		77,457	18,128	18,128	18,128	18,128	18,128	18,128	18,128	18,128
PR	Participation Rate in %	25.1	72.3	62.2	59.5	58.3	56.0	51.6	56.2	52.4
BR1	Breakoff Rate 1 in %	7.9	1.8	1.5	1.9	1.6	1.6	1.5	1.4	1.7
BR2	Breakoff Rate 2 in %	/	2.5	2.3	3.4	3.1	2.2	2.1	2.0	2.7

* Respondents from the 2013 Campaign Panel who did not participate in the first wave of the 2017 Campaign Panel as well as the refreshment sample are not considered.

Table 13: Participation rate statistics of the refreshment sample

Code	Description	WA1	W5	W6	W7	W8	W9
I + P	Complete and partially complete interviews	3,960	3,128	2,900	2,688	2,822	2,335
R1 _{Breakoff}	Break-off after verification process	104	66	49	51	44	42
R2 _{Breakoff}	Break-off before verification process	/	24	34	16	23	28
NC	No Contact	/	679	923	1,172	1,047	1,498
UH	Ambiguous eligibility	27,186	/	/	/	/	/
UO	Others with ambiguous eligibility	31	63	54	33	24	57
NE	Met quota or screened out in wave A1	6,779	/	/	/	/	/
Number of interview invitations		38,060	3,960	3,960	3,960	3,960	3,960
PR	Participation Rate in %	10.4	79.0	73.2	67.9	71.3	57.2
BR1	Breakoff Rate 1 in %	2.6	2.1	1.7	1.9	1.5	1.8
BR2	Breakoff Rate 2 in %	/	2.8	2.8	2.4	2.3	2.9

For the definition of complete interviews in the profile wave of the first refreshment sample (WA1), see footnote 3.

The variable "kpX_compl" in the dataset includes the participation status of the panellists according to AAPOR standards. With the help of those variables, the calculation of nonresponse rates is to be simplified in research works. Table 14 provides an overview of the meaning of the coding.

Table 14: Manifestation of the variables kpX_compl

Code	Value label	Explanation
<u>1 Eligible persons, completed questionnaire</u>		
1.10	Fully completed	People who filled in the complete questionnaire. If there was an "I don't know"- category and it was used, this is also a complete answer. If, however, a respondent clicked through at least once without providing an answer, the case is described as incomplete.
1.20	Interrupted, partly completed	The meaning of this category differs in every wave. In the first wave, it includes those persons who did not provide any answer at least once but finished the interview so that they took part in the study further on (first wave drop-outs were not invited again and are not included in the published dataset). In the waves 2-9, the category includes those people who were identified correctly and answered at least one question after the verification process, but then either a) quit the survey without resuming at a later time or b) clicked through at least once without providing a valid answer.
<u>2 Eligible persons, no interview</u>		
2.11	Implicit refusal of participants with known eligibility	People who had logged on with their correct identification code (query to identify the target person), but immediately quit the survey afterwards and did not answer a single question (only in waves 2 to 9).
2.20	No contact	People who did not respond to the invitation (only in waves 2 to 9).
2.30	No contact, panel membership cancelled	People who did not respond to the invitation because they are no longer part of the Respondi panel (only in waves 2 to 9).
<u>3 Unknown eligibility, no interview</u>		
3.93	Implicit refusal of participants with ambiguous eligibility	People who had quit the survey even before they entered the verification code (only in waves 2 to 9).
3.94	Break-off due to a wrong verification code	People who entered a wrong verification code and were therefore not forwarded to the survey

Explanation: "X" represents the respective wave.

As can be seen from Table 15, the proportion of respondents who participated in all surveys is lower among respondents first surveyed in October 2016 than among the respondents of the refreshment sample. However, it should be noted that the refreshment sample had to complete fewer waves over a shorter period than the respondents recruited in 2016. Nearly 60% of the respondents recruited in 2013 participated in at least one wave from 2016-2017, slightly less than 30% in all nine waves. The variable "teilnahme" provides detailed information in the dataset about the individual development of participation in all waves.

Table 15: Number of completed surveys

Participation in ... waves	Respondent since 2016		Respondent since 2013		Refreshment sample	
	absolute	%	absolute	%	absolute	%
No (complete) participation			1,883	40.9		
1 participation	3,629	23.0	202	4.4	524	13.2
2 participations	1,590	10.1	183	4.0	326	8.2
3 participations	893	5.7	141	3.1	263	6.6
4 participations	702	4.4	107	2.3	274	6.9
5 participations	584	3.7	106	2.3	666	16.8
6 participations	555	3.5	111	2.4	1,907	48.2
7 participations	775	4.9	181	3.9		
8 participations	1,582	10.0	408	8.9		
9 participations	5,492	34.8	1,286	27.9		
Total	15,802	100.0	4,608	100.0	3,960	100.0

Deviation from 100 percent due to rounding-off of the numbers. Only completed interviews considered. For the definition of complete interviews in the profile wave of the 1. refreshment sample (WA1), see footnote 3.

4.10. Interview duration

With a median time of 25:53 minutes, the first wave took much longer than the following waves (see Table 16). As can be seen from the data on the minimum and maximum values, the completion times in every wave varied considerably. Very long completion times, which are probably caused by short interruptions of the survey, are mostly uncritical (longer interruptions that were followed by an automatic log out are not included in the statistic). Very fast interviews are far more problematic as they indicate that the respondents solely clicked through the survey without having read the questions carefully. The handling of fast responding times is described in section 7.1.

Table 16: Interview length

	N	Minimal duration of the survey	Maximum duration of the survey	Arithmetic mean	Median
Wave 1	17,466	00:03:14	06:27:07	00:29:11	00:25:53
Wave 2	12,406	00:02:29	04:29:43	00:20:23	00:17:37
Wave 3	10,753	00:02:11	02:42:54	00:18:06	00:15:28
Wave 4	10,261	00:02:39	02:40:08	00:17:35	00:14:55
1. Refreshment sample	3,960	00:04:50	02:17:52	00:24:32	00:21:34
Wave 5	10,046	00:03:19	03:07:35	00:20:00	00:17:05
Wave 6	9,714	00:02:30	03:34:43	00:21:18	00:18:24
Wave 7	9,009	00:02:17	02:43:10	00:21:05	00:18:01
Wave 8	9,809	00:02:15	03:43:19	00:20:05	00:17:11
Wave 9	9,120	00:01:55	03:07:02	00:19:53	00:17:08

The duration is only shown for those participants who didn't interrupt the survey. According to the survey software, one person spent 13:12:12 and a second person spent even 1:05:32:08 (tt:hh:mm:ss) on the first wave questionnaire without interruption. These two persons were not included in the presentation of the completion times in wave 1. Respondents who participated in the 2013 Campaign Panel but not in the first wave were not considered. Similarly, respondents of the refreshment sample are not taken into account in the calculation for wave 5-9. Presentation: hh:mm:ss.

5 Variables in the dataset

The dataset of the Campaign Panel includes different types of variables:

- *GESIS archive variables* describe the dataset and its creation. They include variables which are required for the archiving [storage; filing] and distribution of the study, i.e. the study number, the version and the field time of data collection.
- *Administrative variables*: Variables concerning the participation of the interviewees in the Campaign Panel (i.e. AAPOR codes or dummies which provide information about the participation in specific waves).
- *Contextual variables* provide information about regional contexts of the participants.
- *Data quality variables* are indicators which can potentially be used to evaluate the participants' answer quality (see chapter 7).
- *Weighting variables*
- *Attitudinal and behavioural variables*: The actually interesting respondents' characteristics which were collected on the basis of a questionnaire.
- *Paradata* document the technical procedure of the actual interview process as well as information about activities in surveys of Respondi members before and during the field time. This includes, for instance, the number of received invitations to take part in a survey, the number of participations, as well as the equipping of the participants (i.e. browser version, JavaScript version).

Table 17 gives a review of the total number of the variables in the individual waves. The following subchapters provide a description of the different variable types.

Table 17: Number of variables in total and detailed for waves

	W1	W2	W3	W4	Wa1	W5	W6	W7	W8	W9	All waves	Total
GESIS archive variables	0	0	0	0	0	0	0	0	0	0	7	7
Administrative variables	6	6	6	6	6	6	6	6	6	6	3	63
Contextual variables	/	/	/	/	/	/	/	/	/	/	6	6
Data quality variables	7	3	2	2	2	4	2	2	4	4	1	33
Weighting variables	/	/	/	/	/	/	/	/	/	/	10	10
Attitudinal and behavioural variables	241	155	197	243	158	284	296	320	296	197	54	2,441
Paradata	8	7	7	7	8	7	7	7	7	7	16	88
Time variables	261	142	172	150	174	188	224	224	222	156	/	1,913
Total	523	313	384	408	348	489	535	559	535	370	97	4,561

Wa1 denotes the profile wave of the first refreshment sample. The variables of this wave have the prefix "kpa1" in the dataset. If another profile wave is added in the future for a new refreshment sample, the corresponding variables will have the prefix "kpa2".

5.1. GESIS archive variables

Table 18: Archive variables of GESIS

Variable	Explanation
study	provides the (ZA) study number of the dataset in four-digit format and the study number and title as value label
version	Dataset version, starting with 1.0.0
doi	Digital Object Identifier
field_start	First day of field time of the study
field_end	Last day of field time of the study
glescomp	Allocation to one of the eleven components of GLES. Here: component 3 for the Campaign Panel
sample	Allocation of respondents to survey groups (1=main sample 2016; 2= re-contact sample 2013 ; 3=refreshment sample 2017)

Explanation: "X" represents the respective wave.

5.2. Paradata

The published dataset contains two different kinds of paradata. The first group of variables provides information about the technical process of the interview. These are saved automatically by Global Park's survey software EFS (see Table 19). The second group includes information about the panellists' survey activities and is provided by Respondi (see Table 20).

Table 19: Paradata of the technical interview process

Variable	Explanation
kpX_browser	States the name of the browser (User Agent) in the same way in which it was transmitted to the EFS survey server. Examples can, i.e. be found at http://de.wikipedia.org/wiki/User_Agent .
kpX_javascript	Contains the result of the check for JavaScript which is optionally carried out at the start of the study (0 = JavaScript is deactivated; otherwise = version, e.g. 10 for JavaScript 1.0)
kpX_flash	Contains the result of the check for the Adobe Flash-Plugin, which is optionally carried out at the start of the survey and is especially important for Flash questions (0 = nonexistent Flash-Plugin; otherwise= version, e.g. 800 for Flash-Plugin 8.0).
kpX_datetime	Date and time of the start of the survey, i.e. the access on the first page of the questionnaire. The information about time in the dataset corresponds to Greenwich Mean Time (GMT).
kpX_date_of_last_access	Date and time of the last access to the survey. The information about time in the dataset corresponds to Greenwich Mean Time (GMT).
lfdn, lfdn09, lfdn13	Every survey participant receives a consecutive number.
kp1_quota & kpa1_quota	Contains the ID of the assigned quota (i.e., the quota which was chosen by (assignment mode). This is used to screen out unsuitable participants.
kpX_lastpage	Indicates the last page submitted by the participant, i.e. if the participant dropped out of the survey.
kpX_duration	The duration of processing, i.e. the time which passes between the first and the last access of the participant to the questionnaire. When a participant interrupts filling out the questionnaire and continues at a later point in time (disposition codes 23, 32), kpX_duration is assigned the value -99, as no reasonable calculation is possible.

Explanation: "X" represents the respective wave.

The paradata, which was requested of Respondi, provide information about the panellists' entry in the panel, the way of recruitment, as well as the survey participation behaviour of the panellists: a) during the last 12 months and b) during the last three months. The cut-off date for the data retrieval of the information concerning temporal issues was November 16, 2016.

Table 20: Paradata provided by Respondi

Variable	Explanation
p_enter_date	Date of joining the Respondi panel
p_numinv2 & p_numinv2_2	Number of survey participation invitations (in the last 12 months; key date respondents recruited in 2016: 11/16/2016 (p_numinv2), key date refreshment sample A1: 08/10/2017 (p_numinv2_2))
p_numcpl2 & p_numcpl2_2	Number of survey participations (in the last 12 months; key date respondents recruited in 2016: 11/16/2016 (p_numcpl2), key date refreshment sample A1: 08/10/2017 (p_numcpl2_2))
p_numstr2 & p_numstr2_2	Number of commenced surveys (in the last 12 months; key date respondents recruited in 2016: 11/16/2016 (p_numstr2), key date refreshment sample A1: 08/10/2017 (p_numstr2_2))
p_numinv3 & p_numinv3_2	Number of survey participation invitations (in the last 3 months; key date respondents recruited in 2016: 11/16/2016 (p_numinv3), key date refreshment sample A1: 08/10/2017 (p_numinv3_2))
p_numcpl3 & p_numcpl3_2	Number of survey participations (in the last 3 months; key date respondents recruited in 2016: 11/16/2016 (p_numcpl3), key date refreshment sample A1: 08/10/2017 (p_numcpl3_2))
p_numstr3 & p_numstr3_2	Number of commenced surveys (in the last 3 months; key date respondents recruited in 2016: 11/16/2016 (p_numstr3), key date refreshment sample A1: 08/10/2017 (p_numstr3_2))

5.3. Administrative variables

Table 21: Administrative variables

Variable	Explanation
partner	Respondent recruited from the GapFish Access-Panel
kpX_compl	Participation status of the panellists according to AAPOR standards
kpX_dispcode	Disposition code of the participants which shows their status in the field, that is e.g. whether the participant has already begun or finished the survey
kpX_dropout	Completed interviews (0=complete; 1=not complete)
kpX_interrupt	Interruption of the interview (0 = no interruption; 1 = with interruption)
wXa	Information whether the respondent has begun the particular wave X or not (0 = no participation; 1 = interrupted / finished interview (after verification))
wXb	Information whether the respondent has finished the particular section X of the survey or not (0 = interrupted / no participation; 1 = finished interview)
n_teilnahmen	Number of interviews which were finished by the Campaign Panel participant
teilnahme	Seven-digit numeric code which indicates in which waves of the Campaign Panel a participant has taken part (0 = no participation; 1 = participation).

Explanation: "X" represents the respective wave.

5.4. Contextual variables: Allocation of constituencies

The only contextual variable included in the Campaign Panel is the constituency of the participants. Respondents were assigned to a constituency on the basis of the postal code of their main residence, which they stated in the survey. Based on the correspondence table of the Federal Returning Officer, the Stata-ado "plztowknr"⁶ assigns postcodes to electoral districts (2017 area status).

When using the information on constituencies, one has to consider that one postal code does not always allow a distinct mapping to one constituency. Especially in urban areas, a postal code can be matched with up to five constituencies in extreme cases. If an unambiguous allocation was possible - 16,962 cases (93.6%) in the main sample and 3,727 cases (94.1%) in the refreshment sample – the constituency was saved in the variable "elecdist". If several allocations were possible – 855 cases (4.7%) in the main sample and 203 cases (5.1%) in the refreshment sample – the up to four possible constituencies are included in the variables "elecdist1" to "elecdist4". The remaining 311 participants (1.7%) in the main sample, - 30 participants (0.8%) in the refreshment sample - did not provide a valid postal code or refused to provide it.

In addition to that, a comparison between the mentioned federal state and the provided postal code was made. In 482 cases, the postal code did not correspond to the entered state. In the case of the refreshment sample, this applies to 12 respondents.

For the ninth wave, the federal state was surveyed for the second time. 150 respondents in the main sample and 26 respondents from the refreshment sample respectively indicated a different federal state than in the first survey. For 12 respondents (four respondents in the refreshment sample), the federal state did not match the constituency, and the federal state was different from the one given in the first survey. These were marked with the variable "kpX_info".

5.5. Data quality variables

The published dataset contains variables which are used to enable the users to assess the quality of the data. Detailed explanations of the quality indicators can be found in chapter 7.

Table 22: Data quality variables

Variable	Explanation	W1	W2	W3	W4	W5	WA1	W6	W7	W8	W9
kpX_speederindex	Considers the response time of one participant in relation to those of all other participants. See also see section 7.1.	X	X	X	X	X	X	X	X	X	X
kpX_4240	Evaluation of the survey	X	X	X	X	X	X	X	X	X	X
kpX_4250	Participation in surveys, number of membership in online panels	X									
kpX_4260	Participation in surveys, number of survey participations in the last month	X									

⁶ Glinitzer, Konstantin, Tobias Gummer, Malte Kaukal, and Joss Roßmann (2018): plztowknr: Stata module to translate German zip codes into electoral districts (Version: 1.0) [Computer Software]. Chestnut Hill, MA: Boston College.

kpX_4270s	Comment field in which the respondents were able to state their remarks on the survey	X	X	X	X	X	X	X	X	X	X
kpX_040q, kpX_050q	Bogus Items	X				X				X	X

Explanation: "X" represents the respective wave.

5.6. Weighting variables

Cross-section weights

With the help of cross-section weights, the distribution of certain variables in the dataset can be adapted to known distributions of the population. The latter is based on the assumption that there is at least a weak correlation between adjustment variables and the attitudinal and behavioural variables. Caution is advised when using weighting variables: For each specific analysis, the researcher must decide for himself or herself whether one of the weighting variables supplied is useful for the study or not.

To calculate the cross-section weights for GLES datasets, socio- as well as regional structural features were chosen. When calculating the cross-sectional weights for GLES, the iterative proportional fitting method (IPF) was used (Deming/Stephan 1940). When the IPF weighting is used, the actual distribution of the individual cells is gradually adjusted to the respective target distribution of the weighting variables. The process of adjustment is finished when the difference between the weighted marginal distribution of all factors and the target distribution undercuts the abort criterion of 0.05⁷. In order to prevent huge weighting factors, the factors are trimmed to the quadruple mean value of the weighting variable (thus five) after every step of the iteration process⁸. However, such a trim was only necessary for the weights created for participation in all waves based on the (N)Onliner-Atlas 2016 (wei4_on) as there were no great derivations of set point and actual value.

Table 23: Overview of weights

Weight	Variable
Social- and regional structural weight (adjustment to MZ2016) (2016 main sample)	wei_mz
Social- and regional structural weight (adjustment to Onliner) (2016 main sample)	wei_on
Social- and regional structural weight (adjustment to MZ2016) (2016 main sample and 2013 re-contact sample)	wei2_mz
Social- and regional structural weight (adjustment to Onliner) (main sample 2016 and 2013 re-contact sample)	wei2_on
Social- and regional structural weight (adjustment to MZ2016) (refreshment sample 2017)	wei3_mz

⁷ The cross-section weights were calculated with Stata whereas one reverted to the ado "ipfweight" by Michael Bergmann.

⁸ This procedure is also applied in the calculation of the weights of ANES (American National Election Study; see: DeBell et al. 2009).

Social- and regional structural weight (adjustment to Onliner) (refreshment sample 2017)	wei3_on
Social- and regional structural weight (adjustment to MZ2016) participation in all waves	wei4_mz
Social- and regional structural weight (adjustment to Onliner) participation in all waves	wei4_on
Social- and regional structural weight (adjustment to MZ2016) participation in wave 5 to 9	wei5_mz
Social- and regional structural weight (adjustment to Onliner) participation in wave 5 to 9	wei5_on

While the cross-section weights of the Campaign Panel were calculated, based on the actual distribution of the first panel wave, it was attuned to the marginal distributions of the (N) Onliner Atlas 2016, as well as to the distributions of the Mikrozensus 2016. Only persons entitled to vote in private households situated in their main residence were included in the calculation of the target distribution of Mikrozensus 2016.⁹

It was attuned to the socio-demographic and the regional structural features: Gender, age, education and the old West German states or the states of former East Germany (Berlin included). Age was divided into four groups: "18 to under 30 years", "30 to under 45 years", "45 to under 60 years" and "60 years and older". The characteristic education was divided into three groups:

- low education: School completed without graduation, Elementary School graduation, lowest formal qualification of Germany's tripartite school system, after 8 or 9 years of schooling ("Hauptschulabschluss", "Volksschulabschluss"), still attending school¹⁰
- intermediate education: Intermediary secondary qualification, after 10 years of schooling ("Mittlere Reife", "Realschulabschluss", or "Polytechnische Oberschule mit Abschluss 10. Klasse")
- high education: Certificate fulfilling entrance requirements to study at a polytechnic college ("Fachhochschulreife (Abschluss einer Fachoberschule etc.)") or higher qualification which entitles holders to study at a university ("Abitur" or "Erweiterte Oberschule mit Abschluss 12. Klasse" ("Hochschulreife"))

⁹ The values reported by Mikrozensus represent absolute numbers of people after bound extrapolation.

¹⁰ For respondents who were recruited in 2013 and who did not participate in wave 1, the weights in wave 5 to 8 (wei5_mz and wei5_on) were constructed using information on education from the previous 2013 Campaign Panel.

Table 24: Actual and target distributions of the variables which were used to calculate the weights

Characteristic	Actual distribution (in percent)			Target distribution (in percent)	
	2016 main sample N=15,802	2016 main sample and 2013 re-contact sample N=18,128 ¹¹	2017 refreshment sample N=3,960	Mikro- zensus 2016	(N)Onliner Atlas 2016
Gender					
male	46.1	46.7	48.3	48.6	51.6
female	53.9	53.3	51.7	51.4	48.4
Age group					
18 up to 30 years	20.0	18.2	21.4	15.9	21.0
30 up to 45 years	27.9	27.1	28.6	20.7	25.2
45 up to 60 years	31.9	32.8	30.0	29.0	33.8
60 years and older	20.3	22.0	20.0	34.4	20.0
Education					
low	28.6	28.1	26.7	37.6	35.5
intermediate	35.4	35.6	36.6	30.4	32.1
high	36.1	36.3	36.7	32.0	32.4
Federal state					
Western German states	76.4	76.5	77.7	79.1	80.7
Newly-formed German states (incl. Berlin)	23.6	23.6	22.3	20.9	19.3

The calculation of the weighting variables for the different sample groups was stopped after each third iteration because the difference between the weighted actual values in the sample and the target values of the Mikrozensus or (N)Onliner Atlas undercut the abort criterion of 0,05. No trimming was necessary.

¹¹ Weights can only be calculated for those respondents who took part in the first wave, as all weighting-relevant content was collected there. Therefore, respondents who were not interviewed again until later waves cannot be taken into account (information on educational attainment can be obtained as of 2013 by merging with the 2013 Campaign Panel).

Table 25: Actual and target distributions of the variables which were used to calculate the weights by participation

Characteristic	Actual distribution (in percent)		Target distribution (in percent)	
	Participation in all waves	Participation in wave 5 to 9	Mikro-zensus 2016	(N)Onliner Atlas 2016
Gender				
male	52.1	51.7	48.6	51.6
female	47.9	48.3	51.4	48.4
Age group				
18 up to 30 years	7.4	9.3	15.9	21.0
30 up to 45 years	22.9	23.8	20.7	25.2
45 up to 60 years	39.1	38.0	29.0	33.8
60 years and older	30.7	28.9	34.4	20.0
Education				
low	21.8	21.9	37.6	35.5
intermediate	37.0	36.4	30.4	32.1
high	41.5	41.7	32.0	32.4
Federal state				
Western German states	76.5	76.8	79.1	80.7
Newly-formed German states (incl. Berlin)	23.5	23.2	20.9	19.3

In addition to the sample weights presented above, weights were also calculated in the same way for respondents who participated in all waves and for respondents who participated in all waves from wave 5 onwards. The calculation of these weighting variables was aborted after the fourth iteration since the difference between the weighted actual values in the sample and the target values of the (N)Online Atlas fell short of the abort criterion of 0.05. The calculation of the weighting variables was then interrupted after the fourth iteration. The weighting factors were trimmed above 5. This affects 54 cases for the weighting variable wei4_on. The weighting variables wei5_mz and wei5_on did not have to be trimmed. The following table provides an overview of all calculated weights.

Table 26: Overview of the weighting factors

	N	Mean	Std.Dev	Min	Max	1.Q	Median	3.Q	Max./Min. ¹²
wei_mz	15,802	1	0.41	0.59	2.27	0.71	0.83	1.15	3.85
wei_on	15,802	1	0.22	0.62	1.55	0.83	0.96	1.13	2.50
wei2_mz	18,128	1	0.37	0.61	2.12	0.72	0.82	1.20	3.48
wei2_on	18,128	1	0.23	0.61	1.72	0.82	0.97	1.12	2.82
wei3_mz	3,960	1	0.42	0.64	2.14	0.69	0.83	1.21	3.34
wei3_on	3,960	1	0.23	0.68	1.58	0.81	0.90	1.14	2.32
wei4_mz	6,778	1	0.57	0.49	4.74	0.64	0.78	1.22	9.67
wei4_on	6,778	1	0.69	0.38	5.00	0.54	0.74	1.16	13.16
wei5_mz	9,428	1	0.51	0.51	3.64	0.66	0.78	1.34	7.14
wei5_on	9,428	1	0.63	0.41	4.79	0.60	0.78	1.19	11.68

5.7. Attitudinal and behavioural variables

A simple scheme was applied to name the individual variables. The first three digits of the variable name are reserved for the respective wave, i.e., "kp1" for the first wave and "kp2" for the second wave¹³. Subsequently, the item number follows. For instance, the variable "Interest in politics" carries the item number "010". The variable "Interest in politics", which was collected in the first wave, can accordingly be found under the designation "kp1_010". Provided that an item was collected unaltered in the 2013 Campaign Panel as well as in the 2017 Campaign Panel, the item number remains the same in both datasets. When the question wording or the answer scales were edited, the last digit of the item number was increased by one.

¹² The value Max/Min indicates the relationship between the highest and the lowest weighting factor. Ideally, the weights do not become too big or too small, so a lower value is to be considered as positive.

¹³ Exceptions to this rule are the socio-structural characteristics of the respondents. As those are expected to be stable beyond the field time, variable names start with the abbreviation kpX, no matter in which waves the variables were initially collected.

Variable	Item	W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
----------	------	----	----	----	----	-----	----	----	----	----	----

Voting behaviour, retrospective		W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
kpX_2780	Turnout federal election 2013	X								X	
kpX_352	Recall previous federal election	X								X	
Ideology		W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
kpX_1490	Left-right assessment, parties		X		X				X		
kpX_680	Left-right assessment, chancellor candidates			X				X			
kpX_1500	Left-right self-assessment	X	X	X	X	X		X	X		
Variable	Opinions about political parties and the government	W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
kpX_430	Scalometer parties	X	X	X	X		X	X	X	X	X
kpX_730	Scalometer government	X	X	X	X	X	X	X	X	X	
kpX_731	Scalometer governing parties	X	X	X	X	X	X	X	X	X	
kpX_5021	Difference parties	X					X		X	X	X
kpX_5020	Difference government	X					X		X	X	X
Valence issues		W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
kpX_840*	Most important problem	X	X	X	X	X	X	X	X	X	X
kpX_850	Ability to solve the most important problem	X	X	X	X	X	X	X	X	X	X
kpX_860*	Second most important problem	X	X	X	X	X	X	X	X	X	X
kpX_870	Ability to solve the second most important problem	X	X	X	X	X	X	X	X	X	X
Position issues		W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
kpX_1070	Socio-economic dimension, parties		X		X				X		
kpX_1080	Socio-economic dimension, chancellor candidates			X				X			
kpX_1090	Socio-economic dimension, ego	X	X	X	X	X		X	X	X	
kpX_1110	Libertarian-authoritarian dimension, parties		X		X				X		
kpX_1120	Libertarian-authoritarian dimension, chancellor candidates			X				X			
kpX_1130	Libertarian-authoritarian dimension, ego	X	X	X	X	X		X	X	X	
kpX_1391	Security and privacy, parties		X		X				X		
kpX_1392	Security and privacy, chancellor candidates							X			
kpX_1411	Security and privacy, ego	X	X		X	X		X	X	X	

kpX_1230	European integration, parties							X			
kpX_1240	European integration, chancellor candidates							X			
kpX_1250	European integration, ego	X				X		X	X	X	
kpX_1260	Globalisation, ego										X
kpX_1270	Climate change, parties		X		X				X		
kpX_1290	Climate change, ego	X	X		X	X			X	X	
kpX_1210	Integration, Ego	X	X		X	X			X	X	
kpX_1100	Socio-economic dimension, importance		X		X				X	X	
kpX_1140	Libertarian-authoritarian dimension, importance		X		X				X	X	
kpX_1220	Integration foreigners, importance		X		X				X	X	
kpX_1300	Climate change, importance		X		X				X	X	
kpX_1421	Security and privacy, importance		X		X				X	X	
kpX_2880	Ego positions	X	X	X	X		X	X	X	X	X
Candidates		W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
kpX_650	Scalometer politicians	X	X	X	X	X	X	X	X	X	X
kpX_671	Preferred chancellor		X	X	X		X	X	X	X	
kpX_661	Attributes of chancellor candidates, Merkel			X			X				
kpX_661	Attributes of chancellor candidates, Schulz			X			X				
kpX_662	Ambivalence politicians, negative feelings				X			X			
kpX_663	Ambivalence politicians, positive feelings				X			X			
Strategic Voting		W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
kpX_910	Scalometer coalitions			X			X	X	X	X	X
kpX_962	Expected coalition			X			X	X	X		
kpX_911	Participation of AfD in a coalition							X	X	X	
kpX_1050	Perception opinion polls						X	X	X		
kpX_1051	Credibility opinion polls						X	X	X		
kpX_3010	Probability of smaller parties entering parliament						X	X	X		

[illegible]

[illegible]

kpX_916	Exploratory talks Jamaica coalition, attitude										X
kpX_917	Exploratory talks Jamaica coalition, responsibility										X
kpX_918	Behaviour parties, overall										X
kpX_919	Behaviour parties, willingness to compromise										X
kpX_920	Behaviour parties, faithfulness to principles										X
kpX_921	Behaviour parties, assertiveness										X
kpX_922	Duration of negotiations, attitude										X
kpX_923	Results of negotiations, attitude										X
Election campaign on the Internet		W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
kpX_3490	Wahl-O-Mat, usage							X	[X]	[X]	
kpX_3500	Wahl-O-Mat, reason for use							X	[X]	[X]	
kpX_3510	Wahl-O-Mat, congruence							X	[X]	[X]	
kpX_1610	Use of social media, politically current				X		X		X		
kpX_1615	Use of social media, party related				X		X		X		
kpX_1616	Online participation, politically current				X		X		X		
Televised debate		W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9
kpX_1800	Televised debate: Reception							X	[X]		
kpX_1810	Televised debate: Perception of performance Merkel							X	[X]		
kpX_1820	Televised debate: Perception of performance Schulz							X	[X]		
kpX_1870	Televised debate: Reception of media response							X	[X]		
kpX_1880	Televised debate: Perceived media response Merkel							X	[X]		
kpX_1890	Televised debate: Perceived media response Schulz							X	[X]		
kpX_1900	Televised debate: Reception of response of social environment							X	[X]		
kpX_1910	Televised debate: Perceived response of social environment Merkel							X	[X]		
kpX_1920	Televised debate: Perceived response of social environment Schulz							X	[X]		
Variable	Personality, values, basic orientations	W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9

kpX_3320	Schwartz values	X				X						
kpX_2180	Big 5	X				X						
kpX_1572	Willingness to take risks	X				X						
kpX_2090	Party identification	X	X	X	X	X	X	X	X	X	X	X
kpX_2100	Party identification, strength	X	X	X	X	X	X	X	X	X	X	X
kpX_2095	Multiple party identification	X	X	X	X	X	X	X	X	X	X	X
kpX_2101	Multiple party identification, strength	X	X	X	X	X	X	X	X	X	X	X
kpX_020	Satisfaction with democracy	X					X			X	X	
kpX_050	Principles of democracy		X			X						X
kpX_040	Attitudes to parties in general	X					X			X	X	
kpX_3103	Populism						X			X	X	
kpX_050	Efficacy and duty to vote	X					X			X	X	
kpX_1483	Foreign policy orientations		X			X						
kpX_060	Extremism/Authoritarianism		X			X						X
kpX_5000	National identity	X	[X]			X						X
kpX_5010	Form of national identity	X				X						X
kpX_160	Confidence in institutions		X			X				X	X	
kpX_2200	Attachment battery			X		X						X
kpX_1555	Political motivation			X				X				
kpX_1570	Need for Cognitive Closure			X		X						
Socio-demographics		W1	W2	W3	W4	WA1	W5	W6	W7	W8	W9	
kpX_2441	Household size	X				X						
kpX_2450	Number members of the household under 18 years	X				X						
kpX_2590	Net household income, with categories	X				X						
kpX_2461	Membership of organisations			X								
kpX_2470	Membership of trade unions, household			X								
kpX_4000	Membership party			X								
kpX_4110	Membership party, other party			X								
kpX_3910	Religiousness	X				X						
kpX_2481	Religious denomination	X				X						

kpX_2491	Church attendance	X		X	
kpX_2580	Subjective perception of class	X		X	
kpX_2601	Federal state	X	[X]	X	X
kpX_2600	Residence	X		X	
kpX_2602	Postal code	X		X	
kpX_2320	School leaving certificate	X		X	X
kpX_2330	Vocational and professional training	X		X	
kpX_2340	Gainful employment	X		X	
kpX_2350	Former gainful employment	X		X	
kpX_2371	Duration of unemployment	X			
kpX_2380	Profession	X		X	
kpX_3610	Employee - differentiation	X		X	
kpX_3620	Worker - differentiation	X		X	
kpX_3630	Independent professional - differentiation	X		X	
kpX_3640	Civil servant - differentiation	X		X	
kpX_3650	Employment sector	X			
kpX_3660	Sector of the economy	X			
kpX_3670	Temporary/agency work	X			
kpX_3680	Fear of job loss	X			
kpX_3690	Fear of losing professional career	X			
kpX_2390	Former profession	X		X	
kpX_3710	Former profession, employee - differentiation	X		X	
kpX_3720	Former profession, worker - differentiation	X		X	
kpX_3730	Former profession, independent professional - differentiation	X		X	
kpX_3740	Former profession, civil servant - differentiation	X		X	
kpX_3750	Former profession, employment sector	X			
kpX_3760	Former profession, sector of the economy	X			
kpX_3920	Country of birth, federal state	X			
kpX_3930	Country of birth, other country	X			
kpX_2550	Age at immigration	X			

kpX_3940	Age at immigration, East Germany	X			
kpX_3950	Age at immigration, West Germany	X			
kpX_3960	Age moving to Berlin	X			
kpX_2520	German citizenship		X	[X]	
kpX_2571	Country of birth mother		X	[X]	
kpX_2571	Country of birth, father		X	[X]	
kpX_4131	Language spoken in household		X	[X]	
kpX_2301	Marital status	X			X
kpX_2311	Respondent has a partner	X			X
kpX_2312	Partner living in the same household	X			
kpX_2391	School leaving certificate partner	X			
kpX_2400	Employment partner	X			
kpX_2410	Former employment partner	X			
kpX_2420	Profession partner	X			
kpX_2430	Former profession partner	X			

Explanation: "X" represents the respective wave. [X] means that only those participants of the election campaign who had not participated in one of the retrospective waves in which this group of issues had already been questioned received the module questions (see also section 6.7.).

5.8. Unpublished variables

When conducting online surveys, variables are collected or transmitted to GESIS that are relevant for the implementation of the online survey but do not have any textual meaning. These variables are not published, but can be obtained from GESIS (gles@gesis.org) if needed. For data protection reasons, some variables cannot be made available for free download (e.g. postcode). These variables can be obtained under specific conditions.

Table 28: Overview of unreleased variables and subscription opportunities

Variable	Explanation	Subscription opportunity
kpX_ats	Absolute time stamp	available on request
kpX_page_history	Page history	available on request
kpX_2293s	Place of birth (first letter)	available on request
kpX_2062/kpa1_2602	Postal code	available on request
kpX_xxxx_org, _c1, _c2	Several original variables and request variables which were created whilst summarizing the request variables, party versions and dependent interviewing variables in order to back up the original data; or are not needed any longer	available on request
kpX_2280_c1	Gender	available on request
kpX_2291_c1	Month of birth	available on request
kpX_2290_c1	Year of birth	available on request
kpX_2293s_c1	Place of birth	available on request
kpX_2280_c2	Gender	available on request
kpX_2291_c2	Month of birth	available on request
kpX_2290_c2	Year of birth	available on request
kpX_2293s_c2	Place of birth	available on request
kpX_4280	Reason for wrong identification code	available on request
Master data		
s0	Gender	available on request
s1	Day of birth	available on request
s2	Month of birth	available on request
s3	Year of birth	available on request
s4	Marital status	available on request
s5	Schooling	available on request
s6	Vocational education	available on request
s7	Employment	available on request
s8	Sector	available on request
s9	Profession	available on request
s10	Professional department	available on request
s11	Net income	available on request
s12	Household income	available on request
sHH	Number of household members	available on request
sHHK	Number of children in the household	available on request

Explanation: "X" represents the respective wave.

6 Data preparation

6.1. Preliminary note

In the first step, the data preparation of the Campaign Panel data was done individually for every wave. The datasets of the individual waves were subsequently converted into a joint dataset. All treatment steps are based on the directive of minimal invasive editing. Data errors were marked with flag variables. The preparation was conducted by means of syntax and can therefore be reproduced in all its steps.

6.2. Encoding of missing values

Missing values were assigned conforming to the uniform encoding scheme of GLES in all cases. This includes negative values from -71 to -99 for SPSS format and Stata codes from .a to .p. A Do-File, with which the SPSS missing codes can be transformed to the intended Stata codes in the Stata dataset, is enclosed with the dataset.

Table 29: Brief overview of the missing value codes of GLES

SPSS-Code	Stata-Code	Label
-71	.p	subject unknown*
-72	.o	not rateable*
-84	.k	no cast of first/second vote
-85	.j	no vote
-92	.h	error in data
-93	.g	not asked, terminated
-95	.e	not participated
-97	.c	not applicable
-98	.b	don't know
-99	.a	no answer

*The variables elecdist – elecdist4 are labelled as -71 "unknown postal code" and -72 "postal code not clearly assignable".

6.3. Encoding of the parties

In order to be able to compare all GLES components better with each other, the encoding of political parties was done according to a uniform encoding scheme. This approach was applied to all questions which included closed or open requests concerning political parties. The encoding scheme can be found on the GLES pages of the GESIS website (<http://www.gesis.org/gles>). All information concerning political parties is released in two variables which are marked version A and version B. Version A only includes parties which were represented in the Bundestag for the 2009-2013 legislative period, as well as the category "other party". Version B differs from A in that way that it provides a detailed list of smaller parties. The respondents were only able to choose these parties if they had chosen "other party" in the previous question. The question only included the parties which are represented in the German Bundestag at first.

Some respondents did not provide an answer to the request to choose between the small parties, or they interrupted the survey. Those persons were assigned the code -99 (no answer) or -93 (not asked,

terminated) in both versions A and B, although partial information was accessible in these cases. They chose "other party" on the first page.

6.4. Encoding of open questions

Open questions (questions without fixed answer categories) were encoded by the polling agency BACES. Encoding was done by using encoding schemes that were developed by the GLES team. This includes the questions about the most important and the second most important issue in Germany (agenda questions, "kpX_840", "kpX_860") and the question about the "Reasons for Voting Decision" ("kpX_260"). The corresponding encoding schemes are accessible on the GLES pages of the GESIS website (www.gesis.org/gles).

6.5. Marking of knowledge questions

The political and economic knowledge of the Campaign Panel participants was tested with several questions. The suitable answer to a knowledge issue was marked with a star (*) in the corresponding value label.

Table 30: Knowledge issues with marked answers

Variable	Item	Wave
kpX_110	Political knowledge: First/second vote	1, 6, A1
kpX_130	Political knowledge: Electoral law	1, 6, A1
kpX_3430q,c,l,p,m,r	Political knowledge: Matching politicians/parties	2, 4, 6
kpX_3430qj,a,s,t,u,v	Political knowledge: Matching politicians/parties	3, 5, 7
kpX_3440	Political knowledge: Unemployment rate	2,7

The question about the 5%-threshold ("kpX_090") is an exception and was asked in the first and sixth as well as in the profile wave of the refreshment sample. As the question was posed openly, the answers were transferred to a new dummy variable ("kpX_090_v1") which indicates whether the respondent answered correctly.

6.6. Handling of questions with "checkboxes" and "slide controls"

Some forms of questions required an enquiry if the respondent did not answer in order to clarify the meaning. This included the forms "checkbox" and "slide control" which both had a default value (mostly zero). If respondents clicked through this type of questions without making adjustments in neither the checkboxes nor the slide controls, it would not be possible to draw a distinct line between possible substantive replies, e.g. "zero days", and non-substantive refusal of answering. Therefore, the reason for the non-given answer was investigated. It was asked whether the respondent did not want to state an answer or if the value "zero" represented a substantial answer which was to be interpreted. In addition, the respondents were given the possibility to correct their answers and to state values which differed from zero. The immediate replies and the answers that were collected by request are summarized in the stem variable in the released dataset. Whether the answer was given on request can be found out by looking for flag variables which consist of the variable stem name and the suffix "flag". The variables which are shown in Table 31 are affected by this issue.

Table 31: List of affected variables whose inquiries were summarized

Variable	Item	Form of Question	Wave
kpX_1661a-h	Use of print media, politically current	Slide control	1, 3-8
kpX_1681a-f	Use of TV, news, current	Slide control	1, 3-8
kpX_1933	Conversations about politics, in general	Slide control	1-8
kpX_1945	Discussion partner, frequency	Slide control	1,-8
kpX_421	Contact with parties I and II	Checkbox	4-8
kpX_1616	Online participation, politically current	Checkbox	4, 5, 7

Explanation: "X" represents the respective wave.

6.7. Summary of dependent interviewing variables

The term "Dependent Interviewing" (DI) describes in this study a special form of filtering, in which it depends on information from *previous waves* whether a question is posed and how it is worded. In the second wave of the Campaign Panel, this technique was used to collect data for respondents recruited in 2013 who had not participated in the first wave.¹⁴

Table 32: List of variables for which a subsequent summary of measured values has been carried out

Variable	Item	First appearance in wave ...	Retaken in wave ...
kpX_2601	Federal state	1	2
kpX_5000a-c	National identity	1	2
kpX_2520	German citizenship	2	3
kpX_2571a	Country of birth mother	2	3
kpX_2571b	Country of birth, father	2	3
kpX_2572a	Country of birth, mother, other country	2	3
kpX_2572b	Country of birth, father, other country	2	3
kpX_4131	Language spoken in household	2	3

¹⁴ The data of the respondents from the profile wave of the refreshment sample were summarized in the variables kp1_2601 and kp1_5000. The point of time of each response can be looked at by using the variable kp1_2601flag or kp1_5000flag.

7 Data quality

Preliminary note

As there are no interviewers present in online surveys, the interview situation is characterized by a high level of anonymity.¹⁵ Anonymity can be an advantage if data about socially undesirable attitudes or behaviour patterns is collected, as respondents tend to answer more honest in self-administrative surveys. (cf. Joinson, 1999: 435; Mühlenfeld, 2004; Taddicken, 2009). However, a disadvantage of anonymity is that nobody controls the carefulness and seriousness of the participants' responsiveness (Gräf/Heidingsfelder, 1999:120). Thus, this is an encouragement for less motivated participants to answer the questions superficially or imprecise ("satisficing", cf. Krosnick, 1991, 1999). In this way, they receive incentives without putting a high amount of effort into the survey. This form of responsiveness manifests itself in various ways in the responsiveness, for example in responding very quickly to the questions, in a high share of refusal, random answers or undifferentiating answers ("straightlining"). Furthermore, it is a fact that a small percentage of respondents in online access panels states wrong answers intentionally (Downes- Le Guin et al., 2006; Baker/Downes-Le Guin, 2007).

Despite several measures to ensure the establishment of data quality, the 2017 Campaign Panel is not immune to these issues. One should consider at the same time that surveys with telephone or personal interviews also struggle with data quality issues (even though another kind of issues, e.g. fake interviews in CAPI studies or issues with non-visualized Likert scales in CATI surveys), but those are less noticeable in the data and are therefore rarely made a subject of discussion. The relatively poor reputation of online surveys in comparison to other modes of data collection partly stems from the high degree of mechanization of the data collection processes, in which numerous measured values arise. However, these data contribute to making the dimension of lacking data quality transparent (e.g. by automatic measurement of response times).

This transparency is not a disadvantage, but a great advantage of online studies: Although the share of problematic data in online surveys is higher than in personal interviews, the studying of poor data raises the awareness for the existing issue and forces research to grapple with it at the same time. Moreover, data quality measurements could be helpful to test the sensitivity of analytical results with regard to different degrees of data quality.

The 2017 Campaign Panel includes a great number of different indicators which can and should be used to evaluate data quality. The release at hand contains information about the speed of given answers (section 7.1.), Bogus Items (section 7.2.) as well as self-assessment of the data quality (section 7.4.).

Data quality issues in online surveys not only result from a lack of motivation on the part of the respondents but also - although quite rarely - due to technical errors. Filter errors are described in section 7.3. In section 7.4., cases will be listed in which the respondents pointed out errors themselves.

¹⁵ This section is strongly oriented towards the statements of Plischke (2014: 197-198).

7.1. Response times

If respondents give fast answers, this does not necessarily indicate lacking data quality. Many people are able to read questions quickly, and their stable opinion enables them to answer the questions in a short amount of time. At the same time, it can be assumed that data quality suffers if the median response time is significantly undershot, as there is not enough time to read the question text and give well-considered answers. Often, fast response times indicate that respondents just click through the survey, giving arbitrary answers or not giving any information.

There are no established standards for the identification of "too fast" responses in the scientific literature. As a general rule, the measures of their identification include the median or average value of the distribution and the scattering. Based on that cut-off, criteria are chosen which must not be undercut (Mayerl/Urban 2008, 58ff). Those respondents are either excluded from the dataset or highlighted by marker variables.

Since the 17th survey of the Longterm-Online-Tracking (T17 (ZA5350)), GLES identifies speeding respondents using a modified algorithm further developed by Roßmann (2010). To this end, a speeder index is generated (variable "kpX_speederindex"¹⁶) which includes the participants' response times on every page of the survey as well as the overall duration per respondent. The index values range between greater 0 and less than 2. An index value of 1 indicates an average response time; values approaching 0 indicate very short response times and values approaching 2 indicate very slow response times. This index variable allows the user to exclude speeding respondents from the analysis.

When analysing the Campaign Panel data, it is recommended to test the impact of the inclusion/exclusion of speeding respondents on the results. On the basis of his/her own experience with the speeder index, the user should decide from which cut-off an exclusion of the respondents appears meaningful (and document this adequately in research work).

7.2. Bogus items

"Satisficing" in online surveys implies that people click through answer categories randomly or - in the particular case of matrix questions - frequently select answer options of the same column ("straightlining") without having read the questions thoroughly. In order to identify such behaviour of responding, control items were included in matrix questions in four waves. In this item, the respondents were not asked to state their political opinion, but to click a particular response category (see Figure 5). If this category was not chosen, this does not necessarily imply poor data quality, as respondents might also have provided wrong information deliberately (e.g. to protest against the control measures). However, preliminary analyses have already shown that wrong control item answers are strongly associated with straightlining and rapid answers. The Bogus items are marked in the dataset by the suffix "q" (e.g. "kp1_050q").

¹⁶ For respondents of the first wave, two speederindex variables are available because the questionnaire of the respondents recruited in 2013 differed from the questionnaire of the new respondents recruited in 2016. From wave 2 onwards, only one speederindex was calculated, as the questionnaires of these two sample groups were identical.

Figure 5: Example of a control question (screenshot from wave 1)

	strongly disagree	rather disagree	neither agree nor disagree	rather agree	strongly agree
Politicians care about what ordinary people think.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Politicians try to get in close contact with the population.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am confident that I could actively participate in a conversation dealing with political issues.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am perfectly able to understand and assess important political questions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Please choose „rather agree“ for testing the functioning of the questionnaire.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In a democracy it is the duty of all citizens to vote regularly in elections.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Note: The orange box was added subsequently to highlight the control question. The participants were asked in German, so the example has been translated.

7.3. Filter errors

Filter errors in wave 1

- kp1_1500: Due to a programming error, all "don't know" values (-98) were saved as "no answer" (-99).
- A respondent has entered an e-mail address instead of letters in the name generator (kp1_1937a-c). This information was recoded to -92 "error in data".
- 121 respondents did not answer the must answer question kp1_1933 (Conversations about politics, in general) and were mistakenly forwarded to the next block. The values for these persons were recoded in the entire module (up to and including variable kp1_1971cb) into -92 "error in data".
- A respondent did not provide any information on the must answer question kp1_1945a-c (Discussion partner, frequency) and was mistakenly forwarded to the next block. The values for this person were converted to -92 "error in data".
- 17 respondents did not answer the must answer question kp1_1681a-f (Use of TV, news, current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".
- Two respondents did not answer the must answer question kp1_1661a-g (Use of print media, politically current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".

Filter errors in wave 2

- 56 participants from the 2013 Campaign Panel were able to verify themselves despite errors in the stored data (see section 4.7.). These can be identified via kp2_info1
- For one person, the slide control for the variables kp2_1933 (Conversations about politics, in general) and kp2_1945a-c (discussion partner, frequency) did not work. The values for this person were recoded to -92 "error in data".

Filter errors in wave 3

- 6 respondents did not answer the must answer question kp3_1681a-f (Use of TV, news, current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".

- 3 respondents did not answer the must answer question kp3_1661a-g (Use of print media, politically current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".
- 22 respondents did not answer the must answer question kp3_1933 (Conversations about politics, in general) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".
- A respondent did not provide any information on the must answer question kp3_1945a-c (Discussion partner, frequency) and was mistakenly forwarded to the next block. The values for this person were converted to -92 "error in data".
- kp3_221a,c-g,i: Due to a programming error in the filter, 7 respondents who indicated to kp3_190bb to choose Liberal-Konservative Reformer (code 331) or Bündnis C (code 351) did not see the question. The values for these persons were converted to -92 "error in data".

Filter errors in wave 4

- A respondent did not provide any information on the must answer question kp4_421ka-ei (Contact with parties) and was mistakenly forwarded to the next block. The values for this person were converted to -92 "error in data".
- 2 respondent did not provide any information on the must answer question kp4_1616a-c, y (Online participation, politically current) and was mistakenly forwarded to the next block. The values for this person were converted to -92 "error in data".
- 6 respondents did not answer the must answer question kp4_1681a-f (Use of TV, news, current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".
- 2 respondents did not answer the must answer question kp4_1661a-g (Use of print media, politically current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".
- 27 respondents did not answer the must answer question kp4_1933 (Conversations about politics, in general) and were mistakenly forwarded to the next block. The values for these persons were recoded in the entire module (up to and including variable kp4_1971cb) into -92 "error in data".
- In the variable kp4_2095 (Multiple party identification), 18 persons could name the same party again from kp4_2090a and kp4_2090b (Party identification). Their data were recoded into -92 "error in data".
- kp4_221a,c-g,i: Due to a programming error in the filter, 8 respondents who indicated to kp4_190bb to choose Liberal-Konservative Reformer (code 331) or Bündnis C (code 351) did not see the question. The values for these persons were converted to -92 "error in data".

Filter errors in wave 5

- 5 respondents did not answer the must answer question kp5_1681a-f (Use of TV, news, current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".
- 3 respondents did not answer the must answer question kp5_1661a-g (Use of print media, politically current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".
- 21 respondents did not answer the must answer question kp5_1933 (Conversations about politics, in general) and were mistakenly forwarded to the next block. The values for these persons were recoded in the entire module (up to and including variable kp5_1971cb) into -92 "error in data".
- kp5_221a,c-g,i: Due to a programming error in the filter, 8 respondents who indicated to kp5_190bb to choose Liberal-Konservative Reformer (code 331) or Bündnis C (code 351) did not see the question. The values for these persons were converted to -92 "error in data".

Filter errors in wave 6

- 5 respondents did not answer the must answer question kp6_1681a-e (Use of TV, news, current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".
- kp6_2880j,ae,ac,x,z,ab,ad,aa,w,k,j: Due to a premature field start, people with a running number of up to 45 have seen temporary screens instead of the final ego positions that were not released for field start. The data of 38 respondents were therefore converted to -92 "error in data".
- 2 respondents did not answer the must answer question kp6_1661a-g (Use of print media, politically current) and were mistakenly forwarded to the next block. The values for these persons were converted to -92 "error in data".
- 21 respondents did not answer the must answer question kp6_1933 (Conversations about politics, in general) and were mistakenly forwarded to the next block. The values for these persons were recoded in the entire module (up to and including variable kp6_1971cb) into -92 "error in data".
- kp6_1941a-c: A respondent was filtered, although he or she specified a first name initial at kp6_1937. His data has been encoded into -92 "error in data".
- kp6_221a,c-g,i: Due to a programming error in the filter, 3 respondents who indicated to kp6_190bb to choose Liberal-Konservative Reformer (code 331) or Bündnis C (code 351) did not see the question. The values for these persons were converted to -92 "error in data".

Filter errors in wave 7

- 15 respondents did not answer the must answer question kp7_1933 (Conversations about politics, in general) and were mistakenly forwarded to the next block. The values for these persons were recoded in the entire module (up to and including variable kp7_1971cb) into -92 "error in data".
- A respondent did not provide any information on the must answer question kp7_1945a-b (Discussion partner, frequency) and was mistakenly forwarded to the next block. The values for this person were recoded into -92 "error in data".
- In the variable kp7_2095 (Multiple party identification), 19 persons could name the same party again from kp7_2090a and kp7_2090b (Party identification). Their data were recoded into -92 "error in data"
- kp7_221a,c-g,i: Due to a programming error in the filter, 6 respondents who indicated to kp7_190bb to choose Liberal-Konservative Reformer (code 331) or Bündnis C (code 351) did not see the question. The values for these persons were recoded into -92 "error in data".

Filter errors in wave 8

- 3 respondents did not answer the must answer question kp8_1681a-e (Use of TV, news, current) and were mistakenly forwarded to the next block. The values for these persons were recoded into -92 "error in data".
- 2 respondents did not answer the must answer question kp8_1661a-g (Use of print media, politically current) and were mistakenly forwarded to the next block. The values for these persons were recoded into -92 "error in data".
- 23 respondents did not answer the must answer question kp8_1933 (Conversations about politics, in general) and were mistakenly forwarded to the next block. The values for these persons were recoded in the entire module (up to and including variable kp8_1971cb) into -92 "error in data".

Filter errors in wave 9

- 4 respondents did not see the question kp9_211a, c-f,i,g although they met the filter condition. Their data was converted to -92 "error in data".

7.4. Respondents feedback

At the end of each survey, respondents were invited to give feedback. In addition to largely positive feedback, some comments provide information about technical problems or incorrect answers. If the mistake was clearly specified incorrect answers were corrected. Display errors affecting individual respondents were indicated in the respective variable with the error code -92 "error in data". These few cases related almost exclusively to user-defined survey pages (e.g., sample ballot paper, pictures of politicians, slide control).

8 Matching of other datasets

8.1. Matching of time variables

The corresponding time variables are made accessible in a separate dataset via download. They can be matched to the Campaign Panel by means of the variable "lfdn" (serial number, i.e. ID). Alternatively, it is possible to download a Stata Do-File or a SPSS syntax file which can be applied to combine the datasets.

The dataset contains two groups of time variables: The first group reflects the participant's response time on a particular page of the questionnaire. These variables are named according to the scheme T_name of the variable (e.g. "T_kp1_010" for political interest).

The second group of time variables contains the cumulated time, which was necessary to get to the respective point in the survey, for every respondent. These cumulated time variables are numbered consecutively corresponding to the questions' order in the questionnaire, whereas the number in the designation after T marks the spot within the survey sequence. For instance, the political interest was queried in the first wave on the eighth screen. Therefore, the cumulated time variable was named "T8_kp1_010".

8.2. Matching of the GLES Campaign Panel 2013

Meanwhile, the dataset "Repeatedly questioned respondents of the Short-term Campaign Panel 2013 and 2017 (GLES)" has been published by GESIS under the number ZA6827, which combines the Short-term Campaign Panel 2013 and 2017, as well as the interim surveys in 2014 and 2015.

8.3. Matching of the control groups

As described in section 2.2., three cross-sectional interviews were carried out simultaneously to the fourth, sixth and eighth interview of the Campaign Panel. The control group included about 1,000 people in each interview and was carried out with an almost identical questionnaire. The control group data was published by GESIS under the ZA numbers 6805, 6806 and 6807. As the variables of the Campaign Panel and the control group have the same names, the datasets can be stacked without any problems. GESIS provides a Stata Do File on its website.

9 Version history

Changes from version 1.0.0 to version 2.0.0

The second wave has been added to the dataset. In a line with the addition of the second wave variables invariant in time have been renamed into the kpx_ pattern. This affects "kp1_2280", "kp1_2291", "kp1_2290", "kp1_2293s", "kp1_2390", "kp1_3710", "kp1_3720", "kp1_3730", "kp1_3740", "kp1_3750", "kp1_3760", "kp1_3921", "kp1_3930", "kp1_2550", "kp1_3940", "kp1_3950", "kp1_3960" and "kp1_2350". Those variable names now start with "kpx_" instead of "kp1_".

Weights regarding speeders have been dropped ("wei_mzz", "wei_mzoz", "wei_onz", "wei_onoz") and are replaced by weights affecting all cases ("wei_mz", "wei_on").

Notifications of filter errors have been changed. Instead of flagging them in kp1_info1 and kp1_info2 affected respondents are recoded into -92 "error in data" (kp1_1681a-e, kp1_1661a-g).

The following variables have been removed from the first wave: "kp1_speederflag_WKP17", "kp1_speederflag_WKP13", "kp1_ats" and "kp1_device_type".

The following variables have been added to the first wave: "kp1_compl", "kp1_dispcode" and "kp1_quota".

A wrong recoding of -97 "not applicable" and -93 "not asked, terminated" has been corrected in the first wave.

Correction of variable and value labels.

Changes from version 2.0.0 to version 3.0.0

The third wave has been added to the dataset.

For the answer of 107 respondents in the verification process who also participated in 2013 and happened to have a typo, the typo instead of the corrected answer was saved. This has been corrected.

Respondents breaking off before or during the verification process of wave 2 have been recoded into -95 "not participated".

Coding of kp1_1702 has been corrected.

Correction of wrong code of party "Bündnis C - Christen für Deutschland" in wave 2.

Correction of variable and value labels.

Changes from version 3.0.0 to version 4.0.0

Waves 4-8 have been added to the dataset.

Between the fourth and the fifth wave, new respondents were added and they were interviewed from wave five onward.

This resulted in the possibility to provide information on time-invariant characteristics in various waves. Variables "kpx_2280flag", "kpx_2291flag", "kpx_2290flag", "kpx_2350flag", "kpx_2390flag", "kpx_3710flag", "kpx_3720flag", "kpx_3730flag" and "kpx_3740flag" give information on when respondents gave these information.

The variable "wkp13" has been replaced by the variable "gruppe", to identify respondents of the 2013 Campaign Panel as well as respondents added between waves four and five.

Recoding of variables kp3_4000 (Membership party) and kp3_4110 (Membership party, other party) in party versions A and B.

For variables kpX_2090a (Party identification) and kpX_2095a (Multiple party identification), CDU and CSU have been listed separately.

Because of an erroneous variable kpX_compl, not all disposition codes had been assigned correctly. This fault has been removed.

Correction of variable and value labels.

Changes from version 4.0.0 to version 5.0.0

Wave 9 has been added to the dataset.

The variables elecdist, elecdist1, elecdist2, elecdist3, and elecdist4 have been updated. The assignment of constituencies is now based on the Stata ado-file "plztownr" and accounts for the geographic boundaries of the electoral districts in 2017.

The variable kpx_info has been updated and gives information on inconsistencies in the reported residence of respondents. This includes not matching federal states and constituencies as well as changes of the federal state between wave 1/a1 and wave 9.

The variable kp8_211 has been recoded because the values were incorrectly replaced by the values of the variable kp8_221 in earlier versions. The error has been corrected.

The variable kp7_compl was erroneous and has been corrected.

An error occurred during programming with the variable kp8_341a. Here CSU/CSU was wrongly specified as label (instead of CDU/CSU). The variable is still in the dataset.

The variable lfdn has been inconsistent. All cases, which were already included in the first wave, were assigned a new consecutive number in each released version of the dataset. This fault has been removed, and the consecutive number corresponds to the latest dataset version (v4.0.0).

Post-stratification weights have been recalculated based on the marginal distributions from the German micro census 2016.

Correction of variable and value labels.

Changes from version 5.0.0 to version 6.0.0

Open questions (questions without given answer possibilities) of all waves were coded by the survey institute BACES on the basis of the coding schemes of the GLES and included in the dataset (kpX_840_c1-5, kpX_860_c1-5, kpX_260_c1-3).

The dichotomous variables "wXa" were recoded based on the short-term campaign-panel 2013. The incorrect coding of respondents who had already aborted the survey before the start page and during verification process into category 1 "interrupted/finished survey" was corrected. In the current coding, all respondents who did not answer any of the survey's content questions - regardless of whether their eligibility is missing or unclear - will be coded as "0 survey not started" because no content questions were seen and/or answered. As direct consequence, 18 respondents were identified, who did - based on the new definition - not participate in any of the waves.

The Variable „kpx_info" was wrongly calculated and has been corrected.

The missing categories of the constituency coding has been updated (elecdist, elecdist1-elecdist4). It now differentiates between respondents, who refused to name their postal code, and those, who named a postal code which could not be assigned to a constituency.

Variable names have been adjusted according to the Short-term Campaign-Panel 2013 (ZA5704). This affects the variables:

Old variable name	New variable name	Explanation
kpX_650x,z,v	kpX_650p,z1,v2	The suffixes in the variable "Politician scalometer" have been adjusted so that each suffix clearly identifies only one politician. In the 2017 Campaign Panel, Frauke Petry was originally assigned the suffix "x", which had already been assigned to Bernd Schlömer in the 2013 Campaign Panel. Petry is given the new suffix "p". Similarly, the new suffix for Martin Schulz is "z1" ("previously "z") and Katja Kipping is given the new suffix "v1" (previously "v").
kpX_250	kpX_252	The variable was originally listed as kpx_250 in waves 8 and 9. However, the question on the reasons for the decision not to vote (with categories) has not changed significantly. Only the time of query (before and after the election) deviates from the 2013 Campaign Panel.
kpX_262*	kpX_260*	In the 2017 Campaign Panel, the respondents were asked about the reasons for their voting decision (open answer) after stating their actual voting decision only.
kpX_352	kpX_350	The variable was listed as kpx_352 in wave 1 and 8. The question text only differs slightly in the 2013 Campaign Panel if respondents previously indicated that they could not remember their voting decision in the previous federal election.
kpX_2550	kpX_2551	No changes in the question text between the Campaign Panel in 2013 and 2017. In the 2017 Campaign Panel, an additional plausibility check was implemented (if entry>2016: "Your specified year is in the future. Please check your entry.").
kpX_3921	kpX_3920	In the 2013 Campaign Panel, the respondents were asked whether their country of birth was Germany and, as a follow-up question, they should indicate the federal state they were born in. In the 2017 Campaign Panel, the filter question was omitted and an additional answer category "not born in the territory of the present-day Federal Republic of Germany" was presented. However, the question text was identical.
kpX_2481	kpX_2480	No changes in the question text between the Campaign Panel in 2013 and 2017. In the 2017 Campaign Panel, however, the answer category of "belonging to a Muslim community/Islam" was offered.
kpX_2590	kpX_2591	The question on the net household income (with categories) was identical in both campaign panels.

Post-stratification weights have been recalculated based on the marginal distributions from the (N)Onliner-Atlas 2016 (wei_on – wei5_on). Several corrections in the variables and labels were implemented.

Changes from version 6.0.0 to version 7.0.0

Due to an error when loading the dat, open question verbatims of the respondents in wave 4 were incorrectly coded to -95 "Errors in the data". The error was corrected affecting the variables kp4_840_c1, kp4_840_c2, kp4_840_c3, kp4_840_c4, kp4_840_c5, kp4_860_c1, kp4_860_c2, kp4_860_c3, kp4_860_c4, kp4_860_c5.

Depending on its structure, each GLES data set is enriched with up to seven variables that identify the data set and describe its main features (metadata). The variable scheme has been updated and implemented in the current version of the data set which results in the following changes:

Old variable	Neue variable	Explanation
study	study	The numerical variable study contains as value the (ZA-)study number of the data set in four-digit format. In addition to the ZA number, the value label now also specifies the title of the respective study.
gruppe	sample	The numerical variable "sample" is used to allocates respondents to their sample or survey group. This variable is newly introduced for all GLES studies and harmonizes these information. In the election campaign panel it replaces the variable "group", which previously contained this information.
field	field_end, field_start	In the string variable "field_start" the date of the first day of the survey is stored in the format 'YYYYY-MM-DD'. In the string variable "field_end" the date of the last day of the survey is stored in the format 'YYYYY-MM-DD'. They replace the variable field, in which this information was previously summarized.
kpX_field	dropped	Information on the field time of the waves can be found in this study description.
year	dropped	The variable "year" contains the same information as the newly created variable "sample" and will not be published.
survey	dropped	The variable "survey" contains the same information as the updated variable "study" and will not be published.

The following variable names have been updated:

Old variable name	New variable name	Explanation
elecdist, elecdist1, elecdist2, etc.	elecdist17_, elecdist17_1, elecdist17_2 etc.	The variable name of the constituency allocation did not contain the year of the federal election, when the constituencies were effective. . This has been adjusted.

An English version of the data sets, the questionnaire and the study documentation were published.

References

- Baker, Reg and Theo Downes-Le Guin (2007): Separating the Wheat from the Chaff: Ensuring Data Quality in Internet Samples. Paper presented at: The Challenges of a Changing World: Proceedings of the Fifth International Conference of the Association for Survey Computing in Southampton, 12-14 September.
- Blumenberg, Manuela S. and Tobias Gummer (2013): Gewichtung in der German Longitudinal Election Study 2009. In: GESIS – Technical Reports 2013/19.
- Blumenstiel, Jan Eric and Tobias Gummer (2012): Langfrist-Panels der German Longitudinal Election Study (GLES): Konzeption, Durchführung, Aufbereitung und Archivierung. GESIS - Technical Reports, 2012/11.
- Brüderl, Josef (2010). Kausalanalyse mit Paneldaten. In: Christof Wolf and Henning Best (Eds.). *Handbuch der sozialwissenschaftlichen Datenanalyse*. Wiesbaden: VS Verlag für Sozialwissenschaften, 963-994.
- Callegaro, Mario and Charles DiSogra (2008): Computing Response Metrics for Online Panels, in: *Public Opinion Quarterly* 72, 1008-1032.
- DeBell, Matthew, Jon A. Krosnick, Arthur Lupia, and Caroline Roberts (2009): User's Guide to the Advance Release of the 2008-2009 ANES Panel Study. Palo Alto, CA and Ann Arbor, MI: Stanford University and University of Michigan.
- Deming, Edwards W. and Frederick F. Stephan (1940): On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. In: *The Annals of Mathematical Statistics* 11: 427-444.
- Downes-Le Guin, Theo, Joanne Mechling, and Reg Baker (2006): Great Results from Ambiguous Sources: Cleaning Internet Panel Data. Paper presented at: ESOMAR World Research Conference: Panel Research 2006 in Barcelona, 27-29 November.
- Gräf, Lorenz and Martin Heidingsfelder (1999): Bessere Datenqualität bei WWW-Umfragen – Erfahrungen aus einem Methodenexperiment mit dem Internet-Rogator. in: Bernad Batinic, Andreas Werner, Lorenz Gräf, and Wolfgang Bandilla (Eds.). *Online Research. Methoden, Anwendungen und Ergebnisse*. Göttingen: Hogrefe, 113-126.
- Green, Donald P., Soo Yeon Kim, and David H. Yoon (2001): Dirty Pool. In: *International Organization* 55, S. 441-468.
- Joinson, Adam (1999): Social Desirability, Anonymity, and Internet-Based Questionnaires. In: *Behavior Research Methods, Instruments, & Computers* 31, 433-438.
- Krosnick, Jon A. (1991): Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. In: *Applied Cognitive Psychology* 5, 213-236.
- Krosnick, Jon A. (1999): Survey research. In: *Annual Review of Psychology* 50, 537-567.
- Mayerl, Jochen and Dieter Urban (2008): Antwortreaktionszeiten in Survey-Analysen. Messung, Auswertung und Anwendungen. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mühlenfeld, Hans U. (2004): Der Mensch in der Online-Kommunikation: Zum Einfluss webbasierter, audiovisueller Fernkommunikation auf das Verhalten von Befragten. Wiesbaden: Deutscher Universitäts Verlag.
- Plischke, Thomas (2014): Wann Wähler entscheiden. Abläufe von Entscheidungsprozessen und der Zeitpunkt der Wahlentscheidung. Nomos: Baden-Baden.
- Roßmann, Joss (2010): Data Quality in Web Surveys of the German Longitudinal Election Study 2009. Paper presented at: 3rd ECPR Graduate Conference. Dublin.
- Schmitt-Beck, Rüdiger, Hans Rattinger, Sigrid Roßteutscher, and Bernhard Weißels (2010): Die deutsche Wahlforschung und die German Lognitudinal Election Study (GLES). In: Frank Faulbaum and Christof Wolf (Eds.). *Gesellschaftliche Entwicklungen im Spiegel der empirischen Sozialforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften, 141-172.
- Steinbrecher, Markus, Joss Roßmann, and Michael Bergmann (2013): Das Wahlkampfpanel der German Longitudinal Election Study 2009. Konzeption, Durchführung, Aufbereitung und Archivierung, GESIS Technical Reports 2013/17.
- Taddicken, Monika (2009): Die Bedeutung von Methodeneffekten der Online-Befragung: Zusammenhänge zwischen computervermittelter Kommunikation und erreichbarer Datengüte. In: Nikolaus Jakob; Harald Schoen and Thomas Zerback (Eds.). *Sozialforschung im Internet*. Wiesbaden: VS Verlag für Sozialwissenschaften, 91-107.
- The American Association for Public Opinion Research (Ed.) (2011): Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th volume. o.O.

Links

Homepage of the GLES: <http://www.gles.eu>

German Society for Electoral Studies e. V. (DGfW): <http://www.dgfw.info>

GESIS – Leibniz-Institute for the Social Sciences: <http://www.gesis.org/gles>

Goethe University Frankfurt am Main: <http://www.uni-frankfurt.de>

University of Mannheim: <http://www.uni-mannheim.de>

Social Science Research Center Berlin (WZB): <http://www.wzb.eu>

Picture credits

Wave 2, 4, 6:

Picture: Horst Seehofer, Ralf Roletschek, Lizenz: CC BY-SA 3.0, via Wikimedia Commons

Picture: Christian Lindner, Martin Kraft, Lizenz: CC BY-SA 3.0, via Wikimedia Commons

Picture: Katrin Göring-Eckardt, ©Laurence Chaperon

Picture: Frauke Petry, <http://www.frauke-petry.com>

Picture: Sahra Wagenknecht, ©DiG/Trialon

Picture: Martin Schulz, ©Susie Knoll

Wave 3, 5, 7:

Picture: Sigmar Gabriel, ©Dominik Butzmann

Picture: Angela Merkel, ©Laurence Chaperon

Picture: Winfried Kretschmann, ©Staatsministerium Baden-Württemberg

Picture: Wolfgang Schäuble, ©Laurence Chaperon

Picture: Katja Kipping, ©Fraktion Die Linke im Deutschen Bundestag

Picture: Björn Höcke, ©AfD-Fraktion im Thüringer Landtag

The study description was authored by the University of Mannheim and GESIS - Leibniz - Institute for Social Sciences.

Mannheim/Cologne, July 2019