gesis Leibniz-Institut
für Sozialwissenschaften

German Longitudinal
Election Study

GLES German Longitudinal Election Study

GLES
Short-Term Campaign Panel 2013

ZA5704, Version 3.2.0

Study description

Contents

1. <u>Preliminary notes</u>

Please note: Working with GLES data

This publication and the corresponding data set are data of the German Longitudinal Election Study (GLES), which are 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. Mistakes will immediately be announced via the GLES mailing list and are documented in the data catalogue's errata list (www.gesis.org/dbk).

If you discover an error whilst 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 data set 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 data set). 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.

<u>Contact</u>

GESIS - Leibniz - Institute for Social Sciences

P.O. Box 122155

68072 Mannheim

E-mail: gles@glesis.org

Citation of GLES data

Please include following citation in your publication with GLES data:

> Rattinger, Hans; Roßteutscher, Sigrid; Schmitt-Beck, Rüdiger; Weßels, Bernhard; Wolf, Christof; Plischke, Thomas; Wiegand, Elena (2016): Short-term Campaign Panel 2013 (GLES). GESIS Data Archive, Cologne. ZA5704 Datafile Version 3.2.0, doi: 10.4232/1.12561.

2. <u>Study characteristics</u>

Study no.

ZA5704 (Version 3.2.0); doi: 10.4232/1.12561

Data set version

GLES Version: 3.2.0; File: ZA5704_v3-2-0.sav, ZA5704_v3-2-0.dta
GLES Version (english): 3.2.0; File: ZA5704_v3-2-0_en.sav, ZA5704_v3-2-0_en.dta

Title

German Longitudinal Election Study, Component 3: Campaign Panel

Principal investigators

Prof. Dr. Hans Rattinger	(University of Mannheim)
Prof. Dr. Sigrid Roßteutscher	(Goethe University Frankfurt am Main)
Prof. Dr. Rüdiger Schmitt-Beck	(University of Mannheim)
Prof. Dr. Bernhard Weßels	(Social Science Research Center Berlin)
Prof. Dr. Christof Wolf	(GESIS – Leibniz-Institute for the Social Sciences)

Date of collection

06/20/2013- 10/04/2013

> Wave 1 (06/20-07/07/2013)
> Wave 2 (07/18-07/28/2013)
> Wave 3 (08/01-08/11/2013)
> Wave 4 (08/15-08/25/2013)
> Wave 5 (09/02-09/12/2013)
> Wave 6 (09/16-09/21/2013)
> Wave 7 (09/24-10/04/2013)

Survey method

Online survey with standardized questionnaire

Survey software

GlobalPark, EFS Umfragecenter Version 7.1

Survey institute

The organization and implementation of the data collection process was done by the Bamberg Center for Empirical Studies (Bamberg Centrum für Empirische Studien, BACES) at the University of Bamberg. Furthermore, BACES acted as intermediary to the online access panel provider Respondi AG, carried out the programming of the questionnaires and hosted the web surveys.

Funding agency

German Research Foundation (Deutsche Forschungsgemeinschaft e. V., DFG)

3. Conception and implementation of the Campaign Panel 2013

3.1. Study design

The campaign panel of the German Longitudinal Election Study enables to analyze individual changes of political attitudes and behavior patterns during the election campaign. For this purpose, a steady selection of citizens was interviewed during the election campaign up to seven times - six times before the election and one time after the election had taken place.

In addition to the Campaign Panel 2009, control groups are implemented into the Campaign Panel 2013 design, i.e. three independent samples were drawn simultaneously to the third, fifth and seventh interview of the Campaign Panel. The samples included round about 1200 people who were interviewed with an almost identical[1] questionnaire (see Figure 1). By comparing panel participants and control group participants, it can be analyzed whether and to which extent the repeated interviewing of the same participants is influential to the measuring of political attitudes and behavior patterns. The present study description only refers to the seven waves of the Campaign Panel. The documentation and the data of the control group were published separately under the ZA-numbers 5753, 5754 and 5755.

Figure 1: Waves and control groups in the Campaign Panel 2013
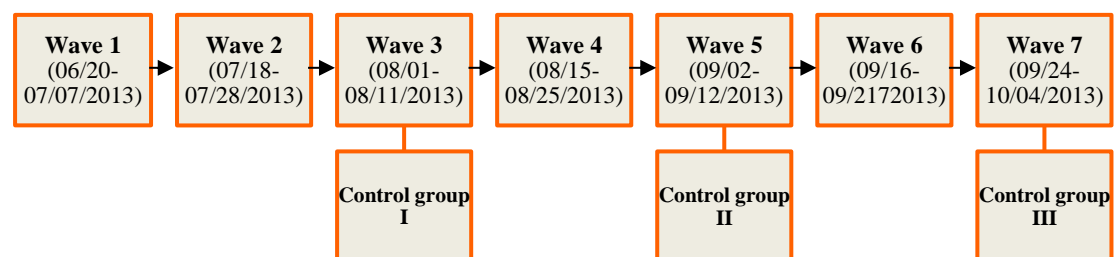
Table 1 illustrates the main focus of the seven interviews. The "core questionnaire" includes those questions which are requested in unchanged/original form in every single interview, no matter in which wave it takes place. This core questionnaire represents about 35% of the questionnaire of one wave and primarily includes questions concerning the voting and communication behavior. Additionally, every wave focuses on different features, e.g. questions concerning strategic voting, questions concerning the candidates or questions which deal with Europe and the European debt crisis.

---

[1] The battery of questions concerning socio-demographics were divided into three blocks and assigned to the waves (see Table 1). The first block contained questions about the household, the family status, religiosity, and also about the profession and vocational training. The second block dealt with the respondents' migrant backgrounds and the third collected information about the profession and the vocational training of the respondents' partners. The three control groups each only received the questions which were included in the first set of questions.

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

Contents	Wave 1 (06/20-07/07)	Wave 2 (07/18-07/28)	Wave 3 (08/01-08/11)	Wave 4 (08/15-08/25)	Wave 5 (09/02-09/12)	Wave 6 (09/16-09/21)	Wave 7 (09/24-10/04)
Core questionnaire	X	X	X	X	X	X	X
Voting behavior, retrospective	X						
Ideology, parties	X		X		X		X
Issues, parties		X		X		X	
Candidates		X		X	X		X
Strategic voting		X		X		X	
Europe			X				X
Economy	X			X			X
State election in Bavaria		X	[X]	X	X	X	[X]
State election in Hessen		X	[X]		X	X	X
Justice, in general	X						
Justice, income			X		X		X
Justice, taxes		X		X		X	
TV debate, prospective				X			
TV debate, retrospective					X	X	X
Socio-demographics, ego*	X						
Socio-demographics, migrant background		X	[X]	[X]			
Socio-demographics, partner					X	[X]	[X]

[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. Control group participants did not receive these module questions.

*participants of the control groups at the third, fifth and seventh panel wave were asked for their socio-demographics (see footnote 1).

3.2. Sampling

3.2.1 Population - the online access panel of the Respondi AG

The target population of the Campaign Panel comprises all German citizens who were eligible to vote in the 2013 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 who were entitled to vote. In 2013, this panel included 96,445 active German users. According to the definition of

Respondi, active users have completed the master data questionnaire after (double-opt-in) registration[2] and have participated in at least one survey within the last twelve months.

Table 2: Distribution of the socio-demographic characteristics in the Respondi access-panel

	Percentage
Sex	
Female	60.3
Male	39.7
Education	
Low (i.e., no graduation or graduation after 8 or 9 years of schooling) ("Hauptschulabschluss", "Volksschulabschluss")	14.9
Intermediate (i.e., secondary qualification, graduation after 10 years of schooling) ("Realschulabschluss")	36.7
High (i.e. Abitur, advanced technical college certificate, studies)	47.2
Age group	
18-29 years	32.6
30-39 years	23.3
40-49 years	21.6
50-59 years	14.8
60 years and older	7.6

100% missing: no answer

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.

---

[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.

Table 3: Ways of recruitment at Respondi

	Percentage
Mingle Trend Blog	2.6
Via online advertisements (of that 50% advertised topic-related surveys; 50% direct publicity for the panel, including affiliate marketing)	29.8
Self-entries	13.5
Facebook fan page	26.2
Search engines	21.1
Recommendations (tell-a-friend advertising)	2.1
Cooperations	4.6
Acquisition by telephone (CATI)	0.1

Respondi indicates to pursue efficient quality management. The participants' responsiveness in the Respondi panel is continuously evaluated by means of an internal evaluation system. According to Respondi's reasons for the removal of members are e.g. lacking participation over a period of twelve months, as well as deliberate submittance of false information in various surveys and double registrations.

Quality management as well as just a moderate number of requests for participation is to help avoid undesirable effects like the distortion of the panel (panel attrition) or professionalization of the members. Average members participate for 18 months in the panel. Within one year, about 36% of the panel members are removed from the data base due to quality control measures and panel attrition.

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 finished 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 2013 Campaign Panel the participants received a higher amount, see chapter 3.6). As soon as the participants have collected 1500 mingle points, the amount can be received in cash, or vouchers and it can be donated. For panel care, additional raffles are conducted periodically amongst the members.

3.2.2 Selection method and quota

The participants of the Campaign Panel 2013 consist of two subgroups:

- A sample of 4,226 people which was drawn based on a quota sampling.
- A group of 1,030 people who had already participated in the Campaign Panel 2009 and who were not included in predefined quotas for their re-invitation to the Campaign Panel 2013. By interviewing some participants of the 2009 panel again, the data sets of the Campaign Panel 2009 and the Campaign Panel 2013 can be matched (see section 7.2). Therefore, a data set with 14 panel waves is created, which includes the short-term as well as the long-term perspective.

Sampling quotas of the first subgroup are based on the categories sex, 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 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 a sufficient representation of particular combinations. For instance, it was very challenging to motivate younger males with low educational status to participate.

Independently, all participants of the Campaign Panel 2009 who had participated in at least four waves in 2009 and were still active Respondi members in 2013 were re-invited. A total of 1,527 people met those criteria. 1,011 people accepted the invitation to take part in the first survey and therefore qualified themselves as participants of the Campaign Panel 2013. Furthermore, 19 additional participants who took part in the Campaign Panel 2009, but did not meet the criteria were found. Initially, it was planned to discourage them from taking part with the help of a blacklist which did not prove to be working. As those participants were not identified until the second version of the dataset, quoting still includes these 19 participants. However, a recalculation showed that the derivations of the distribution are only marginal.

Table 4: Quoting in the Campaign Panel 2013

Quotation features	Target distribution in percent	Actual distribution* in percent	
		N=4245 (without participants of CP 2009)	N=5256 (with participants of CP 2009)
Sex			
Female	50.0	51.4	51.4
Male	50.0	48.6	48.6
Education			
Low (i.e. no graduation or graduation after 8 or 9 years of schooling) ("Hauptschule", "Volksschule")	33.3	35.2	30.8
Intermediate (i.e. secondary qualification) ("Realschule", "mittlere Reife")	33.4	32.1	34.3
High (i.e. Abitur, advanced technical college certificate, studies)	33.3	32.3	35.0
Age group			
18-29 years	20.0	18.5	17.5
30-39 years	20.0	17.8	18.0
40-49 years	20.0	21.1	22.1
50-59 years	20.0	21.8	19.8
60 years and older	20.0	20.7	19.8

* Actual distribution in the dataset without the participants of the Campaign panel 2009 who were not considered in quoting.

A typical issue arising in panel survey occurred from second panel wave two up to last wave: panel attrition. Not every first wave participant also took part in the following waves (see section 3.7). As can be seen in Table 5, panel attrition only had a minimal influence on the distribution of quota features. The proportion of male participants increased about one percentage point in the process of the Campaign Panel. The proportion of people with low educational status decreased about 1.5 percentage points in favor of the highly educated people. The highest systematic non-response rate can be registered in relation to the five age groups: As the proportion of the youngest participants decreased about nearly 3 percentage points during the panel, the proportion of the oldest age group increased about 2 percentage points.

However, the dimension of the systematic bias seems a bit bigger if the mutual distribution of the quota features is considered (which is not included in the table). Through panel attrition, male 18-29 year olds with lower education were affected most: In this group, only 35% of the first wave participants also participated in wave six. However, other groups were hardly affected at all: At least 88% of females in the age group over sixty with higher education participated in every wave.

Table 5: Distribution of gender, education and age, separated by wave, in percent (with participants of CP 2009)

Wave		W1	W2	W3	W4	W5	W6	W7
Sex	Male	48.6	48.8	49.0	49.2	49.3	49.6	49.5
	Female	51.4	51.2	51.0	50.8	50.7	50.4	50.5
Education	Low	30.8	29.4	29.3	29.1	28.6	28.7	28.3
	Intermediate	34.3	34.9	34.9	35.2	34.8	34.7	34.8
	High	35.0	35.7	35.8	35.7	36.6	36.7	36.9
	18-29 years	17.5	15.9	15.2	15.0	15.0	14.5	14.7
	30-39 years	18.0	17.4	16.9	17.1	16.7	16.7	16.8
Age Group	40-49 years	22.1	22.1	22.1	22.2	22.4	22.4	22.4
	50-59 years	22.7	23.5	23.9	23.8	23.9	24.2	24.1
	60 years >	19.8	21.1	21.9	22.0	22.0	22.2	22.1
N		**5256**	**4598**	**4432**	**4355**	**4257**	**4112**	**4231**

3.2.3 About the handling of lacking representativeness in the Campaign Panel

The participants of the Campaign Panel do not represent the people eligible to vote. Mostly young, well-educated and internet-orientated people are present in online access panels. This group differs from the total population by the collected opinions and behavior patterns. Therefore, it is not allowed to generalize proportions which are calculated based on Campaign Panel data.

The decisive advantage of panel surveys is their potential to analyze causal mechanisms.[3] One cannot rule out that lacking representativeness also leads to distortion to a certain extent, even when using panel causal analyses. Nevertheless, it has to be assumed that these distortions are considerably lower than in causal analyses with representative control group data whose effect estimation values are significantly influenced by the issue of "unobserved heterogeneity" (also "omitted variable bias") (cf. Greene et al. 2001, Brüderl 2010). Based on panel data and so-called fixed effects panel models, it is possible to eliminate unobserved heterogeneity to a large share (see Brüderl 2010). For the purpose of causal analyses, fixed-effects regressions, which are based on non-representative panel data, normally provide less distorted effect estimation values than common control group regression techniques that are based on representative data.

3.3. Invitation of the participants

At the beginning of the first wave, only those persons were invited who also participated in the Campaign Panel 2009. Based on the response rate, it was possible to calculate how many people should be involved in the new sample. Six days later, the first "new" participants were invited to take part in the Campaign Panel 2013. 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 6).

Table 6: Recruitment of participants in wave 1

Date	Group	Total	Sex male	Sex female	Age 18-29	Age 30-39	Age 40-49	Age 50-59	Age 60+	Education low	Education intermediate	Education high
06/20	1: Participants CP 2009 (E)	1527	718	809	260	305	412	336	214	183	641	703
06/22	1: Participants CP 2009 (R)	883	415	468	150	177	238	194	124	106	371	406
06/23	2 (E)	3922	1961	1961	942	745	745	745	745	1530	1373	1019
06/25	3 (E)	8000	3600	4400	4320	480	800	640	1760	2800	2400	2800
06/26	1 (R)	594	279	315	101	119	160	131	83	71	250	273
06/26	2 (R)	926	417	509	630	74	130	92		926		
07/01	2 (R)	2084	1042	1042	584	500	500	500		813	729	542
07/01	4 (E)	3000	1500	1500		1080	990	930			1800	1200
07/01	3 (R)	834	417	417	667				167	834		
07/03	4 (R)	477	239	238			239	238				477
07/06	1, 2, 3, 4 (R)	512	307	205	512					512		
Total Invitations		**16449**										

E: Invitation; R: Reminder email ; CP: Campaign Panel

---

[3] This paragraph is strongly oriented towards the work of Plischke (2014: 198-199)

From the second wave on, only those persons were re-invited who had fully completed the first survey (wave 1 drop-outs were not invited again). The number of forwarded invitations decreased progressively as in between the waves a small number of people withdrew from the Respondi panel and therefore was not available for further surveys. From the third wave on, participants were split into two equal groups and were invited separately in order to prevent server overload. In addition to reminding emails for people who did not fill out the questionnaire yet, the people who began the questionnaire but did not complete it received a reminder one day later (see Table 7).

Table 7: Invitations and reminders in waves 2 to 7

	Wave 2		Wave 3		Wave 4		Wave 5		Wave 6		Wave 7	
	Date	**N**	**Date**	**N**	**Date**	**N**	**Date**	**N**	**Date**	**N**	**Date**	**N**
Invitation	07/18	5242	08/01	2870	08/15	2844	09/02	2864	09/16	2923	09/24	2799
Invitation			08/01	2343	08/15	2360	09/02	2324	09/16	2257	09/24	2368
First reminder for not yet approached panelists	07/21	1367	08/04	1556	08/18	1703	09/05	1507	09/18	1653	09/27	1386
First reminder for interrupters	07/22	244	08/05	206	08/19	166	09/06	186	09/19	136	09/27	232
Second reminder for not yet approached panelists	07/25	735	08/07	955	08/21	1006	09/08	1057	09/20	1196	09/30	1068
Second reminder for interrupters	07/26	179	08/08	174	08/22	125	09/09	150	09/21	114	10/01	120

Whilst formulating invitation and reminding emails, importance was attached to highlight the special characteristics of the study. The survey was presented to the panelists 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, the standard text in emails was changed and the GLES logo was included. The layout and exact wording of the invitations and the homepage of every wave can be reviewed in the documents "Screen Views" which are enclosed.

3.4. 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 email account of the target person can participate instead. In order to prevent incorrect participation, the invitational email 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: sex, the year of birth, the month of birth and the first letter of the place of birth. In the first wave, the data given by the participants in order to verify their access to the survey were compared to their

master data. [4]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 2). 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 and that the problem was to be solved and they would be re-invited later. Before the participants were rejected, they were able to give reason why the verification code did not correspond to the statements given in the first wave (see "Reason for wrong verification code" in Figure 2). The target persons were re-invited to the survey in the course of reminder emails for interrupters and could try again to enter their correct verification code.

---

[4] Master data are personal data which the participants of the Campaign Panel initially provide about themselves when they join the Respondi Online Access Panel. They are periodically asked to update their data. Those data were collected by Respondi before the Campaign Panel and therefore enable to control invitations. The by Respondi already collected master data (see 4.9) were mostly re-collected as there is only little information on the datas' topicality.

Figure 2: Verification process
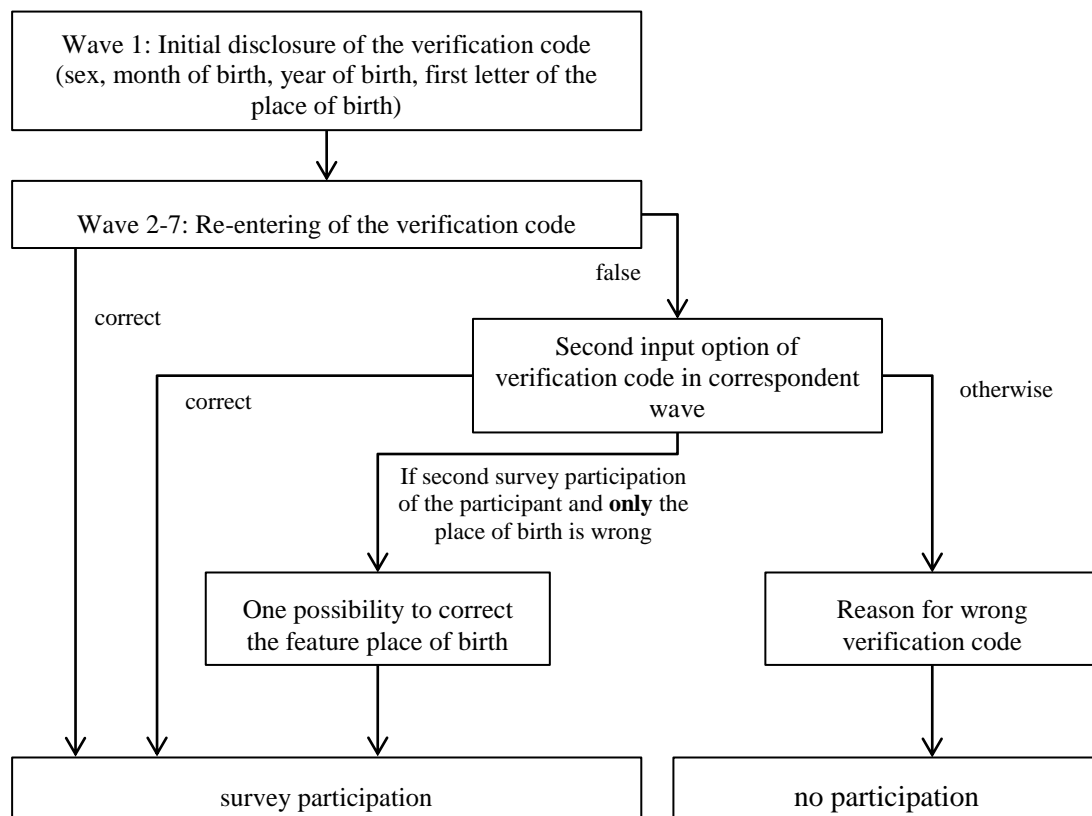
A one-off exception in the verification process was granted with the feature "first letter of place of birth". As this is an open statement, there is the possibility that a typing error occurred in the first wave. Therefore participants had the single possibility to correct the in wave 1 given information if only the letter was incorrect. 77 people corrected the first letter of their place of birth in this way ($n_{wave2} = 72$; $n_{wave3} = 2$; $n_{wave4} = 2$; $n_{wave5} = 0$; $n_{wave6} = 0$; $n_{wave7} = 1$).

3.5. Interview duration

Median completion times in all panel waves were about 20 minutes each on average. With a median time of 24:37 minutes, the first wave took the most amount of time, while the fifth wave with 17:37 minutes took the least (see Table 8).

Apparently, 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 where 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 are described in section 6.2.

Table 8: Interview length

	N	Minimal duration of the survey	Maximum duration of the survey	Arithmetic mean	Median
Wave 1	5002	00:03:33	03:08:55	00:27:38	00:24:37
Wave 2	4236	00:02:18	02:09:10	00:24:31	00:21:07
Wave 3	4143	00:01:54	02:33:03	00:21:48	00:19:18
Wave 4	4089	00:01:55	02:47:13	00:21:34	00:18:52
Wave 5	4041	00:01:34	05:23:54	00:20:19	00:17:37
Wave 6	3906	00:01:39	02:17:08	00:20:46	00:18:19
Wave 7	3926	00:01:43	02:41:24	00:24:14	00:21:21

The duration is only shown for those participants who didn`t interrupt the survey. According to the survey software, one person spent 105:44:24 on the second wave questionnaire without interruption. This person was not included in the presentation of the completion times in wave 2. Presentation: hh:mm:ss

3.6. Incentives

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 accordance 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. Violating the agreement, Respondi only paid out 100 Mingle Points for the participation in the first wave. As an offset, participants of the second wave received a singular payment of 200 Mingle Points (see Table 9).

Furthermore, the participants were told that a frequent participation in the Campaign Panel would enable them to receive up to 350 additional Mingle Points (see Table 10). Thus the participants of the Campaign Panel were able to receive up to a maximum of 1400 Mingle Points (7*150 + 350) when they took part in the Campaign panel seven times.

Table 9: Mingle-Points for participation in single waves

Participation in…	Mingle-Points
Wave 1	100
Wave 2	200
Wave 3	150
Wave 4	150
Wave 5	150
Wave 6	150
Wave 7	150

Table 10: Additional Mingle-Points for at least four participations

Participation in … waves	Mingle-Points
4	120
5	170
6	220
7	350

3.7. Participation statistics

3.7.1 Frequency of participation

3,487 persons (66%) from 5,256 persons who completed the first wave questionnaire of the Campaign panel took part in all seven surveys (see Table 11). Only about 8 percent of the participants solely participated in the first wave. The variable "teilnahme" provides detailed information in the dataset about the individual development of participation in all seven waves (also see Chapter 4.3).

Table 11: Number of completed surveys

Participation in … waves	absolute	%
1 participation	413	7.9
2 participations	235	4.5
3 participations	185	3.5
4 participations	185	3.5
5 participations	244	4.6
6 participations	507	9.6
7 participations	3487	66.3
Total	**5256**	**100.0**

Deviation from 100 percent due to rounding-off of the numbers

3.7.2 Distribution of participations over field time

In the first wave, the field time was a week longer than in the other waves, as not all invitations were sent on the same day. The participations are spread on the whole length of the field time. 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, in which on average nearly 78% of all surveys were completed (see Figure 3). On the fourth field day, in most waves a light increase of participations can be noticed, which is due to reminder emails.

Figure 3: Distribution of the finished interviews during the field time (in percent of the number of all finished interviews per wave)
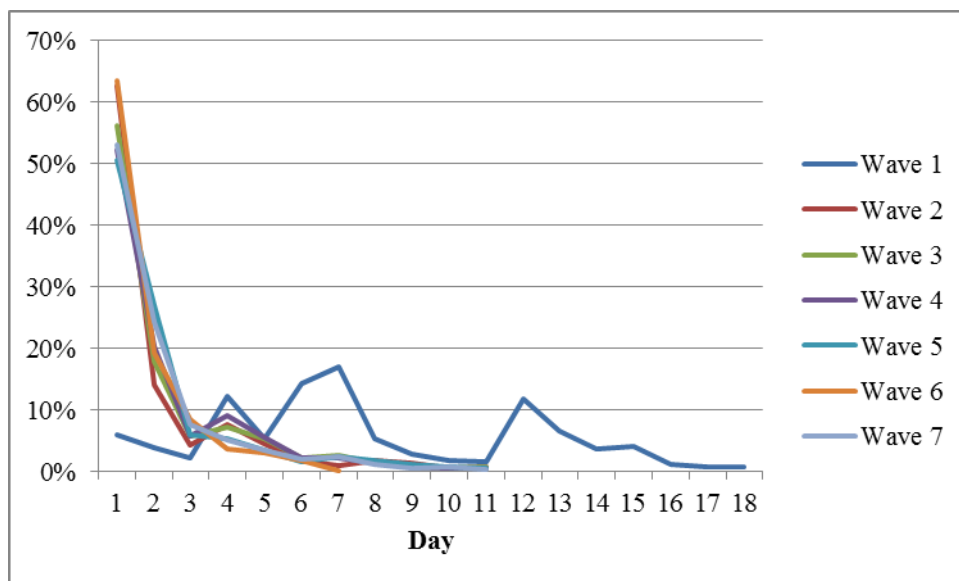
3.7.3   Participation rates

Systematic unit non-response is a source of error in surveys which can substantially reduce the data's quality.[5] When the collected variables are correlated with the panelists' 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 panelists' participation behavior ensures the comparability of these patterns beyond other studies. For the GLES Campaign Panel, indicators of the participation behavior are stated according to the standards of the American Association for Public Opinion Research, which were 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 online access panel, it is not reasonable to refer to response rates (see The American Association for Public Opinion Research 2011).Therefore, AAPOR suggests to state 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

---

[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).

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). Respondi members who quit the access panel completely during active field work and were therefore no longer available for further surveys are defined as others (O). These cases are included in the calculation of the participation rate as they had possibly been able to give an interview.

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 approval of the owner. Since participants who entered a wrong verification code were invited again several times in order to enter a correct code (see section 3.4), 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 false 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 as others with unclear eligibility (UO). The calculation of the participation rates is as follows:

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

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.

$$\text{Breakoff Rate 1 (BR1)} = \frac{R1_{\text{Breakoff}}}{(I + P) + (R1_{\text{Breakoff}})})$$

$$\text{Breakoff Rate 2 (BR2)} = \frac{(R1_{\text{Breakoff}} + R2_{\text{Breakoff}})}{(I + P) + (R1_{\text{Breakoff}} + R2_{\text{Breakoff}})}$$

The gross sample for the first wave of the Campaign Panel was drawn from the online access panel by Respondi. The chosen panelists were invited to participate by Respondi. The gross sample for the second to seventh wave of the Campaign Panel formed the panelists who completed the first panel wave. 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 12: Participation rate statistics of the Campaign Panel

Code	Description	W1	W2	W3	W4	W5	W6	W7
I + P	**Complete and partially complete interviews**	5256	4517	4332	4285	4205	4062	4192
R1$_{Breakoff}$	**Break-off after verification process**	271	81	100	70	52	50	39
R2$_{Breakoff}$	**Break-off before verification process**	/	34	51	32	34	27	21
NC	**No Contact**	/	596	729	811	890	1037	912
O	**Other**	/	14	43	52	68	76	89
UH	**Ambiguous eligibility**	8841	/	/	/	/	/	/
UO	**Others with ambigiuous eligibility**	/	14	1	6	7	4	3
NE	**Met quota or screened out in wave 1**	2081	/	/	/	/	/	/
Number of interview invitations sent by Respondi		16449	5242	5213	5204	5188	5180	5167
PR	**Participation Rate in %**	36.6	85.9	82.4	81.5	80.0	77.3	79.8
BR1	**Breakoff Rate 1 in %**	4.9	1.8	2.3	1.6	1.2	1.2	0.9
BR2	**Breakoff Rate 2 in %**	/	2.5	3.4	2.3	2.0	1.9	1.4

The variable "kpX_compl" in the dataset includes the participation status of the panelists according to AAPOR standards. With the help of those variables, the calculation of nonresponse rates is to be simplified in research works. Table 13 provides an overview of the meaning of the coding.

Table 13: Manifestation of the variables kpX_compl

Code	Value label	Explanation
1 Eligible persons, completed questionnaire		
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 data set). In the waves 2-7, 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.
2 Eligible persons, no interview		
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 7).
2.20	No contact	People who did not respond to the invitation (only in waves 2 to 7).
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 7).
3 Unknown eligibility, no interview		
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 7).
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

4 Variables in the data set

4.1 Preliminary note

The data set of the Campaign Panel includes different types of variables:

- *GESIS archive variables* describe the data set 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 6).
- *Weighting variables* include cross-section and panel weights (see section 4.6).
- *Attitudinal and behavioral 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 and participations, as well as the equipping of the participants (i.e. browser version, JavaScript version).

Table 14 gives a review on the total number of the variables in the individual waves. The following subchapters provide a description of the different variable types.

Table 14: Number of variables in total and detailed for waves

	W1	W2	W3	W4	W5	W6	W7	All waves	Total
GESIS archive variables	1	1	1	1	1	1	1	6	13
Administrative variables	6	6	6	6	6	6	6	4	46
Contextual variables	/	/	/	/	/	/	/	9	9
Data quality variables	17	13	13	13	13	13	15	1	98
Weighting variables	/	/	/	/	/	/	/	8	8
Attitudinal and behavioral variables	216	263	303	298	338	325	330	76	2149
Paradata	8	7	7	7	7	7	7	13	63
Time variables	280	222	252	268	312	308	356	/	1998
Total	528	512	582	593	677	660	715	120	4384

4.2 GESIS archive variables

Table 15: Archive variables of GESIS

Variable	Explanation
study	provides the (ZA-) study number of the dataset in four-digit format.
version	Data set version, starting with 1.0.0
year	Survey year of the study
field	Field time of the study. Because of the panel design, the dataset additionally includes the variables kpX_field which reflect the field time of the correspondent wave.
glescomp	Allocation to one of the eleven components of GLES. Here: component 3 for the Campaign Panel
survey	Differentiation between several surveys which were implemented within one GLES component. In the Campaign panel 2013, these are the cumulated data set of all seven waves (coded as 1) and the control groups of wave 3, 5 and 7 (coded as 2 to 4).

Explanation: "X" represents the respective wave.

4.3 Administrative variables

Table 16: Administrative variables

Variable	Explanation
wkp2009	Participation in Campaign Panel 2009 (0 = no; 1 = yes)
mut09	Marking of cases which were identified as panel mutants in 2009.
kpX_compl	Participation status of the panelists 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.

4.4 Contextual variables: Allocation of constituencies

The only contextual variable included in the Campaign Panel is the constituency of the participants. The assignment of respondents to a constituency was based on the postal code of their main residence which they stated in the survey. Therefore a correspondence table by the Federal Returning Officer was accessible.

When using 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 - in 4.438 cases (84.4 %) - the constituency was saved in the variable "elecdist". If several allocations were possible - 694 cases (13.2%) - the up to five possible constituencies are included in the variables "elecdist1" to "elecdist5". The remaining 124 participants (2.4%) did not provide a valid postal code or refused to give the relevant information.

In addition to that, a comparison between the mentioned federal state and the provided postal code was made. In 15 cases, the postal code did not correspond to the entered state. These were marked with the variable "kpx_info".

4.5 Data quality variables

The published data set contains variables which are used to enable the users to assess the quality of the data. It is planned to include additional indicators for the quality of the data in succeeding versions of the data set. Detailed explanations on the quality indicators are to be found in chapter 6.

Table 17: Data quality variables

Variable	Explanation	W1	W2	W3	W4	W5	W6	W7
kpX_speederindex	Considers the response time of one participant in relation to those of all other participants. Also see section 6.2 "Quick Responses".	X	X	X	X	X	X	X
kpX_dev	Index indicating deviating answering patterns	X	X	X	X	X	X	X
kpX_itnrp	Proportion of answers with "no answer" (item non - response))	X	X	X	X	X	X	X
kpX_strl	Straightlining index	X	X	X	X	X	X	X
kpX_mtime	Average response time on one site of the survey in seconds	X	X	X	X	X	X	X
kpX_mtimex	Average response time on one site oft he survey in seconds (adjusted, to be able to compare the values of all waves)	X	X	X	X	X	X	X
kpX_dokn	Proportion of answers with "don't know"	X	X	X	X	X	X	X
kpX_qual	Overall quality indicator	X	X	X	X	X	X	X

kpX_4210	Attention during survey participation	X	X	X	X	X	X	X
kpX_4220	Accuracy during survey participation	X	X	X	X	X	X	X
kpX_4230, s	Reason for survey participation	X						X
kpX_4240	Evaluation of the survey	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						X
kpX_4270s	Comment field in which the respondents were able to state their remarks on the survey	X	X	X	X	X	X	X
kpX_	Bogus Items	050q; 1571q	2900q	050q	040q	2910q	2940q	050q

Explanation: "X" represents the respective wave.

4.6 Weighting variables

4.6.1 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 behavioral variables.

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 0.05[6]. 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 process[7]. However, it was not necessary to trim the cross-sectional and panel weights in the Campaign Panel as there were no great derivations of set point and actual value.

---

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

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

Table 18: Overview about weights

Weight	Variable
Social- and regional structural weight (adjustment to MZ 2012)	wei_mz
Social- and regional structural weight (adjustment to Onliner)	wei_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 2012, as well as to the distributions of the Mikrozensus[8]. Only persons entitled to vote in private households situated in their main residence were included in the calculation of the target distribution of Mikrozensus 2012.[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 school
- 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"))

---

[8] The marginal distributions of the (N)Onliner Atlas (for people over 18 years) for the weighting process were kindly provided by TNS Infratest.

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

Table 19: Actual and target distributions of the variables which were used to calculate the weights

Characteristic		Actual distribution (in percent) N=5256		Target distribution (in percent) Mikrozensus 2012	(N)Onliner Atlas 2012
Sex					
male			48.6	48.6	52.5
female			51.4	51.4	47.6
Age group (crossed with education)					
18 up to 30 years	low		4.1	4.2	
	intermediate	17.5	5.6	5.4	22.6
	high		7.7	6.8	
30 up to 45 years	low		8.5	5.3	
	intermediate	29.8	10.9	8.3	28.9
	high		10.3	8.5	
45 up to 60 years	low		11.7	10.1	
	intermediate	33.0	11.3	10.2	30.9
	high		10.0	7.8	
60 years and older	low		6.4	22.1	
	intermediate	19.8	6.4	5.9	17.6
	high		7.0	5.3	
Education					
low			30.8		34.6
intermediate			34.3		32.1
high			35.0		33.3
Federal state					
Western German states			78.2	78.6	79.9
Newly- formed German states (incl. Berlin)			21.8	21.4	20.1

As a joint distribution of the characteristics age and education of the Mikrozensus is additionally available, it was possible to attune more differentiated to the population. One beneficial aspect of this process is that it is possible to adjust individual educational categories to every age group. As a result, one is able to counteract possible contortions caused by especially strongly or weakly represented groups. However, this is only possible for the adjustment to Mikrozensus as (N) Onliner Atlas does not provide any information about the mutual distribution of both characteristics. Correspondingly, the characteristics

were included individually in order to be able to attune them to the online population (see Table 19).

The calculation of the two weighting variables 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 in both cases. The following table provides an overview of the calculated weights.

Table 20: Overview of the weighting factors

	N	Mean	Std.Dev	Min	Max	1.Q	Median	3.Q	Max./Min.[10]
wei_mz	5256	1	0.65	0.61	3.65	0.77	0.85	0.90	5.98
wei_on	5256	1	0.19	0.71	1.67	0.85	0.96	1.12	2.35

4.6.2 Panel weights

Panel weights correct systematic distortions which arise in panel surveys due to irregular participation (panel attrition) in the panel waves. In most cases these are Propensity Score Weights. In the following section the methods of calculation, the used models and the calculated weights will be explained.

When the Propensity Score Weighting is applied, one estimates individual probabilities of each respondent to participate in the survey. In the case of a panel survey, this has to be referred to as the probability of the respondents to remain in the panel. Here, one estimates the probability to remain in the panel waves of all respondents (except for the first wave in which all respondents participated). Logistic regressions are commonly used to model the drop-out process. The estimated model provides an explanation of the drop-out process and allows estimating the probability to remain in the panel for each respondent. The probability's inverse establishes the individual weighting factor for a respondent, which can be used to control for panel attrition. In concrete terms: Respondents with a lower probability to remain in the panel receive a higher weighting, whereas respondents with a higher probability to stay receive a lower weighting. A new model is estimated for every wave of the panel, namely for the participation in the waves two, three, four, five, six and seven in the case of the Campaign Panel. Thereby the information of the last wave is used, as it must be assumed that the gained experience in the last survey is to influence the participation in the following waves. That means that the probabilities for respondents to stay in the panel for wave three are estimated with the information collected in wave two. As some characteristics are to be assumed temporally invariant, e.g. sex or the year of birth, the information of the first wave can be used in these cases.

Two problems arise when the weights are calculated. Firstly, it is possible that respondents contain missing values in the variables which were used in the model. Therefore, it is not possible to estimate the probability to stay in the panel for those participants. In such cases, they are assigned the average probability of other respondents. Secondly, it is also possible that respondents did not participate in the previous wave. In this case, one is not able to estimate the probability that the respondents will remain in the panel either. If probabilities from previous waves are available for the respondent, these are used to calculate the

---

[10] 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.

weighting factor. For example, a respondent participates in the first, second, third and fifth wave. As there is no information accessible about the respondent in wave 4, it is not possible to determine the probability of the respondent to remain in the panel for wave 5. Therefore, the probability to stay in the panel of wave three is used to calculate the weighting factor for wave five. If this is not possible because the respondent only took part in wave 1, the average probability of other respondents is imputed.

The variables which were used for modeling are displayed in Table 21. The selection is oriented towards several theoretical approaches which explain the behavior patterns of respondents in surveys. The panel weighting in GLES 2009 is also based on these approaches (cf. Blumenberg/Gummer 2013 p. 16ff). Due to the model, the questionnaires, and additional accessible information, the explanatory model was adjusted for the current Campaign Panel. The models which were calculated on this basis are given in Table 22.

Table 21: Used variables

Explanation of the variables	Encoding
Age (and age²)	18-99
Sex (male / female)	0/1
Education (low / intermediate / high)	0/1 & 0/1 & 0/1
Spouse	0/1
Labor status (Employed person / homemaker / pensioner)	0/1 & 0/1 & 0/1
Household size	1-8
East Germany	0/1
Political interest	0-4
Frequent conversations about politics	0/1
Strength of party identification	0-5
Satisfaction with democracy	0/1
Preferred chancellor	0/1
Likelihood of Voting	0/1
Political knowledge: correct assignment of the top candidates	0/1
Internet usage (intensity)	0-7
Motivation to participate: Index consisting of subjective deliberation and accuracy in answering the questions and evaluation of the survey	factor scores
Bogus Items (have been recognized correctly)	W2: 0-2, W3-7: 0/1
Big 5	factor scores
Monetary motivation: Reason for participation "Mingle Points"	0/1
Experience with surveys: number of completed surveys (0; 1-4; 5-10; 11+)	0-3
Number of participations in the Campaign Panel	1-6 (depending on the [respective] wave)
Frequency of Item Nonresponse (refusal of providing an answer)	0-122
Duration of the last interview (and duration²)	Time in seconds

Used input device (PC / smartphone / tablet) 0/1 & 0/1 & 0/1

Table 22: Logistic regression model in order to explain the participation in the waves 2-7.

	Participation in					
	Wave 2 b (se)	Wave 3 b (se)	Wave 4 b (se)	Wave 5 b (se)	Wave 6 b (se)	Wave 7 b (se)
Age	0.039 [a] (0.021)	-0.050 (0.033)	-0.000 (0.035)	0.017 (0.033)	0.039 (0.031)	0.037 (0.044)
Age²	-0.000 (0.000)	0.001* (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Education: intermediate	0.430*** (0.117)	0.131 (0.163)	-0.047 (0.187)	0.093 (0.179)	-0.183 (0.179)	0.041 (0.250)
Education: high	0.353** (0.126)	0.155 (0.175)	-0.032 (0.204)	0.564** (0.213)	-0.068 (0.198)	0.460 (0.288)
Spouse	-0.042 (0.110)	-0.316* (0.157)	0.048 (0.173)	0.206 (0.172)	-0.159 (0.168)	-0.420 (0.262)
Employment: Homemaker	0.491* (0.215)	0.118 (0.297)	0.327 (0.324)	-0.315 (0.284)	0.087 (0.282)	-0.389 (0.386)
Employment: Pensioner	0.116 (0.187)	-0.059 (0.253)	0.185 (0.306)	0.278 (0.339)	0.029 (0.288)	-0.191 (0.378)
Household size	-0.097* (0.040)	-0.088 (0.056)	-0.109 a (0.066)	-0.091 (0.067)	0.029 (0.067)	-0.053 (0.093)
East Germany	-0.032 (0.114)	0.073 (0.158)	0.253 (0.186)	-0.012 (0.179)	0.356* (0.179)	0.512 a (0.272)
Political interest	0.093 (0.059)	0.085 (0.078)	0.128 (0.091)	0.337*** (0.090)	0.076 (0.086)	0.084 (0.126)
Frequency of political conversations	-0.111 (0.106)	-0.176 (0.145)	-0.134 (0.170)	-0.432* (0.170)	-0.105 (0.165)	0.089 (0.240)
Strength of party identification	-0.029 (0.028)	-0.062 (0.039)	0.016 (0.045)	0.030 (0.045)	0.026 (0.043)	0.050 (0.063)
Satisfaction with democracy	0.141* (0.056)	-0.079 (0.077)	0.038 (0.087)	0.086 (0.087)	0.183* (0.083)	0.063 (0.123)
Preferred chancellor	0.078 (0.101)	0.185 (0.140)	-0.070 (0.162)	-0.004 (0.162)	-0.202 (0.165)	0.019 (0.235)
Likelihood of voting	-0.040 (0.139)	0.200 (0.185)	0.290 (0.208)	-0.480* (0.227)	0.395 a (0.209)	0.067 (0.317)
Political knowledge	0.198 a (0.114)	0.090 (0.150)	0.161 (0.177)	-0.267 (0.176)	0.124 (0.168)	-0.180 (0.245)
Internet usage	0.057 (0.043)	0.137* (0.056)	0.055 (0.068)	0.094 (0.065)	0.125* (0.060)	0.227** (0.074)
Motivation to participate	0.248*** (0.061)	0.162* (0.072)	0.171* (0.085)	0.309*** (0.082)	0.330*** (0.074)	0.167 (0.113)
Control questions	0.160* (0.078)	-0.159 (0.209)	-0.114 (0.248)	0.006 (0.200)	0.141 (0.219)	0.074 (0.389)
Big 5: Extraversion	-0.198*** (0.059)	-0.102 (0.078)	-0.173 a (0.092)	-0.003 (0.092)	-0.023 (0.088)	0.006 (0.130)
Big 5: Openness	-0.060 (0.067)	-0.020 (0.088)	-0.084 (0.104)	-0.023 (0.102)	-0.130 (0.098)	-0.008 (0.150)

Big 5: Conscientiousness	0.097 (0.067)	-0.016 (0.089)	-0.010 (0.106)	-0.116 (0.104)	-0.016 (0.099)	0.247 a (0.148)

Continuation of Table 22

	Participation in					
	Wave 2 b (se)	**Wave 3** b (se)	**Wave 4** b (se)	**Wave 5** b (se)	**Wave 6** b (se)	**Wave 7** b (se)
Big 5: Neuroticism	0.112 a (0.064)	-0.115 (0.086)	0.108 (0.100)	-0.114 (0.103)	0.035 (0.097)	-0.028 (0.145)
Big 5: Compatibility	0.079 (0.067)	0.006 (0.090)	0.293** (0.108)	-0.148 (0.106)	-0.008 (0.103)	0.085 (0.153)
Monetary motive	0.189 a (0.114)	0.022 (0.151)	-0.066 (0.173)	0.238 (0.181)	0.092 (0.168)	-0.210 (0.238)
Experience with surveys	0.341*** (0.058)	0.351*** (0.080)	0.185* (0.091)	0.109 (0.090)	0.100 (0.086)	0.016 (0.126)
Previous participations	---	---	1.910*** (0.201)	1.269*** (0.140)	1.130*** (0.100)	1.103*** (0.098)
Item non response	-0.038*** (0.008)	-0.007 (0.012)	-0.012 (0.009)	-0.023** (0.007)	0.000 (0.013)	-0.014 a (0.009)
Length of the last interview	-0.000** (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.001 a (0.000)	-0.000 (0.000)	-0.000 (0.000)
Length of the last interview²	0.000a (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Device: Smartphone	-0.228 (0.228)	0.215 (0.389)	0.452 (0.517)	-0.431 (0.384)	0.918 (0.572)	-0.626 (0.521)
Device: Tablet	-0.211 (0.258)	0.245 (0.438)	0.311 (0.477)	1.055 a (0.610)	0.016 (0.437)	-0.354 (0.569)
Constant	-0.321 (0.596)	1.691* (0.827)	-3.893*** (1.118)	-3.401** (1.106)	-5.267*** (1.002)	-5.147*** (1.388)
N	4810	4142	3959	3914	3823	3719
Nagelkerke R²	0.12	0.08	0.11	0.13	0.14	0.19

a $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 23 sums up the calculated panel weights. They can be used when variables of previous waves, for which has to be assumed that they are distorted due to panel attrition, are to be calculated. For every wave, there is a corresponding panel weight "wei_wX" in which X symbolizes the particular wave. With the help of multiplication, it is possible to combine panel weights and cross-section weights, in order to additionally adjust to the population of eligible people in Germany (Mikrozensus) or the online population ((N) Onliner-Atlas).

Table 23: Panel weights for the waves 2 to 7

	N	Mean	Std.Dev	Min	Max	1.Q	Median	3.Q	Max./Min.[11]
wei_w2	4598	1	0.13	0.88	2.90	0.93	0.97	1.02	3.30
wei_w3	4432	1	0.05	0.93	1.64	0.96	0.99	1.02	1.76
wei_w4	4355	1	0.07	0.95	2.19	0.97	0.99	1.00	2.31
wei_w5	4257	1	0.10	0.94	4.13	0.97	0.99	1.00	4.39
wei_w6	4112	1	0.09	0.94	3.11	0.97	0.98	1.00	3.31
wei_w7	4231	1	0.01	0.97	2.78	0.98	0.99	1.00	2.87

4.7 Attitudinal and behavioral 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 wave. [12] Subsequently, the item number follows. For instance, the variable "Political Interest" carries the item number "010". The variable "Political Interest", 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 Campaign Panel 2009 as well as in the Campaign Panel 2013, the item number remains the same in both data sets. When the question wording or the answer scales were edited, the last digit of the item number was increased by one.

Table 24: Attitudinal and behavioral variables

Variable	Item	W1	W2	W3	W4	W5	W6	W7
kpX_screen 1	Right to Vote	X						
kpX_screen 2	Right to Vote, request	X						
kpX_2280	Gender	X	X	X	X	X	X	X
kpX_2291	Month of birth	X	X	X	X	X	X	X
kpX_2290	Year of birth	X	X	X	X	X	X	X
	Voting behavior	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_170	Intention to vote	X	X	X	X	X	X	
kpX_180	Actual turnout							X
kpX_190	Voting intention	X	X	X	X	X	X	
kpX_191	Voting of postal voters			X	X	X	X	X
kpX_200	Actual voting behavior							X
kpX_270	Reliability of voting decision	X	X	X	X	X	X	
kpX_252	Reasons for decision not to vote, closed	X	X	X	X	X	X	X
kpX_260s	Reasons for voting decision, open	X	X	X	X	X	X	
kpX_261	Reasons for voting decisions, closed			[X]	[X]	[X]	[X]	X

---

[11] See footnote 10.

[12] 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	W5	W6	W7
kpX_211	Consideration set, hypothetical	X		X			X	X
kpX_221	Consideration set	X		X			X	X
kpX_2751	Consideration set on the day of the postal vote			X	X	X	X	X
kpX_2760	Reason for absentee voting, closed			X	X	X	X	X
kpX_320	Moment of voting decision?							X
kpX_321	Moment of decision not to vote							X
kpX_331	Difficulty of participating in elections			[X]	[X]	[X]	[X]	X
kpX_330	Difficulty of participating in elections			[X]	[X]	[X]	[X]	X
kpX_340	Satisfaction with election result							X
kpX_2770	Hypothetical vote after absentee voting				X	X	X	X
kpX_280	Hypothetical voting behavior							X
kpX_290	Hypothetic vote after parliamentary elections							X
kpX_341	Election result: Winners and losers							X
	Voting behavior, retrospective	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_2780	Turnout parliamentary elections [for the Bundestag] 2009	X						
kpX_2790	Reliability memory/reminiscence of parliamentary elections 2009	X						
kpX_350a	Recall previous parliamentary elections	X						
	Ideology	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_1490	Left-right classification, parties	X		X		X		X
kpX_680	Left-right classification, top candidates		X		X	X		X
kpX_1500	Left-right self-classification	X	X	X	X	X	X	X
	Opinions about political parties and the government	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_430	Scalometer parties	X	X	X	X	X	X	X
kpX_730	Scalometer government	X	X	X	X	X	X	X
kpX_2800	Ambivalence parties, negative feelings	X		X			X	
kpX_2801	Ambivalence parties, positive feelings	X		X			X	
	Valence issues	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_840s	Most important issue	X	X	X	X	X	X	X
kpX_850	Ability to solve the most important issue	X	X	X	X	X	X	X
kpX_851	Ability to solve the most important issue, other party	X	X	X	X	X	X	X
kpX_860s	Second most important issue	X	X	X	X	X	X	X
kpX_870	Ability to solve second most important issue	X	X	X	X	X	X	X
kpX_871	Ability to solve second most important issue, other party	X	X	X	X	X	X	X
	Position issues	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_1070	Socio-economic dimension, parties		X		X		X	
kpX_1080	Socio-economic dimension, chancellor candidates		X		X	X		X

Variable	Item	W1	W2	W3	W4	W5	W6	W7
kpX_1090	Socio economic dimension, ego	X	X	X	X	X	X	X
kpX_1110	immigration opportunities for foreigners, parties		X		X		X	
kpX_1120	Immigration opportunities for foreigners, Chancellor candidates		X		X	X		X
kpX_1130	Immigration opportunities for foreigners, self	X	X	X	X	X	X	X
kpX_1270	Climate protection, parties		X		X		X	
kpX_1280	Climate protection, chancellor candidates		X		X	X		X
kpX_1290	Climate protection, ego	X	X	X	X	X	X	X
kpX_1100	Socio-economic dimension, importance		X		X		X	
kpX_1140	Immigration opportunities for foreigners, importance		X		X		X	
kpX_1300	Climate protection, importance		X		X		X	
kpX_2810	European aid, parties			X				X
kpX_2820	European aid, chancellor candidates			X		X		X
kpX_2830	European aid, ego			X		X		X
kpX_2840	European aid, importance			X				X
kpX_1210	Integration, ego			X			X	
kpX_1250	European integration, ego			X			X	
Variable	**Item**	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_1482	Foreign policy, ego			X			X	
kpX_2850	Ego positions I	X			X			
kpX_2860	Ego positions II		X			X		
kpX_2870	Ego positions III; Current topics			X			X	
	Candidates	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_650	Scalometer politicians	X	X	X	X	X	X	X
kpX_650	Scalometer politicians		X		X		X	
kpX_661	Characteristics of chancellor candidates, Merkel		X		X	X		X
kpX_661	Characteristics of chancellor candidates, Steinbrück		X		X	X		X
kpX_671	Preferred chancellor	X	X	X	X	X	X	X
kpX_3000	Assessment standards for the evaluation of candidates		X		X	X		X
	Strategic voting	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_910	Scalometer coalitions		X		X		X	X
kpX_961	Expected coalition		X		X		X	
kpX_1040	Constituency winner		X		X		X	
kpX_1050	Perception of surveys		X		X		X	
kpX_1051	Credibility of surveys		X		X		X	
kpX_3010	Probability of small parties entering the Bundestag		X		X		X	
kpX_3020	Rating of the candidates in constituency		X				X	
	Europe	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_3030	Fear, European financial crisis			X				X

Variable	Item	W1	W2	W3	W4	W5	W6	W7
kpX_3040	Affected by European financial crisis, self			X				X
kpX_3050	Affected by European financial crisis, Germany			X				X
kpX_3060	Performance of the federal government fighting the European financial crisis			X				X
kpX_3070	Performance of the Chancellor candidates, European financial crisis			X		X		X
kpX_3080	Main share of the blame European financial crisis			X				X
kpX_3090	Returning to D-Mark			X				X
	Economy	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_780	Own economic situation, current	X			X			X
kpX_760	own economic situation, retrospective	X			X			X
kpX_770	Responsibility for own economic situation	X			X			X
kpX_790	Own economic situation, prospective	X			X			X
kpX_820	general economic situation, current	X			X			X
kpX_800	general economic situation, retrospective	X			X			X
kpX_810	Responsibility for general economic situation	X			X			X
kpX_830	General economic situation, prospective	X			X			X
kpX_3110	Economic situation in Europe, current			X				X
	Fairness	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_2250	Fair social order	X						
kpX_2270	Fairness, ego	X						
kpX_3330	fairness, development	X						
kpX_3340	Importance of justice	X			X			
kpX_3350	Justice, comprehension	X						
kpX_3360	Justice [Fairness] , Merkel	X			X	X		
kpX_3360	Justice [Fairness], Steinbrück	X			X	X		
kpX_3370	Fairness, income				X	[X]		[X]
kpX_3380	Importance of fair incomes				X	X		X
kpX_3390	Fairness of income, parties				X	X		X
kpX_3400	Justice, taxes		X		[X]		[X]	
kpX_3410	Importance of fair taxes [tax rates]		X		X		X	
kpX_3420	Fair taxes, parties		X		X		X	
kpX_2700	Old vs. young, statements			X				
	Media use	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_1620	Most important information source, habitually	X						
kpX_1631	Average Internet usage	X						
kpX_1640	Political use of the Internet, habitually	X						
kpX_3540	Duration of Internet usage, habitually	X						

Variable	Item	W1	W2	W3	W4	W5	W6	W7
kpX_1651s	Internet use, web pages in general, open	X						
kpX_1721	Political use of print media, habitually	X						
kpX_1741	TV use, news, habitually	X						
kpX_1760	News magazines, habitually	X						
kpX_1600	Internet use, politically relevant		X	X	X	X	X	X
kpX_1661	Use of print media, politically relevant		X	X	X	X	X	X
kpX_1681	TV use, news, up to date		X	X	X	X	X	X
kpX_1700	News magazines, up to date		X	X	X	X	X	X
kpX_396	Media Reliance							X
	Conversations about politics	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_1931	Conversations about politics, in general	X						
kpX_1951	Interlocutor 1, knowledge of politics	X						
kpX_1961	Interlocutor 1, conflicts	X						
kpX_1991	Conversational network, party preference in general	X						
kpX_1932	Conversations about politics, up to date		X	X	X	X	X	X
kpX_1970	Interlocutor 1, voting decision		X	X	X	X	X	X
	Political knowledge	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_090	Political knowledge: 5% threshold		X			X	X	X
kpX_110	Political knowledge: first/second vote		X		X		X	X
kpX_130	Political knowledge: Electoral law		X		X		X	X
kpX_3430	Political knowledge: Mapping of politicians and parties	X		X				X
kpX_3430	Political knowledge: mapping of politicians and parties		X		X		X	
kpX_3440	Economic knowledge: unemployment rate	X			X			X
kpX_010	Interest in politics, in general	X	X	X	X	X	X	X
kpX_3450	Economic interest	X			X			X
kpX_380	Importance of election results		X	X	X	X	X	
kpX_390	Interest in the election campaign		X	X	X	X	X	
kpX_391	Interesting election campaign							X
kpX_392	Paid attention to election campaign		X	X	X	X	X	X
kpX_395	Helpful election campaign			[X]	[X]	[X]	[X]	X
kpX_421	Contact with parties I		X	X	X	X	X	X
kpX_421	Contact with parties II		X	X	X	X	X	X
kpX_070	Other forms of participation, retrospective		X					
	Virtual election campaign	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_3490	Wahl-O-Mat, usage					X	[X]	[X]
kpX_3500	Wahl-O-Mat, reasons for usage					X	[X]	[X]
kpX_3510	Wahl-O-Mat, congruence[accordance]					X	[X]	[X]
	TV debate	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_1781	Watching of TV debate				X			

Variable	Item	W1	W2	W3	W4	W5	W6	W7
kpX_1790	TV debate: expected result				X			
kpX_1800	TV debate: reception					X	[X]	[X]
kpX_1810	TV debate: perception of Merkel's performance					X	X	X
kpX_1820	TV debate: perception of Steinbrück's performance					X	X	X
kpX_1870	TV debate: reception of media response					X	X	X
kpX_1880	TV debate: perceived media response Merkel					X	X	X
kpX_1890	TV debate perceived media response Steinbrück					X	X	X
kpX_1900	TV debate: reception of the response of the social environment					X	X	X
kpX_1910	TV debate: reception of social environment's response about Merkel					X	X	X
kpX_1920	TV debate: reception of social environment's response about Steinbrück					X	X	X
kpX_3460	TV debate: number of people that watched with the respondent					X		
kpX_3470	TV debate: emotional closeness to people that watched with the respondent					X		
kpX_3480	TV debate: political knowledge of people that watched with the respondent					X		
kpX_3520	TV debate: intensity of interaction					X		
kpX_3530	TV debate: opinion leadership 1					X		
kpX_3531	TV debate: opinion leadership 2					X		
kpX_3550	TV debate: perceived outcome of people that watched with the respondent					X		
kpX_3560	TV debate: preferred chancellor of the people that watched with the respondent					X		
kpX_3570	TV debate: attention					X	[X]	[X]
kpX_3580	TV debate:specific performance in debate Merkel/Steinbrück					X	[X]	[X]
kpX_3600	TV debate statements Merkel/Steinbrück					X	[X]	[X]
	State election in Bavaria	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_3120	Intention to vote, Bavaria		X	[X]	X	X		
kpX_3130	Voting intention, Bavaria		X	[X]	X	X		
kpX_3140	Satisfaction with government, Bavaria		X	[X]	X	X		
kpX_650	Scalometer politicians: Christian Ude		[X]	X	X	X		
kpX_3160	Attention in election campaign, Bavaria		X	[X]	X	X		
kpX_3170	Anticipated majority, Bavaria		X	[X]	X	X		
kpX_3180	Turnout, Bavaria						X	[X]
kpX_3190	Actual voting, Bavaria						X	[X]
kpX_3200	Moment of voting decision, Bavaria						X	[X]

Variable	Item	W1	W2	W3	W4	W5	W6	W7
kpX_3210	Knowledge of result: Bavaria						X	[X]
kpX_3220	Election results: Winners and losers, Bavaria, retrospective						X	[X]
	State election in Hessen	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_3230	Intention to vote, Hessen		X	[X]		X	X	
kpX_3240	Voting intention, Hessen		X	[X]		X	X	
kpX_3250	Satisfaction with government, Hessen		X	[X]		X	X	
kpX_650h, i	Scalometer politicians: Volker Bouffier/Thorsten Schäfer-Gümbel		[X]	X		X	X	
kpX_3270	Attention in election campaign: Hessen		X	[X]		X	X	
kpX_3280	Anticipated majority, Hessen		X	[X]		X	X	
kpX_3290	Turnout, Hessen							X
kpX_3300	Actual voting, Hessen							X
kpX_3310	Moment of voting decision, Hessen							X
	Current topics	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_2900	Opinions about the NSA scandal		X	X	X	X	X	X
kpX_2910	Current socio-economic questions, opinion about criminal case against Wulff		X			X		
kpX_2920	Opinions about the drone project			X				
kpX_2930	Speculations about a red-red-green coalition			X		X	X	
kpX_2940	Opinions about car toll, veggie day and program to prevent rent increases				X		X	X
kpX_2950	Good election poster				X			
kpX_2960	Good election poster: party				X			
kpX_2961	Good election poster: other party				X			
kpX_2970	Opinions about the conflict in Syria					X	X	X
kpX_2980	"Wahlarena": reception						X	
kpX_2990	"Wahlarena": perceived result Merkel/ Steinbrück						X	
kpX_2991	"Wahlarena": statements, knowledge						X	
kpX_2992	"Wahlarena": statements, evaluation						X	
kpX_3590	Knowledge of taxation plans by the Green party						X	X
kpX_3591	Opinion on taxation plans by the Green party						X	X
kpX_4010	Good election advertising						X	
kpX_4020	Good election advertising: party						X	
kpX_4021	Good election advertising: other party						X	
kpX_4030	Awareness of middle finger cover of Steinbrück						X	X
kpX_4031	Opinion about middle finger cover of Steinbrück						X	X
kpX_4040	Opinion about Jürgen Trittin's pedophilia affair							X
kpX_4050	Evaluation of final stage of election campaign							X

Variable	Item	W1	W2	W3	W4	W5	W6	W7
kpX_4060	Preferred coalition							X
	Personality, values, basic orientations	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_1570	Needs	X						
kpX_2180	Big 5	X						
kpX_2090	Party identification	X		X		X		X
kpX_2091	Party identification, other party	X		X		X		X
kpX_2100	Party identification, strength	X		X		X		X
kpX_2110	Party identification, period of time	X						X
kpX_2121	Kind of party identification	X						X
kpX_2081	Inglehart items				X	[X]	[X]	
kpX_2170	General life satisfaction			X				
kpX_2200	Attachment items			X				
kpX_020	Satisfaction with democracy	X						X
kpX_030	Notion of democracy			X				
kpX_040	Disenchantment with politics			X	X			
kpX_050	Efficacy and voting norm	X		X				X
kpX_050	Democracy principles			X				
kpX_060	Extremism items			X				
kpX_160	Trust in institutions			X				
	Socio-demographics	**W1**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**
kpX_2441	Household size	X						
kpX_2450	Number of household members under 18 years	X						
kpX_2591	Net income household, categories	X						
kpX_2461	Membership in organizations	X						
kpX_2470	Trade union membership household	X						
kpX_4000	Party membership	X						
kpX_4110	Party membership, other party	X						
kpX_3910	Religiosity	X						
kpX_2480	Religious affiliation [religion]	X						
kpX_2491	Church attendance	X						
kpX_2580	Social class, subjective	X						
kpX_2601	Federal state	X						
kpX_2600	Residence	X						
kpX_2602	Postal code	X						
kpX_2320	School-leaving qualification	X						
kpX_2330	Vocational [&professional training] education	X						
kpX_2340	Employment status	X						
kpX_2350	Previous employment status	X						
kpX_2360	Unemployment last ten years	X						
kpX_2370	Duration of unemployment	X						
kpX_2380	Job [profession]	X						
kpX_3610	Employee - differentiation	X						
kpX_3620	Worker - differentiation	X						
kpX_3630	Self - employment, graduate profession - differentiation	X						

Variable	Item	W1	W2	W3	W4	W5	W6	W7
kpX_3640	Civil servant - differentiation	X						
kpX_3650	Employment sector	X						
kpX_3660	Economic sector	X						
kpX_3670	Temporary employment	X						
kpX_3680	Fear of losing job	X						
kpX_3690	Fear of losing business	X						
kpX_2390	Previous profession	X						
kpX_3710	Former profession employee - differentiation	X						
kpX_3720	Former profession worker - differentiation	X						
kpX_3730	Former profession self- employment, graduate profession - differentiation	X						
kpX_3740	Former profession civil servant - differentiation	X						
kpX_3750	Former profession - employment sector	X						
kpX_3760	Former profession - economic sector	X						
kpX_2541	Country of birth	X						
kpX_3920	Country of birth, federal state	X						
kpX_3930	Country of birth, other country	X						
kpX_2551	Immigration	X						
kpX_3940	Immigration, East Germany	X						
kpX_3950	Immigration, West Germany	X						
kpX_3960	Immigration to Berlin	X						
kpX_2520	German citizenship since birth		X	[X]	[X]			
kpX_2571	Country of birth, mother		X	[X]	[X]			
kpX_2571	Country of birth, father		X	[X]	[X]			
kpX_2572	Country of birth, mother, other country		X	[X]	[X]			
kpX_2572	Country of birth, father, other country		X	[X]	[X]			
kpX_3980	Immigration mother		X	[X]	[X]			
kpX_3990	Immigration father		X	[X]	[X]			
kpX_3970	Registration of 3rd generation migrants		X	[X]	[X]			
kpX_4120	Spoken language in household		X	[X]	[X]			
kpX_4130	Spoken language in household, other language		X	[X]	[X]			
kpX_4140	Identification of displaced people ("Aussiedler")		X	[X]	[X]			
kpX_4150	Identification of asylum seekers		X	[X]	[X]			
kpX_2301	Marital status	X				X	[X]	[X]
kpX_2311	Partner	X				X	[X]	X
kpX_2312	Cohabitant in household	X				X	[X]	X
kpX_3890	Voting intentions cohabitant	X				X	[X]	X
kpX_3900	Voting intentions cohabitant, other party	X				X	[X]	X
kpX_2391	School-leaving qualification, partner					X	[X]	[X]
kpX_2400	Employment ,partner					X	[X]	[X]
kpX_2410	Previous employment, partner					X	[X]	[X]

Variable	Item	W1	W2	W3	W4	W5	W6	W7
kpX_2420	Job, partner					X	[X]	[X]
kpX_2430	Previous profession, partner					X	[X]	[X]
kpX_3700	Vocational education, partner					X	[X]	[X]
kpX_3770	Employee - differentiation, partner					X	[X]	[X]
kpX_3780	Worker - differentiation, partner					X	[X]	[X]
kpX_3790	Self-employment, graduate profession - differentiation, partner					X	[X]	[X]
kpX_3800	Civil servant - differentiation, partner					X	[X]	[X]
kpX_3810	Employment sector, partner					X	[X]	[X]
kpX_3820	Economic sector, partner					X	[X]	[X]
kpX_3830	Former profession employee - differentiation, partner					X	[X]	[X]
kpX_3840	Former profession worker - differentiation, partner					X	[X]	[X]
kpX_3850	Former profession self-employment, graduate profession - differentiation, partner					X	[X]	[X]
kpX_3860	Former profession civil servant - differentiation, partner					X	[X]	[X]
kpX_3870	Former profession employment sector, partner					X	[X]	[X]
kpX_3880	Former profession economic sector, partner					X	[X]	[X]
kpX_4200	Current health condition			X	X		X	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 (also see Chapter 5.8).

4.8 Paradata

The published data set 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 25). The second group includes information about the panelists' survey activities and is provided by Respondi (see Table 26).

Table 25: 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 data set correspond 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 data set correspond to Greenwich Mean Time (GMT).
lfdn	Every survey participant receives a consecutive number.
kp1_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" symbolizes the particular wave.

The paradata, which was requested of Respondi, provide information about the panelists' entry in the panel, the way of recruitment, as well as the survey participation behavior of the panelists: 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 September 21, 2013.

Table 26: Paradata provided by Respondi

Variable	Explanation
p_enter_date	Date of joining the Respondi panel
p_recruit	Way of recruitment
p_recruit1	General type of recruitment
p_numinv2	Number of survey participation invitations (in the last 12 months; key date: 09/21/2013)
p_numcpl2	Number of survey participations (in the last 12 months; key date: 09/21/2013)
p_numstr2	Number of commenced surveys (in the last 12 months; key date: 09/21/2013)
p_numinc2	Number of interrupted surveys (in the last 12 months, key date: 09/21/2013
p_numinv3	Number of survey participation invitations (in the last 3 months; key date: 09/21/2013)
p_numcpl3	Number of survey participations (in the last 3 months; key date: 09/21/2013)
p_numstr3	Number of commenced surveys (in the last 3 months; key date: 09/21/2013)
p_numinc3	Number of interrupted surveys (in the last 3 months; key date: 09/21/2013)

4.9 Unpublished variables

Whilst carrying out online surveys, variables that are relevant for the implementation of the survey but do not carry any textual meaning are collected or transmitted to GESIS. Those variables are not published, but can be drawn from GESIS (gles@gesis.org) if needed. Some collected data cannot be downloaded independently due to reasons of data protection (e.g. postal codes) as well. These variables can be obtained under specific conditions.

Table 27: Overview of unreleased variables and subscription opportunities

Variable name	Explanation	Subscription opportunity
kpx_ats	Absolute time stamp	available on request
kpx_page_history	Page history	available on request
kpx_lfdpagenr	Current page number	available on request
kpx_invcode	Personal invitation code for every participant	available on request
kpx_codelink	Personal link to the survey for the respondent. Consists of the URL of the survey and the personal invitation code.	available on request
kpx_2293s	Place of birth (first letter)	available on request
kpx_2061	Postal code	available on request
kpx_xxxx_org, _c1-2	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_2293_c3	Correction of birth code	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.

5 Data preparation

5.1 Preliminary note

In a 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 therefore it is possible to reproduce all steps.

5.2 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 28: Brief overview of the missing value codes of GLES

SPSS code	Stata code	Label
-71	.p	haven't heard of term
-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

5.3 Encoding of the parties

In order to be able to compare all GLES components better with each other, 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 are 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 (AfD, die Grauen, Freie Wähler, NDP, oedp, Piraten, REP). [13] The respondents were only able to choose these parties if they had chosen "other party" in the

---

[13] Initially the list of options also included the party "DVU". DVU answers were added to the NPD category, because the DVU joined the NPD.

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 version A and B, although partial information was accessible in these cases. They chose "other party" on the first page. If one wishes to restore this information, Table 29 can be used as it documents the respondents' consecutive numbers, whose codes were changed as mentioned above, in the data set (variable "lfdn").

The question about the strength, the duration and the type of party identification ("kpX_2100", "kpX_2110", "kpX_2121") was only filtered for participants who stated any information in the question "kpX_2090" (party identification). Therefore, those participants whose value of this variable was recoded to -99 [still] feature values in the variables "kpX_2100", "kpX_2110" and "kpX_2121". In order to be consistent, the information that was given in these cases was recoded to -97 "not true". The concerned cases are to be found in Table 29 in the lines that deal with party identification.

Table 29: Respondents whose party variables were changed subsequently because of refusal of response or discontinuation

Variable	Item	Consecutive number of the respondents (lfdn)
Wave 1		
kp1_190ba-b	Intended voting: second vote	1278, 2686, 3465, 3745, 3761, 4519
kp1_850a-b	Ability to solve most important issue	1845, 4424, 4988
kp1_870a-b	Ability to solve second most important issue	104
kp1_350aa-b	Recall previous parliamentary elections, first vote	2598
kp1_350ba-b	Recall previous parliamentary elections, second vote	718, 973, 1513, 4323
kp1_2090a-b	Party identification	95, 261, 326
Wave 2		
kp2_190ba-b	Intended voting: second vote	190, 1227, 1311, 1980, 2021, 2183, 2536, 2892, 3758,
kp2_850a-b	Ability to solve most important issue	1778, 2640, 4582
kp2_870a-b	Ability to solve most important issue	4630, 5125
Wave3		
kp3_190ba-b	Intended voting: second vote	255, 1227, 1289, 3136
kp3_850a-b	Ability to solve most important issue	1090, 3251
kp3_870a-b	Ability to solve second most important issue	1138, 3251, 4873
kp3_2090a-b	Party identification	3753
Wave 4		
kp4_190ba-b	Intended voting: second vote	2519, 2536, 3787, 4308
kp4_850a-b	Ability to solve most important issue	4346
kp4_870a-b	Ability to solve second most important issue	3251

Wave 5		
kp5_190ba-b	Intended voting: second vote	707, 2825
kp5_850a-b	Ability to solve most important issue	907
kp5_870a-b	Ability to solve second most important issue	2341, 4364
kp5_2090a-b	Party identification	1032
kp5_3890a-b	Voting intention, partner	4381
Wave 6		
kp6_190ba-b	Intended voting: second vote	89, 1593, 2023, 2091, 2175, 3096, 3286, 3473, 4330
kp6_2770ba-b	Hypothetic second vote after postal vote	1311
kp6_870a-b	Ability to solve second most important issue	1140, 2255, 2661
Wave 7		
kp7_200ba-b	Factual voting behavior after parliamentary elections: second vote	4423, 4906
kp7_870a-b	Ability to solve second most important issue	284, 1182, 3595
kp7_2090a-b	Party identification	4982, 5038
kp7_3890	Voting intention, partner	568, 2656

5.4 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 project team. This includes the variables "Reasons for Voting Decision" ("kpX_260"), the question about the most important and the second most important issue in Germany (agenda questions, "kpX_840", "kpX_860"), the question about the most used internet site for political purposes ("kp1_1651s") as well as the reason for the survey participation ("kp1_4230s").

As encoding is finished, encodings of open statements are now accessible as well. They are marked by the suffix _c1, whereas the number 1 represents the position of the mention. Up to five mentionings were encoded, therefore suffixes from _c1 (e.g. "kp1_4230_c1") to _c5 were used. The open statements are still included in the dataset. They were checked on content that falls under the Data Protective Directive and were censored if necessary to protect the respondents' privacy.

In the course of the encoding, statements that did not contain a contextual message were encoded -99. Those statements were however counted as data by the survey software, so the corresponding respondents provided additional answers to questions like the one about solution capability ("kpX_850a,b"). The user is able to decide whether they take the information into account for the analysis or not.

In order to work with open answers, we recommend the usage of the SPSS dataset, as all open statements in the Stata dataset are cut off after 244 digits due to a software restriction.

The corresponding encoding schemes are accessible on the GLES pages of the GESIS web site (http:/www.gesis.org/gles).

5.5 Marking of knowledge issues

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	2, 4, 6, 7
kpx_130	Political knowledge: German Electoral Law	2, 4, 6, 7
kpx_3430a-f	Political knowledge: Mapping politicians and parties	1, 3, 7
kpx_3430j-o	Political knowledge: Mapping politicians and parties	2, 4, 6
kpx_3440	Economic knowledge: Unemployment rate	1, 4, 7

The question about the five- percent rule ("kpX_090") is an exception and was asked in the second, fifth, sixth and seventh wave. 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.

5.6 Handling of questions with "checkboxes" and "slide controls"

Some forms of questions required an enquiry if the respondent gave no 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 be possible to draw a distinct line between possible substantive replies like 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
kp1_1721a-g	Political use of print media, habitually	Slide control	1
kp1_1741a-e	TV usage, news habitually	Slide control	1
kp1_1931a-e	Conversations about politics, in general	Slide control	1
kpX_421aa-lf	Contact with political parties I and II	Checkbox	2-7
kpX_1661a-h	Use of print media, politically up to date	Slide control	2-7
kpX_1681a-f	TV usage, news, up to date	Slide control	2-7
kpX_1932a-f	Conversations about politics, up to date	Slide control	2-7

Explanation: "X" represents the respective wave.

5.7 Summary of time spans

The medium daily internet use of the Campaign Panel participants was collected in hours and minutes in two answer lists. Also, the unemployment duration in the last ten years was collected in two answer lists in years and months. In both cases, the two time spans were summarized in a time variable: The duration of daily internet use is displayed in minutes ("kp1_3450"), and the duration of unemployment is displayed in months ("kp1_2370").

5.8 Summary of the 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. This technique was used in two different contexts in the Campaign Panel:

1. *To avoid duplication in data collection.* A good example is postal voting: When the respondents stated in one wave that they already had had a postal vote, they did not receive any questions concerning their participation in the election anymore. They received questions that were suitable for absentee voters instead.
2. *To retake missing measurements.* It seemed appropriate to collect comparatively stable characteristics (e.g. socio-structural characteristics, value orientation, remembering behavior patterns in the past) only once in the Campaign Panel, but from preferably *all* participants. It was not possible to request all stable characteristics in the first wave (in which all respondents participated) due to limited questionnaire space. Therefore, the requesting was partly also done in later waves. If respondents did not participate in the particular wave, their data was collected in a consecutive wave. In this way, it was possible to collect data of nearly every participant - despite of panel attrition.

The summary outlined here only includes those DI variables which were collected in a subsequent measurement (bullet point 2). In order to simplify working with retaken measurements, these were subsequently saved in the variables of that wave in which the corresponding characteristics had been requested for the first time.

One example: Opinions about fair taxation were collected in the second wave ("kp2_3400a-d"). A respondent who did not participate in the second wave then received the questions in the fourth wave instead ("kp4_3400a-d"), otherwise in the sixth wave (" kp6_3400a-d"). In the course of data preparation, the measured data from the fourth and sixth wave was transferred to the variable "kp2_3400a-d". The variables "kp2_3400a-d" and "kp6_3400a-d" were deleted afterwards. A flag variable in the dataset ("kp2_3400flag") helps to reproduce the wave in which the information was initially collected. [14][15]

---

[14] Exceptions to this rule are the variables "kpX_2301", "kpX_2311" and "kpX_2312". These were each requested in the first wave and then in wave 5 as well, as one had to made sure whether the family status had changed. A non-participation in wave five led to the fact that those respondents received the question "kpX_2301" as well in the waves six and seven. Correspondingly, the data of the sixth and seventh wave were combined in wave five and the statements from wave 1 were maintained. For the same reason - to include possible changes of the family status - the questions "kpX_2311" and "kpX_2312" were also given to wave 7 participants who had already answered those questions in wave 5. Only the data of those participants who had not participated in wave five and six were summarized. Correspondingly, "kp7_2311" and "kp7_2312" remain in the dataset.

[15] In order to simplify the DI filtering, one accepted that duplication in the collection of the corresponding characteristics was possible in exceptional cases (these were dropouts of previous waves whose status was unclear; one did not know whether they dropped out before or after the collection of the corresponding item). In such a case, the *last* answer of the respondent was used when summarizing the variables.

Therefore it is possible in exceptional cases that a variable from a distinct wave includes potential interpretable measured data of the respondents who did not take part in that wave. The affected variables are listed in Table 32.

Table 32: List of affected variables

Variable	Item	First appearance in wave …	Retaken in wave …
kpx_3400a-d	Fairness, taxes	2	4, 6
kpx_2520	German citizenship since birth	2	3, 4
kpx_2572a	Country of birth, mother, other country	2	3, 4
kpx_2572b	Country of birth, father, other country	2	3, 4
kpx_2571a,b	Country of birth, mother/father	2	3, 4
kpx_3970	Registration of 3rd generation migrants	2	3, 4
kpx_3980	Immigration, mother	2	3, 4
kpx_3990	Immigration, father	2	3, 4
kpx_4120	Language spoken in household	2	3, 4
kpx_4130	Language spoken in household, other language	2	3, 4
kpx_4140	Identification of displaced people ("Aussiedler")	2	3, 4
kpx_4150	Identification of asylum seekers	2	3, 4
kpx_3370a-d	Fairness, income	3	5, 7
kpx_2081a-d	Inglehart items	4	5, 6
kpx_1800	TV debate: reception	5	6, 7
kpx_3570	TV debate: attention	5	6, 7
kpx_3580a-l	TV debate: specific performance Merkel/Steinbrück	5	6, 7
kpx_3600a-d	TV debate: statements Merkel/Steinbrück	5	6, 7
kpx_2301	Marital status	1/5*	6, 7
kpx_2311	Existing partner	1/5*	6, (7)**
kpx_2312	Partner living in household	1/5*	6, (7)**
kpx_2391	School-leaving qualification, partner	5	6, 7
kpx_2400	Employment, partner	5	6, 7
kpx_2410	Previous employment, partner	5	6, 7
kpx_2420	Profession, partner	5	6, 7
kpx_2430	Previous profession, partner	5	6, 7
kpx_3700	Vocational education	5	6, 7
kpx_3770	Employee - differentiation, partner	5	6, 7
kpx_3780	Worker - differentiation, partner	5	6, 7
kpx_3790	Self-employment, graduate profession - differentiation partner	5	6, 7
kpx_3800	Civil servant - differentiation, partner	5	6, 7
kpx_3810	Employment sector	5	6, 7

Variable	Item	First appearance in wave…	Retaken in wave …
kpx_3820	Economic sector	5	6, 7
kpx_3830	Former profession employee - differentiation partner	5	6, 7
kpx_3840	Former profession worker - differentiation, partner	5	6, 7
kpx_3860	Former profession civil servant - differentiation partner	5	6, 7
kpx_3850	Former profession self-employment, graduate profession - differentiation partner	5	6, 7
kpx_3870	Former profession employment sector, partner	5	6, 7
kpx_3880	Former profession economic sector, partner	5	6, 7
kpx_3180	Turnout, Bavaria	6	7
kpx_3190a,b	Actual voting, Bavaria	6	7
kpx_3200	Moment of voting decision, Bavaria	6	7
kpx_3210	Knowledge of election, Bavaria	6	7

* Demand from the waves 6/7 is summarized in wave 5.
** Information from wave 7 is only summarized in wave 5 if the respondent did not participate in the waves 5 and 6.

6 Data quality

6.1 Preliminary note

As there are no interviewers present in online surveys, the interview situation is characterized by a high level of anonymity.[16] Anonymity can be an advantage if data about socially undesirable attitudes or behavior patterns is collected, as respondents tend to answer more honest in self-administrative surveys. (c.f. 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", c.f. 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 Campaign Panel 2013 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 Campaign Panel 2013 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 (see section 6.2), self-assessment of the data quality (section 6.3) as well as Bogus Items (section 6.4). In later releases of the Campaign Panel, further quality indicators will be published. Furthermore, a report will be released in which measures for providing and evaluating data quality and the general handling of the latter will be described.

Data quality issues in online surveys are not only a result of lacking motivation of the respondents, but also - although quite rarely - due to technical errors. Filter errors are described in section 6.5. In section 6.6 cases will be listed in which the respondents pointed out errors themselves.

---

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

6.2 Indicators for data quality

6.2.1 Self-assessment of data quality by the respondents

At the end of every interview, the respondents of the Campaign Panel were asked whether they answered the questions attentively ("kpX_4210") and thoroughly ("kpX_4220"). Although it must be considered that those questions were not answered truthfully by every respondent, a small number admitted the lack of care and attention. First analyses indicate that those respondents also perform poorly in objective data quality indicators.

6.2.2 Response times

If respondents give quick answers this does not necessarily indicate lacking data quality. A lot of 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 one knows that there are some respondents in all online surveys who just click through the survey without reading the questions and the answers at all. As a general rule, these persons provide very quick response times.

There are no established standards for the identification of "too quick" responses in 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 data set or highlighted by marker variables.

The Campaign Panel dataset contains three indicators for quick response times. These measure different aspects but have a strong empiric correlation:

- The indicator by Roßmann (2010; variable "kpX_speederindex") includes the participants' response times on every page of the survey, as well as the overall duration per respondent. The index values assume values between greater 0 and less than 2. An index value of 1 indicates an average response time, whereas values approaching 0 indicate very quick response times, and values approaching 2 indicate very slow ones.

- The variable "kpX_mtime" contains the average period of time (in seconds) a respondent needed to answer the questions on one page of the survey. When calculating this mean value, residence times of more than 20 seconds per page were trimmed to 20 seconds to ensure that outliers do not influence the mean value.

- The variable "kpX_mtimex" is a variant of "kpX_mtime" that was slightly modified. The variable was created to make the individual mean response times comparable over all waves of the Campaign Panel. Therefore, the individual response times per page were adjusted for the average response time of this page, as there was the need to consider that the different pages of the survey require varying levels of effort. Afterwards, one calculated the mean value for every participant over the adjusted times that the respondents needed to work on the questionnaire. Example for an interpretation: A value of -2 indicates that a person answered 2 seconds quicker per page than the average respondent.

6.2.3 Straightlining

The term "Straightlining" implies a special response pattern of answering matrix questions in a way which generally implies that the questionnaire was not answered with care. According to the definition of straightlining, respondents tend to choose the same value within one column of the matrix question (mostly the middle category); thus the answers represent a

vertical line. Obviously, such a pattern can also be explained by content, as all "true" answers are randomly in line. If, however, straighlinging patterns arise repeatedly in several matrix questions, the cause is more likely to be the insufficient [careless] implementation of the interview.

Matrix questions with at least three lines [rows] were used to calculate a straightlining index. Then one calculated how many matrix questions were straight-lined (in percent) for every respondent. The percentages are included in the variable "kpX_strl".

6.2.4   Item Non-Response

Item non response occurs when respondents refuse to answer a question (answer code -99). A high proportion/share of item non response suggests that the respondents only clicked through the questionnaire without having paid attention to the question and answer stimuli. The variable "kpX_itnrp" includes the share of all questions that the respondents received in one wave and to which the answer was refused.

6.2.5   Don't - know answers

A high share /proportion of don't-know answers in political surveys leads to the assumption of a poor political knowledge. However, don't-know answers are also considered an indicator for satisficing in scientific literature (Krosnick 1991). Therefore, an indicator was established to measure the dimension of don't- know answers in a wave: The variable "kpX_dokn" contains the share/proportion of all don't-know answers, as measured by the overall number of don't - know answer options that the respondents received in one wave.

6.2.6   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 behavior of responding, at least one control item per wave was included in matrix questions. In this item, the respondents were not asked to state their political opinion, but to click a particular response category (see Figure 4). 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").

Figure 4: Example for a control question (screenshot from wave 1)

Note: The orange box serves to highlight and was added subsequently. The partcipants were asked in german, so the example have been translated.

6.2.7　A combined index

Finally, an overall quality indicator ("kpX_quality"), which contains several of the individual indicators listed above, was created. This indicator disposes of a theoretical range of 0 to 1 and indicates the probability by which respondents answered a bogus item incorrectly (see 6.2.6). Thus, the quality of an answer has to be assessed lower the higher the index value is.

Fixed- effects panel regression was used to calculate the index. The answers to bogus items functioned as dependent variable (0 = correct answer; 1 = incorrect answer). All quality indicators that were characterized in the sections 6.2.1 to 6.2.5 were teste das potential explanatory variables.

In the end, however, one model was maintained that only included response times (third index, 6.2.2) and the indicator for straightlining (6.2.3), as well as a complex, non-linear interaction between both included indicators. Finally, the predicted probabilities, whose values create the combined index, were calculated based on the model. A full description of the index will be provided in a separate report.

6.2.8　A deviation index

The index "kpX_dev" indicates the extend to which the answers of one participant in one wave deviate from their answers in the other waves. It is based on the answers to all questions which the participants of the Campaign Panel received at least twice in all seven panel waves.

The background of the index is the issue that it is very challenging to determine whether a participant of an online survey really is the targeted person (see section 3.4). With the index possible respondents who are not the targeted person shall be identified: The higher the index value, the higher the possibility that the interviewed person is not the targeted person who should originally have been interviewed. Admittedly it can be assumed that the verification mechanism, which was introduced in section 3.4, helped to prevent most of these cases.

The logic of the index can be illustrated with the help of an example: Granted that the target person took part in the firstv wave, in the second wave their spouse participated instead and in all remaining waves the target person filled out the questionnaire again. In this case, the answers to the questions of the second wave would differ significantly from the answers to the same questions in the other waves. Correspondingly, the deviation index registers a higher value in the second wave. As a rough rule of tumb, index values which are larger than 2 can be seen as "suspicious".A full description of the index will be provided in a separate report.

6.3 Filter errors

Filter errors in wave 1

- The respondent with lfdn 2313 did not see the question "kp1_2470" (Trade union member in household), although filter conditions were correct. We assume problems with the internet connection.

- The value of the variable "kp1_190a" (intended voting, first vote) was subsequently changed from "-97 - not applicable" to "801 - other party" for the respondent with the lfdn 80, as they had provided a valid answer on that screen which demanded replies about other parties. Furthermore, other answers of the respondent with the lfdn 60 were subsequently changed: The variable "kp1_350a-b" (recall preceding parliamentary elections: first and second vote) was recoded from "-97 - not true" to "801 - other party", as the person gave valid answers on the screen concerning other

parties. The reason for the error is not constructible in both cases. We assume problems with the internet connection.

- One person (lfdn 3356) stated to want to go vote, but was filtered for some strange reasons over the item "kp1_190a-b" (intended voting). Furthermore, the respondent did not see the items "kp1_650o,h,j" (scalometer politicians) for reasons that could not be explained (the variables contained a system missing), although they answered the first three items of the matrix ("kp1_650a,k,c"). The system missings of the items "kp1_650o,h,j" were subsequently recoded to -92 "error in data".

- One person (lfdn 3356) did state that they wanted to take part in the election but was for some strange reasons filtered over the item "kp1_190a-b" (intended voting).

- Wave 1: One person (lfdn 700) gave contradictory information regarding item "kp1_1931a-f" (conversations about politics, in general). On the one hand, they stated to have talked to friends about politics twice a week and on the other hand that they did not have conversations at all. Due to plausibility checks that did not work in this case for unknown reasons, this combination of answer possibilities would not have been allowed to occur. This case is marked in "kpx_info".

Filter errors in wave 2

- One respondent (lfdn 2820) shows a system missing value of the variable "kp2_3990" although this person actually saw the site. The system missing was recoded to -92 "error in data".

Filter errors in wave 4

- A respondent (lfdn 5233) did not see the question about their likelihood of voting ("kp4_170") due to a filter error, although they stated in "kp3_170" that they consider going to the polls. The case is documented in the variable "kpx_info".

- 24 respondents received the question "Reasons for decision not to vote, closed" (variable kp4_252) due to a filter error although they stated in the question concerning their likelihood of voting (variable "kp4_170") that they "will surely go to the polls". These cases are documented in the variable "kpx_info".

- Two respondents (lfdn 553 and 1800) received the question "Reasons for decision not to vote, closed" (variable "kp4_252") due to a filter error, although they had already stated in the question concerning their likelihood of voting (variable "kp4_170") that they "already had a postal vote". The cases are documented in the variable "kpx_info".

- Ten respondents received the question "Reasons for decision not to vote, closed" (variable "kp4_252") due to filter errors, although they stated in the question which requested an answer concerning their likelihood of voting (variable "kp4_170") "to probably go to the polls". These cases are documented in the variable "kpx_info".

- Four respondents (lfdn 590, 1099, 2883, 4713) received the question "Reasons for decision not to vote, closed" (variable "kp4_259") due to filter errors, although they had already stated a reason in wave 1 ("kp1_252") and did not participate in wave 2 and 3. These cases are documented in the variable "kpx_info".

Filter errors in wave 5

- One respondent (lfdn 3677) received the question "kp5_170" (likelihood of voting) due to a filter error, although they had already stated in wave 4 to have had a postal vote.

- One respondent (lfdn 3996) received the questions "kp5_170", "kp5_191a-b", "kp5_2751a-g", "kp5_2760", "kp5_330", "kp5:331" and "kp5_395" (likelihood of voting, postal votes, consideration set, difficulty of voting decision and the difficulty of the turnout, election campaign was helpful) due to a filter error, although they stated in wave 4 to already have had a postal vote. Instead they did not receive the question "kp5_2770a-b" (hypothetic vote after postal vote). The information concerning postal votes in between the waves 4 and 5 are partially inconsistent.

6.4 Respondents' feedback

At the end of every interview, the respondents were able to give feedback. Besides lots of positive feedback, occasional answers that included technical issues or corrected wrong answers were also collected. All reported issues are listed below.

Technical problems

One respondent (lfdn in wave 2: 3666) reported that instead of pictures of the politicians (items "kpX_3430j-o") red crosses were displayed. The values of the variables "kpX_3430j-o" were recoded in -92 "error in data". Another respondent (lfdn in wave 6: 5157) was not able to see the pictures in the lower half of the browser window. Correspondingly, variables that contained a -99 value "not applicable" (variables "kp6_3430j,l,m") were recoded to -92 "error in data".

Several people reported problems concerning incorrect text display (lfdn in wave 3: 5216; in wave 4: 3206, 1821, 5216; in wave 5: 1432, 1436, 1821 ,2790; in wave 6: 2790; in wave 7: 428, 2790, 5202). No changes in the data were made.

One person reported that the names of the politicians were not displayed correctly. However, it is unclear to which question they refer (lfnd in wave 4: 5006). No changes in the data were made.

One respondent reported technical issues while answering the questionnaire (lfdn in wave 6: 50669). No changes in the data were made.

One person reported that the slide control in the survey failed to work (lfdn in wave 2: 649). No changes in the data were made.

One person (lfdn in wave 6: 5040) reported that the slide control in the question about the current use of political print media did not work. Therefore the values of the variable "kp6_1661a-h" were recoded in -92 "error in data".

Several people reported that the survey loading time had taken a long amount of time (lfdn in wave 2: 1199, 4690; in wave 6: 1700, 4009). No changes in the data were made.

Several people reported that they had issues handling the survey when using a smartphone or a tablet (lfdn in wave 1: 153, 1752; in wave 4: 2906, 3549). No changes in the data were made.

Respondents who reported mistakes in their answers

Several respondents stated that they intended to refuse the answer to the question concerning the voting intention, but they did not find a suitable category and therefore chose the option "I don't know" (lfdn in waves 1 to 6: 2436; in wave 1: 1804, 2436, 4424; in wave 3 1804). These cases' values were recoded to -99 "no answer".

Two respondents mixed up the first vote and the second vote ("kp2_190aa-b,ba-b": lfdn 720; "kp3_190aa-b,ba-b":2332).The answers were corrected corresponding to the respondents' statements.

In wave 1, one person pointed out that they entered a wrong postal code (lfdn 872). The postal code ("kp1_2602") was recoded in -99 "no answer".

In wave 1, one person (lfdn 2462) pointed out that they provided wrong information concerning the size of their place of residence ("kp1_2600"). The variable was recoded according to the number of inhabitants stated in the comments.

In wave 1, one person additionally stated that they were involved in a church-related group. This information was correspondingly inserted in "kp1_2461e".

In wave 1, four people (lfdn 872, 1956, 3325, 3529) admitted that they did not provide correct information concerning the height of their income ("kp1_2591"). The values were recoded to -99 "no answer".

In wave 1, eight people complained about the question asking to state their income ("kp1_2591"), yet answered the question. It is possible that they provided wrong information (lfdn 1738, 1884, 2321, 3345, 3378, 4031, 4285, 4685). No changes in the data were made.

In wave 2, a person mixed up the answer alternatives "rather correct" and "rather not correct" of the variables "kp2_661a-c" (lfdn 4555). The values were accordingly recoded.

In wave 2, a person restated the information that their mother had been born in a foreign country (lfdn 44). Therefore the information contained in the variables "kp2_2570a" (Country of birth, mother") was recoded to "Yes" for Germany and the remaining variables "kp2_2570a", "kp2_3980", "kp2_4120", "kp2_4140", "kp2_4150" were accordingly recoded to "not correct".

In wave 3, a respondent (lfdn 1539) accidentally chose SPD instead of FDP (variable "kp3_190ba-b"). This was corrected subsequently.

In wave 4, a respondent referred to the change of scale legends and that they therefore often chose answers that do not correspond to their opinion (lfdn 2111). No changes in the data were made.

In wave 4, one respondent stated that they mixed up the poles "left" and "right" in the left and right classification of Angela Merkel and Peer Steinbrück ("kp4_680a-b"). The provided information was recoded mirror-inversely.

One respondent (lfdn 4928) stated in wave 5 that they mixed up the poles when classifying the parties on a left-right-spectrum ("kp5_1490a-i"). The provided information was recoded mirror-inversely.

In wave 5, one respondent (lfdn 2518) stated that they answered in the wrong direction when replying to the question concerning the respondent's opinion on the influx of immigrants. However, the respondent furthermore added that they were actually of the opposite opinion. The data was correspondingly recoded.

In wave 6, two respondents (lfdn 211, 4337) stated that they had scaled their negative feelings concerning the political parties ("kp6_2800a-f") in the same way as their positive feelings by mistake. They had not noticed that the scaling had been changed. The statements are recoded to -92 "error in data" because the intended answers cannot be reconstructed.

In wave 7, one person stated that they did skip two questions ("kp7_820" and "kp7_3030") and provided the information in the comments box (lfdn 5045). The two variables were recoded to their statement correspondingly.

In wave 7, one person stated that they had seen the posters of all parties ("kp7_421ga-f"; lfdn 832). The variables were recoded correspondingly.

In wave 7, one respondent (lfdn 1385) corrected their assignment of Angela Merkel to the CSU in the comments box. Instead they assigned her to CDU. Therefore, the value of "kp7_3430a" was recoded.

Other

Wave 1: Although a person (lfdn 970) stated in the comments field that they would vote invalid, they nevertheless stated an answer to "kp1_190aa-b,ba-b". No changes were made.

Waves 2 and 3: One person (lfdn 2315) stated in the open comments field in the waves 2 and 3 that they would vote invalid. Yet, they made valid statements about the parties in the variables "kp2/kp3_190aa-b, ba-b". Another person (lfdn 1298) stated the same remarks in wave 3 for the variable "kp3_190aa-b, ba-b". No changes were made.

Wave 5: One person (lfdn 2127) stated that they used the middle categories more frequently because of a missing "don't know"- category. No changes were made.

Wave 6: The respondent with the lfdn 3717 might possibly not be the target person, as their statements concerning the TV debate are contradictory. They stated in wave 5 that they have seen the TV debate and in wave 6 they stated in the comments that they had to answer questions concerning the TV debate although they had not seen it. No changes were made.

6.5 Version history

Changes from version 1.0.0 to version 2.0.0

Open answers have been encoded ("kpX_840_c1-5", "kpX_860_c1-5", "kpX_260_c1-3", "kpX_4230_c1-3", "kp1_1651_c1-3").

System missings in the variables "wei_w2-7" (weights) and "kp1_2180a-l" (Big 5) were coded with the GLES standard missing values.

By comparing IDs of the Short-Term Campaign Panels 2009 and 2013, participants, who had taken part in 2009 but were not marked in "wkp2009", have been identified and those who had not attended in 2009, but were marked in "wkp 2009" were identified as well. Now, those cases are assigned to the right categories.

Variable "mut09" which identifes panel mutants from 2009 to 2013 has been added.

Variables "kpX_1230" (European Integration, Ego) and "kpX_1781" (Televised debate: Probability of reception) were named "kpX_1250" and "kpX_1780" in 2009. Due to a consistent use of variable names, those variables were renamed according its 2009 version.

Kp4_speederindex was calculated wrong and has been fixed.

Respondents may exhibit missing values for the party identification, due to the creation of Version A and B for all party inquiries, nevertheless they made an informative reply in "kpX_2100" (Party identification strength). This circumstance relies on the two-stage inquiry of party identification. Answers in "kpX_2100" (Party identification strength), which refer to no mentioned party in "kpX_2091", were recoded into -97 "not applicable".

Variables "w1a" (Wave 1 started) and "w1b" (Wave 1 finished) were added in order to complete the list.

Misspellings in labels and questionnaire documentations were corrected.

Changes from version 2.0.0 to version 3.0.0

The major changes in version 3.0.0 are made to make the Short-Term Campaign Panel 2013 comparable to the Short-Term Campaign Panel 2009:

Variable "lfdn09" was added. It contains the serial number of the participants in 2009 which can be used to merge both datasets.

The coding of "kpX_260" (Reasons for voting desicions) was adjusted to the coding used in 2009.

The coding of "kp1_2480" (Religious affiliation) were adjusted to the general GLES scheme and are thus identical with the encoding in 2009.

Due to a better comparability several variables were renamed. If the wording and the answer options differ, but have the same variable name as in 2009, the variable was given a new name. If the questions were collected totally equal but had erroneously different variablenames, they have been renamed into the version of 2009. Following variables were renamed:

"kpX_2570a, b" (Country of birth, Father, Mother)	→	"kpX_2572a,b"
"kpX_250" (Reason for decision not to vote, closed)	→	"kpX_252"
"kpX_660a-u" (Characteristics of chancellor candidates)	→	"kpX_661a-u"
"kpX_960" (Expected coalition)	→	"kpX_961"
"kpX_1650c1-3" (Internet use, web pages in general)	→	"kpX_1651c1-3"
"kpX_1720a-h" (Political use of print media, habitually)	→	"kpX_1721a-h"
"kpX_1740a-f" (TV use, news, habitually)	→	"kpX_1741a-f"
"kpX_1660a-h" (Use of print media, politically relevant)	→	"kpX_1661a-h"
"kpX_1680a-f" (TV use, news, up to date)	→	"kpX_1681a-f"
"kpX_2080a-d" (Inglehart Items)	→	"kpX_2081a-d"
"kpX_2460a-l" (Membership in organizations)	→	"kpX_2461a-l"
"kpX_2540" (Country of birth)	→	"kpX_2541"
"kpX_2550" (Immigration)	→	"kpX_2551"
"kpX_2590" (Net income household, categories)	→	"kpX_2591"
"kpX_2181a-m" (Big 5)	→	"kpX_2180a-m"
"kpX_1571a-q" (Needs)	→	"kpX_1570a-q"

Furthermore, suffixes were adjusted to their use in 2009 (suffixes that are not shown have not been changed):

"kpX_661"	"kpX_261"	"kpX_650"
i → t	b → d2	b → k
k → u	c → e2	d → o
	d → b	e → h
	e → h	f → j
	f → c	g → q
	g → f	h → r
	h → g2	i → z
	i → k2	j → t
	j → i2	o → u
	k → j	k → v
		l → w
		n → x
		m → y

The suffix "a" of the variables "kpX_1070a", "kpX_1110a", "kpX_1270a" and "kp7_341a" was renamed to "j", as there were two separate question concerning the party CDU/CSU in 2009.

The data quality indicators "kpX_dev", "kpX_itnrp", "kpX_strl", "kpX_mtime", "kpX_mtimex", "kpX_dokn" and "kpX_qual", that are mentioned in chapter 6.2 ,were added to the data set.

The paradata variables "p_numinv", "p_numcpl", "p_numstr" and "p_numinc" were deleted from the datatset. They should have contained statistics about the panel behavior since the entry into Respondi-Panel of the particpants, but in fact they did not. They just contained information on the last 372 days and thus only differed slightly from the variables "p_numinv2" till "p_numinc2".

The dataset, questionnaire and study description are now also available in English.

Several corrections in the variables and labels were implemented corresponding to the GLES standards.

Changes from version 3.0.0 to version 3.1.0

Variable and value labels have been corrected.

In variable kp7_3591 the coding was twisted and have been corrected.

In contrast to the questionnaire the wording of "kpX_2820a,b" was different in wave 5. This has been documented in the questionnaire.

Changes from version 3.1.0 to version 3.2.0

In the english version of the dataset the value labels of "kpX_421aa-kf,y" (Party contacts) are wrong. Instead of 0 "mentioned"/ 1 "not mentioned" it is 0 "not mentioned"/ 1 "mentioned". This has been corrected.

7 Matching of other data sets

7.1 Matching of time variables

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

The data set 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). If several items are queried on one page of the survey, as in matrix questions, the time variable of the page is only labelled with the stem of the variable (e.g. "T_kp1_430", not "T_kp1_430a-f").

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".

7.2 Matching of the GLES Campaign Panel 2009 data

To include the Campaign Panel 2009 data (ZA no. 5305), a variable "wkp2009" was established which indicates whether the respondent has taken part in the Campaign Panel 2009. In order to include the data by the variable "lfdn09", the variable names have to be changed in one of the data sets. Otherwise, a doubling of variables, which is not admissible, might occur (for instance, the variable "kp2_010" exists in both datasets). Alternativly by now, there is a cumulated version of both datasets (ZA5757) which can be downloaded at GESIS.

7.3 Matching of the control groups

As described in chapter 3.1, three cross-sectional interviews were carried out simultaneously to the third, fifth and seventh interview of the Campaign Panel. The control group included about 1200 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 5753, 5754 and 5755. 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.

Literature

Baker, Reg und Theo Downes-Le Guin (2007): Separating the Wheat from the Chaff: Ensuring Data Quality in Internet Samples. Konferenzpapier vorgestellt auf: The Challenges of a Changing World: Proceedings of the Fifth International Conference of the Association for Survey Computing in Southampton, 12. bis 14. September.

Blumenberg, Manuela S. und Tobias Gummer (2013): Gewichtung in der German Longitudinal Election Study 2009. In: GESIS – Technical Reports 2013|19.

Blumenstiel, Jan Eric und 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 und Henning Best (Hrsg.). *Handbuch der sozialwissenschaftlichen Datenanalyse.* Wiesbaden: VS Verlag für Sozialwissenschaften, S. 963-994.

Callegaro, Mario und Charles DiSogra (2008): Computing Response Metrics for Online Panels, in: Public Opinion Quarterly 72, 1008-1032.

DeBell, Matthew, Jon A. Krosnick, Arthur Lupia und 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. und 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 (4): 427-444.

Downes-Le Guin, Theo; Joanne Mechling und Reg Baker (2006): Great Results from Ambiguous Sources: Cleaning Internet Panel Data. Konferenzpapier vorgestellt auf: ESOMAR World Research Conference: Panel Research 2006 in Barcelona, 27. bis 29. November.

Gräf, Lorenz und Martin Heidingsfelder (1999): Bessere Datenqualität bei WWW-Umfragen – Erfahrungen aus einem Methodenexperiment mit dem Internet-Rogator. in: Bernad Batinic; Andreas Werner; Lorenz Gräf und Wolfgang Bandilla (Hrsg.). *Online Research. Methoden, Anwendungen und Ergebnisse.* Göttingen: Hogrefe, S. 113-126.

Green, Donald P.; Soo Yeon Kim und 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, S. 433-438.

Krosnick, Jon A. (1991): Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. In: *Applied Cognitive Psychology* 5, S. 213-236.

Krosnick, Jon A. (1999): Survey research. In: *Annual Review of Psychology* 50, S. 537-567.

Mayerl, Jochen und Urban, Dieter (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. 3rd ECPR Graduate Conference. Dublin.

Steinbrecher, Markus/ Roßmann, Joss und Michael Bergmann (2013): Das Wahlkampf-Panel 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 Jackob; Harald Schoen und Thomas Zerback (Hrsg.). *Sozialforschung im Internet.* Wiesbaden: VS Verlag für Sozialwissenschaften, S. 91-107.

The American Association for Public Opinion Research (Hrsg.) (2011): Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Siebte Auflage. o.O.

Links

GLES project page: http://www.gles.eu

German Society of Electoral Research (Deutsche Gesellschaft für Wahlforschung e. V., DGfW): http://www.dgfw.info

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

University of Frankfurt: http://www.uni-frankfurt.de

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

Social Science Research Center Berlin (Wissenschaftszentrum Berlin für Sozialforschung GmbH; WZB): http://www.wzb.eu

Bamberg Center for Empirical Studies (Bamberger Centrum für Empirische Studien, BACES) http://www.uni-bamberg.de/zentren/baces/

The study description was authored by GESIS - Leibniz - Institute for Social Sciences and the University of Mannheim, in collaboration with BACES.

Mannheim, June 2016