

gesis

Leibniz-Institut
für Sozialwissenschaften

German Longitudinal
Election Study



GLES 2017

Wahlkampf-Panel

Kontrollquerschnitt II

(zu Welle 6)

ZA6806, Version 2.0.0

Studienbeschreibung

Vorbemerkungen

Arbeiten mit GLES Daten

Bei dieser Publikation und dem zugehörigen Datensatz handelt es sich um Daten der German Longitudinal Election Study (GLES), die von GESIS in Kooperation mit der Deutschen Gesellschaft für Wahlforschung (DGfW) herausgegeben werden. Vor der Veröffentlichung werden die Daten sorgfältig geprüft. Leider kann es trotz gründlicher Überprüfung der Daten passieren, dass Fehler in den Datensätzen unentdeckt bleiben. Fallen diese zu einem späteren Zeitpunkt auf, werden sie dokumentiert (Errata-Liste im Datenbestandskatalog, www.gesis.org/dbk) und zeitnah behoben.

Um hochqualitative Datensätze zur Verfügung stellen zu können, freuen wir uns über Ihre Mithilfe. Wenn Ihnen bei Ihrer Arbeit mit den GLES-Daten ein Fehler auffällt, helfen Sie uns sehr, wenn Sie eine kurze E-Mail an gles@gesis.org schicken. Bitte schicken Sie uns zusätzlich zu der Beschreibung des Fehlers auch die Studiennummer (ZA-Nummer) sowie die Versionsnummer des Datensatzes.

Wir empfehlen, stets mit der aktuellen Version der GLES-Daten zu arbeiten. Sie können diese unkompliziert über den Datenbestandskatalog herunterladen. Die Links zum direkten Download finden Sie auch auf den Seiten der GLES bei GESIS (www.gesis.org/gles).

Meldung von Veröffentlichungen

Um einen Überblick über die tatsächliche Nutzung der Daten zu erhalten, bitten wir Sie um eine kurze Mitteilung bei Veröffentlichungen, die Daten der GLES verwenden (bibliographische Angaben, Studiennummer des verwendeten Datensatzes). Veröffentlichungen, die vollständig oder teilweise auf Daten der GLES beruhen, werden in der offiziellen Bibliographie der GLES aufgeführt. Wenn es sich dabei um Konferenzpapiere o.ä. handelt, die nur schwer zugänglich sind, freuen wir uns über die Überlassung eines Exemplars bzw. eines PDF-Dokuments.

Kontakt

GESIS – Leibniz-Institut für Sozialwissenschaften

Postfach 122155

68072 Mannheim

E-Mail: gles@gesis.org

Zitation bei Veröffentlichungen

Wird in Publikationen auf Daten der GLES zurückgegriffen, bitten wir diese wie folgt zu zitieren:

Roßteutscher, Sigrid; Schmitt-Beck, Rüdiger; Schoen, Harald; Weißels, Bernhard; Wolf, Christof; Preißinger, Maria; Kratz, Agatha; Wuttke, Alexander; Manger, Lea (2019): Wahlkampf-Panel 2017 – Kontrollquerschnitt II (zu Welle 6) (GLES). GESIS Datenarchiv, Köln: ZA6806 Datenfile Version 2.0.0, doi: 10.4232/1.13422.

1 Studieninformationen

1.1. Studiennummer

ZA6806

1.2. Titel der Studie

German Longitudinal Election Study, Komponente 3: Wahlkampfpanel Kontrollquerschnitt II

1.3. Version

2.0.0, 11.12.2019, doi: 10.4232/1.13422

1.4. Erhebungszeitraum

04. September bis 13. September 2017

1.5. Primärforscher/innen

Prof. Dr. Sigrid Roßteutscher	(Goethe-Universität Frankfurt am Main)
Prof. Dr. Rüdiger Schmitt-Beck	(Universität Mannheim)
Prof. Dr. Harald Schoen	(Universität Mannheim)
Prof. Dr. Bernhard Weßels	(Wissenschaftszentrum Berlin für Sozialforschung)
Prof. Dr. Christof Wolf	(GESIS – Leibniz-Institut für Sozialwissenschaften)

1.6. Finanzierende Stelle

Deutsche Forschungsgemeinschaft e. V. (DFG)

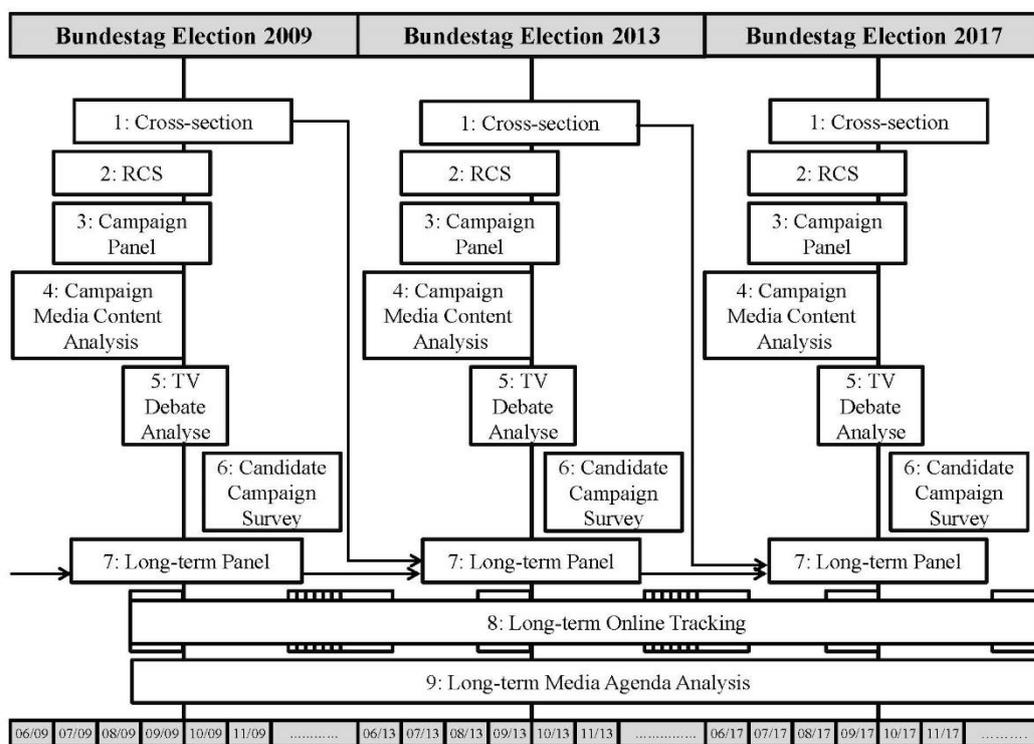
2 Konzeption

2.1. Hintergrund und Zielsetzung

Die German Longitudinal Election Study (GLES) ist eine großangelegte wissenschaftliche Wahlstudie, welche die gesellschaftlichen und politischen Einstellungen, Überzeugungen und Verhaltensweisen der wahlberechtigten Bürgerinnen und Bürger in Deutschland untersucht. Seit der Bundestagswahl 2009 beobachtet und analysiert die Studie, wie die Wählerschaft auf neue komplexe politische Konstellationen reagiert.

Die GLES erlaubt die Analyse des Wahlverhaltens in quer- als auch längsschnittlicher Perspektive sowie in Hinblick auf kurzfristige Dynamiken während des Wahlkampfs und auch langfristige soziale Wandlungsprozesse über den Wahlzyklus hinweg (Schmitt-Beck et al. 2010).

Abbildung 1: Das Design der German Longitudinal Election Study (GLES)



Das Wahlkampfpanel (Komponente 3) ermöglicht es, individuelle Veränderungen von politischen Einstellungen und Verhaltensweisen im Verlauf des Wahlkampfs zu untersuchen. Zu diesem Zweck wurde eine gleichbleibende Auswahl von Bürgern im Verlauf des Wahlkampfs und unmittelbar nach der Wahl in kurzen Abständen befragt. Eine Neuerung im Design des Wahlkampfpanels 2017 im Vergleich zu 2009 und 2013 besteht darin, dass bereits ein Jahr vor der Bundestagswahl mit der Befragung begonnen wurde. Dadurch können politische Einstellungen und Verhaltensweisen auch in Zeiträumen außerhalb des Wahlkampfs gemessen und mit jenen in der Wahlkampfphase verglichen werden.

2.2. Aufbau der Studie

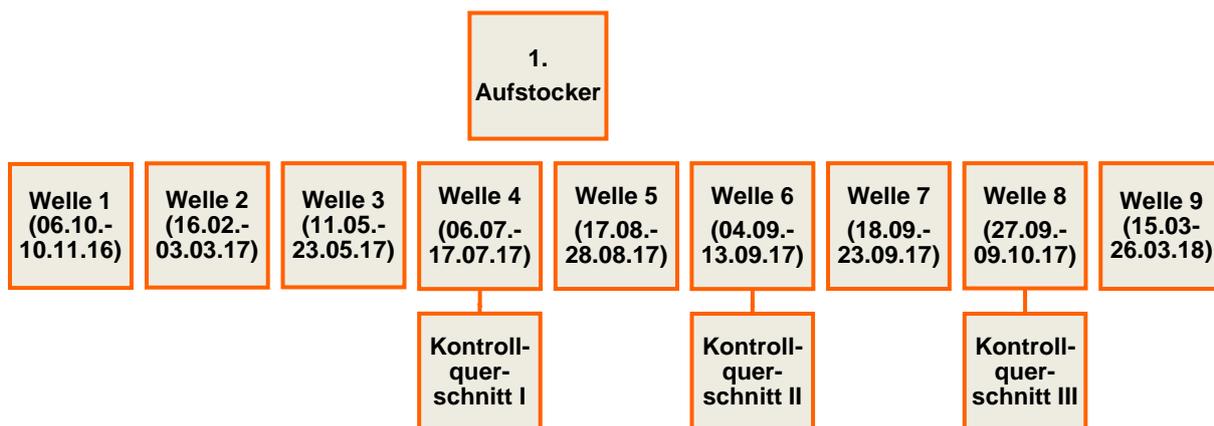
Insgesamt wurden bisher acht Wellen durchgeführt: Die ersten beiden Befragungen im Oktober 2016 und Februar 2017 sollten die politischen Einstellungen und Verhaltensweisen vor Beginn des Wahlkampfs erheben. Zwischen Mai und September 2017 wurden in kurzen Abständen fünf Wellen durch-

geführt, sodass mögliche Wahlkampfeffekte auf die Einstellungen der Befragten nachverfolgt werden können. Unmittelbar nach der Kanzlerwahl vom 14. März 2018 wurde die vorerst letzte Befragung durchgeführt.

Zusätzlich zu den Wiederholungsbefragten wurden zeitgleich zu der vierten, sechsten und achten Befragung drei unabhängige Stichproben mit etwa 1.000 Personen gezogen, die dann anhand eines fast identischen¹ Fragebogens interviewt wurden (siehe Abbildung 2). Durch den Vergleich von Panel- und Querschnittsbefragten lässt sich untersuchen, ob und in welchem Ausmaß die wiederholte Befragung derselben Personen Einfluss auf die Messung von politischen Einstellungen und Verhalten nimmt. Die vorliegende Studienbeschreibung bezieht auf den zweiten dieser Kontrollquerschnitte, der parallel zu Welle 6 des Panels erhoben wurde.

Zwischen der vierten und der fünften Welle wurde das Wahlkampfpanel mit neuen Teilnehmern aufgefrischt. In einer Profilstichprobe wurden gezielt jene Personengruppen eingeladen, die zu diesem Zeitpunkt die höchsten Ausfallquoten im Panel aufwiesen. Teilnehmer mit einem beendeten Interview in dieser Profilstichprobe wurden ab der fünften Welle des Wahlkampfpanels miteingeladen und mit den übrigen Befragten zusammen befragt. Da in Zukunft vermutlich noch weitere Auffrischungsstichproben gezogen werden, bezeichnen wir diese Gruppe im Folgenden als „1. Aufstocker“.

Abbildung 2: Die Wellen und Kontrollquerschnitte des Wahlkampfpanels 2017



¹ Der Fragebogen der Kontrollquerschnitte war etwas länger als der Fragebogen der jeweiligen Panelwelle, weil zusätzlich soziodemographische Merkmale erhoben wurden. Für die Panelbefragten waren diese Merkmale bereits aus früheren Wellen bekannt.

2.3. Inhalt der Studie

In den ersten beiden Befragungen wurden sozialstrukturelle Merkmale und psychologische Konzepte abgefragt, von denen angenommen werden kann, dass sie im Zeitverlauf weitgehend stabil sind. Als „Kernfragebogen“ wird der Bestand jener Fragen bezeichnet, der unverändert in jeder der Befragungen gestellt wurde. Dazu zählen beispielsweise Fragen zum Wahl- und Kommunikationsverhalten, Bewertungen von Parteien und Politikern oder als subjektiv wichtig wahrgenommene Sachfragen. Während der Wahlkampfphase haben die einzelnen Wellen zudem inhaltliche Schwerpunkte. Die Profiwelle der Aufstocker glich im Aufbau der ersten Welle und enthielt neben dem Kernfragebogen Fragen zu soziostrukturellen Merkmalen und psychologischen Konstrukten. Tabelle 1 gibt einen Überblick über die thematischen Schwerpunkte der bereits durchgeführten Befragungen.

Tabelle 1: Übersicht über die thematischen Schwerpunkte der Panelwellen

Inhalte	Welle 1 (06.10.- 10.11.16)	Welle 2 (16.02.- 03.03.17)	Welle 3 (11.05.- 23.05.17)	Profilwelle Auf- stocker(20.0 7.-09.08.17)	Welle 4 (06.07.- 17.07.17)	Welle 5 (17.08.- 28.08.17)	Welle 6 (04.09.- 13.09.17)	Welle 7 (18.09.- 23.09.17)	Welle 8 (27.09.- 09.10.17)	Welle 9 (15.03.- 26.03.18)
Kernfragebogen	X	X	X	X	X	X	X	X	X	X
Wahlverhalten, retrospektiv	X								X	X
Ideologie, Parteien		X			X			X		
Ideologie, Kandidaten			X				X			
Positionsissues, Ego	X	X	X	X	X		X	X	X	X
Positionsissues, Parteien		X			X			X		
Positionsissues, Kandidaten			X				X			
Kandidaten			X		X	X	X			
Strategisches Wählen			X			X	X	X	X	X
Wirtschaft	X		X			X	X		X	
Landtagswahl in Schleswig-Holstein			X							
Landtagswahl in NRW			X							
Landtagswahl in Niedersachsen						X	X	X	X	
TV-Duell							X	[X]		
Koalitionsverhandlungen										X
Psychologische Konstrukte	X			X						
Soziodemografie, Ego und Partner	X			X						
Soziodemografie, Migrationshintergrund		X	[X]							

Erläuterung: „X“ steht für die jeweilige Befragungswelle. [X] bedeutet, dass nur die Teilnehmer des Wahlkampfpanels die Fragen des Moduls erhalten haben, die bei der Abfrage dieses Themenblocks in einer der vorangegangenen Wellen nicht teilgenommen haben (siehe hierzu auch Kapitel 6.7).

3 Stichprobenziehung und Erhebung

3.1. Untersuchungsgebiet

Deutschland (DE)

3.2. Grund- und Auswahlgesamtheit

Die Grundgesamtheit des Kontrollquerschnittes zu der dritten Befragung des Wahlkampfpanels bildet die zur Bundestagswahl 2017 wahlberechtigte Bevölkerung. Da die Studie online durchgeführt wurde, hatten jedoch nicht alle wahlberechtigten Bürger die Chance, für die Befragung ausgewählt zu werden. Die Auswahlgesamtheit umfasst daher ausschließlich die bei der Bundestagswahl wahlberechtigten Bundesbürger des von GapFish betriebenen Online-Access-Panels. Dieses Panel umfasste 2017 in Deutschland 113.000 aktive durch Double-Opt-In-Verfahren registrierte Personen, deren Stammdatenangaben nicht älter als ein Jahr alt sind und die im letzten Jahr an mindestens einer Befragung teilgenommen haben.

Tabelle 2: Zusammensetzung des GapFish-Panels nach soziodemographischen Merkmalen (Anteil in Prozent)

	GapFish
Geschlecht	
Weiblich	55
Männlich	45
Bildung	
Niedrig (d.h. kein Abschluss, Haupt- und Volksschule)	16
Mittel (d.h. Realschule, Mittlere Reife)	34
Hoch (d.h. Abitur, Fachhochschulreife, Studium)	48
Altersgruppen	
18-29 Jahre	34
30-39 Jahre	22
40-49 Jahre	17
50-59 Jahre	16
60 und älter	10

Fehlend zu 100%: keine Angabe

Die Mitglieder des Online-Access-Panels werden von GapFish auf unterschiedlichen Wegen rekrutiert. Zu diesem Zweck tritt GapFish unter dem Markennamen „EntscheiderClub“ an die Personen heran. Dies erfolgt überwiegend online, und nutzt dafür beispielsweise Suchmaschinen-Marketing, E-Mail-Kampagnen, soziale Netzwerke oder Webseitenpartnerschaften.

GapFish gibt an, ein wirkungsvolles Qualitätsmanagement zu betreiben. Das Qualitätsmanagement sowie eine maßvolle Einladungsfrequenz sollen dazu beitragen, unerwünschte Effekte wie Panel-Verzerrungen oder Professionalisierung der Mitglieder zu vermeiden. Ein durchschnittliches Mitglied verweilt 24 Monate im GapFish-Panel. Innerhalb eines Monats werden ca. 10 Prozent der Panelmitglieder aufgrund von Qualitätskontrollen und Panel-Mortalität aus der Datenbank entfernt.

Die Mitglieder des Access-Panels werden von GapFish für die Teilnahme an Befragungen incentiviert. Die Befragten erhalten eine Incentivierung von 0,10 € pro Minute. Die Panel-Mitglieder sammeln durch Umfrageteilnahmen Gutschriften, die ab einem Betrag von 10 € ausbezahlt oder gespendet werden können.

3.3. Auswahlverfahren und Quotierung

Die Stichprobe für den Kontrollquerschnitt wurde durch eine Quotenauswahl aus der Grundgesamtheit der im Online-Access-Panel zur Verfügung stehenden Personen generiert. Dabei wurden nur volljährige Personen mit deutscher Staatsangehörigkeit berücksichtigt. Weiterhin wurde darauf geachtet, dass nur Mitglieder des GapFish-Panels zu dem Kontrollquerschnitt eingeladen werden, welche nicht schon an dem parallel laufenden Wahlkampfpanel teilnehmen.

Die Teilnehmer des Kontrollquerschnitts wurden nach den Merkmalen Geschlecht, Alter (in fünf Alterskategorien: 18-29, 30-39, 40-49, 50-59, 60+) und Bildung (niedrig: kein Abschluss, Haupt- oder Volksschule; mittel: Realschule, mittlere Reife; hoch: Abitur, Fachhochschulreife) kreuzquotiert. Jede Kombination dieser drei Merkmale sollte in einem etwa gleichen Anteil repräsentiert sein (siehe Tabelle 3). Um diese Quotenvorgaben möglichst genau zu erreichen, wurden die Mitglieder des GapFish-Online-Access-Panels in mehreren Stufen eingeladen.

Bei Feldbeginn wurde zunächst ein sogenannter „Soft-launch“ mit nur ca. 700 Mitgliedern des GapFish-Online-Access-Panels vorgenommen. Anhand der Antworten, die von diesen Personen gegeben werden, kann der ordnungsgemäße technische Ablauf und die Programmierung abschließend überprüft werden. Anschließend wurde die eigentliche Einladungswelle (Full-launch) gestartet.

Um den Quotenvorgaben zu genügen, wurden anfangs Personen mit Merkmalskombinationen überproportional stark eingeladen, die vergleichsweise schwer zu erreichen sind (z.B. ältere Personengruppen, Personen mit niedrigen Bildungsabschlüssen).

Je nach Bedarf bzw. bei noch offenen Quoten wurden Teilnehmer, welche die Umfrage noch nicht begonnen hatten, nach wenigen Tagen (siehe Tabelle 4) an die Einladung erinnert und erneut zur Teilnahme aufgefordert.

Tabelle 3: Quotierung im Wahlkampfpanel 2017 Kontrollquerschnitt II

Quotierungsmerkmale	Soll in Prozent	Ist* in Prozent
		N=1.173
Geschlecht		
Männlich	50,0	50,1
Weiblich	50,0	49,9
Bildung (exkl. Schüler)		
Niedrig (d.h. kein Abschluss, Haupt- oder Volksschule)	33,3	33,1
Mittel (d.h. Realschule, Mittlere Reife)	33,4	33,2
Hoch (d.h. Abitur, Fachhochschulreife, Studium)	33,3	33,7
Altersgruppe		
17-29 Jahre	20,0	19,9
30-39 Jahre	20,0	20,0
40-49 Jahre	20,0	20,3
50-59 Jahre	20,0	20,1
60 Jahre und älter	20,0	19,7

* Tatsächliche Verteilung im Datensatz

Tabelle 4: Statistik über die Rekrutierung

Datum	N	Geschlecht		Alter							Bildung			
		männlich	weiblich	<18	18-29	30-39	40-49	50-59	60+	k. A.	niedrig	mittel	hoch	Sonstiges
04.09.2017	1.244	643	601	0	457	282	223	166	116	0	353	441	450	0
05.09.2017	4.505	2.081	2.424	1	1.962	1.181	685	359	315	2	1.815	1.531	1.159	0
06.09.2017	4.111	1.870	2.241	6	2.510	901	436	151	106	1	2.127	1.187	797	0
07.09.2017	3.367	727	2.640	4	2.822	273	131	85	52	0	3.077	245	44	1
11.09.2017	278	2	276	0	266	9	3	1	2	0	1	276	1	0
Summe Einladungen	13.505													

*Die Informationen über die soziodemographische Verteilung der Einladungen wurde erst im Frühjahr 2018 von GapFish angefordert. Zu diesem Zeitpunkt waren nicht mehr alle Teilnehmer, die damals eingeladen wurden, samplebar. Tatsächlich belief sich die Zahl der Einladungen für den Kontrollquerschnitt II auf 14341 Stück.

Die Befragung war für die eingeladenen Mitglieder des GapFish-Panels vom 04.09.2017 bis zum 13.09.2017 geöffnet. In den ersten beiden Tagen nahmen bereits 50,4 Prozent der Stichprobe an der Umfrage teil (siehe Tabelle 5).

Tabelle 5: Teilnahmestatistik nach Tagen

	Datum	Begonnen	
		Relativ	Absolut
	Mo 04.09.2017	10,0%	117
	Di 05.09.2017	40,4%	474
	Mi 06.09.2017	32,8%	385
	Do 07.09.2017	13,5%	158
	Fr 08.09.2017	1,7%	20
	Sa 09.09.2017	0,5%	6
	So 10.09.2017	0,3%	4
	Mo 11.09.2017	0,8%	9
Gesamt		100,0%	1.173

3.4. Ausschöpfung

Systematischer Unit Nonresponse ist eine Fehlerquelle in Umfragestudien, die die Qualität der erhobenen Daten erheblich mindern kann.² Sind die in der Umfrage erhobenen Variablen mit der Teilnahmewahrscheinlichkeit korreliert, so sind Verzerrungen der Umfrageergebnisse („Nonresponse Bias“) die Folge. Die Angabe von Ausschöpfungsquoten („Response Rates“) erlaubt eine Einschätzung darüber, wie stark eine Umfragestudie von möglicherweise systematischer Nichtteilnahme betroffen ist. Die Angabe von standardisierten Response Rates und verwandten Indikatoren für das Teilnahmeverhalten der Personen in der Stichprobe sichert die Vergleichbarkeit dieser Maße über Studien hinweg. Für das Wahlkampf-Panel der GLES werden Indikatoren für das Teilnahmeverhalten nach den 2011 publizierten Standards der American Association for Public Opinion Research (AAPOR, <http://www.aapor.org>) sowie Empfehlungen zur Berechnung der Ausschöpfung von Callegaro und DiSogra (2008) angegeben. Bei Online-Befragungen ist die Interpretation von Response Rates in vielen Fällen problematisch, da in der Regel keine zufallsbasierten Stichprobenziehungsverfahren angewendet werden. Ein gängiger Ausweg ist daher, die Mitglieder von Online-Access-Panels zu befragen. Bei der Befragung von Teilnehmern aus selbstrekrutierten Online-Access-Panels, wie dem Online-Access-Panel von GapFish, ist es nicht sinnvoll, bei der Berechnung der Ausschöpfung von einer Response Rate zu sprechen. Die AAPOR empfiehlt daher für Befragungen von Teilnehmern aus selbstrekrutierten Online-Panels die Angabe einer „Participation Rate“, die sich als Anteil der verwertbaren Interviews an der absoluten Anzahl aller versendeten persönlichen Befragungseinladungen berechnet. Wir greifen diese Empfehlung hier auf und bestimmen die Participation Rate als Anzahl aller vollständigen und partiell voll-ständigen Interviews geteilt durch die Anzahl aller Einladungen zur Befragung (vollständige (I) und partiell vollständige (P) Interviews, abgebrochene Interviews (RBreakoff), Nichtkontakte (NC) und andere Fälle (O) sowie alle Fälle, bei denen unklar ist, ob sie die Einladung erhalten und gesehen haben (UH & UO)).

²Dieser Abschnitt orientiert sich stark an den Ausführungen zur Ausschöpfung im Technical Report zum Wahlkampf-Panel 2009. Siehe 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, S.25f.

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

Nicht berücksichtigt werden bei dieser Bestimmung alle Fälle, die zwar zur Teilnahme eingeladen, auf Grund der Quotierung jedoch als nicht teilnahmeberechtigt klassifiziert wurden.

Als weiteren informativen Indikator berichten wir den Anteil der Befragungsabbrüche in der Online-Befragung. Hierfür wird die „Breakoff Rate“⁷ berechnet als

$$\text{Breakoff Rate} = \frac{R_{\text{Breakoff}}}{(I + P) + R_{\text{Breakoff}}}$$

wobei R_{Breakoff} die Anzahl der Befragungsabbrüche, I die Anzahl der vollständigen Interviews und P die Anzahl der partiellen Interviews ist. Die Breakoff Rate gibt mit anderen Worten den Anteil der Befragungsabbrüche an allen begonnen Interviews wieder.

Tabelle 6: Ausschöpfung im Kontrollquerschnitt II des Wahlkampfpanels 2017

Code	Beschreibung	Anzahl
I + P	Vollständige und partiell vollständige Interviews	1.173
R_{Breakoff}	Abbruch nach der Verifizierung	156
UH	Unklare Teilnahmeberechtigung	12.435
NE	Quote voll bzw. in Welle 1 ausgescreent	577
Anzahl der versandten Befragungseinladungen		14.341
PR	Participation Rate in %	8,2
BR	Breakoff Rate in %	11,7

3.5. Interviewdauer

Mit einer Medianzeit von 23:18 Minuten dauerte die Beantwortung des Kontrollquerschnitts etwas länger als die parallel abgefragte (hier nicht ausgewiesene) Welle 6 im Wahlkampfpanel (vgl. Tabelle 7). Wie aus den Angaben zu den Minimal- und Maximalwerten ersichtlich wird, gab es beträchtliche Variation in den Befragungszeiten. Relativ unkritisch sind dabei sehr lange Befragungszeiten, die vermutlich durch kurzzeitige Interview-Unterbrechungen hervorgerufen wurden (längere Unterbrechungen, die eine automatische Abmeldung von der Umfrage zur Folge hatten, sind in der Statistik nicht enthalten). Wesentlich problematischer sind hingegen sehr schnelle Interviews, weil diese darauf hindeuten, dass sich einige Befragte lediglich durch die Umfrage geklickt haben, ohne die Fragen sorgfältig gelesen zu haben.

Tabelle 7: Interviewlänge im Kontrollquerschnitt II des Wahlkampfpanels 2017

	N	Minimale Befra- gungs- dauer	Maximale Befragungsda- uer	Arithmetisches Mittel	Median
Bearbeitungsdauer *	1.153	00:05:03	03:04:33	00:26:18	00:23:18

* Die Bearbeitungsdauer wird nur für diejenigen Befragten ausgewiesen, die die Befragung weder unter- noch abgebrochen haben. Wurde eine Befragung unterbrochen, ohne jedoch den Browser zu schließen, wurde dies nicht als Unterbrechung gewertet.

4 Variablen im Datensatz

Der Datensatz des Kontrollquerschnitts enthält verschiedene Variablen-Typen:

- *GESIS Archiv-Variablen* beschreiben den Datensatz und seine Erstellung. Sie umfassen Variablen, die für die Archivierung und Distribution der Studie benötigt werden, z.B. die Studiennummer, die Versionskennung und die Feldzeit der Erhebung.
- *Administrative Variablen*: Variablen zur Teilnahme der Befragten im Wahlkampfpanel (z.B. AAPOR-Codes oder Dummies, die Auskunft über die Teilnahme an bestimmten Wellen geben).
- *Paradaten* dokumentieren den technischen Ablauf des tatsächlichen Interviewprozesses. Hierzu zählt beispielsweise die technische Ausstattung der Befragten (z.B. Browser-Version, Javascript-Version).
- *Datenqualitätsvariablen* sind Indikatoren, die potenziell dazu verwendet werden könnten, um die Qualität der gegebenen Antworten eines Befragten zu beurteilen.
- *Gewichtungsvariablen* umfassen Querschnitts- und Panelgewichte.
- *Inhaltliche Variablen*: Die eigentlich interessierenden Merkmale der Befragten, die anhand des Fragebogens erhoben wurden.
- *Kontextvariablen* geben Auskünfte über den regionalen Kontext der Befragten.
- *Zeitvariablen* geben die benötigte Zeit pro Seite wieder

Einen Überblick über die Gesamtzahl der Variablen in den einzelnen Wellen gibt Tabelle 8. In den folgenden Unterkapiteln wird der Inhalt der verschiedenen Variablentypen genauer beschrieben.

Tabelle 8: Anzahl der Variablen im Kontrollquerschnitt II des Wahlkampfpanels 2017

Variablentyp	Anzahl
GESIS-Archiv-Variablen	7
Administrative Variablen	4
Paradaten	10
Datenqualitätsvariablen	3
Gewichtungsvariablen	4
Inhaltliche Variablen	314
Kontextvariablen	6
Zeitvariablen	262
Gesamt	610

4.1. Metadaten

Tabelle 9: Archivvariablen von GESIS

Variablenname	Erläuterung
study	Gibt die (ZA-)Studennummer des Datensatzes in vierstelligem Format an, unter der die Studie im Archivsystem der GESIS abgelegt und zu finden ist
version	Version des Datensatzes, beginnend mit 1.0.0
doi	Digital Object Identifier
field_start	Beginn der Feldzeit der Studie
field_end	Ende der Feldzeit der Studie
glescomp	Zuteilung zu einer der elf Komponenten der GLES. Hier Komponente 3 für das Wahlkampfpanel
sample	Differenzierung zwischen den verschiedenen Stichproben, die innerhalb einer GLES-Studie untersucht wurden.

4.2. Paradata

Tabelle 10: Paradata zum technischen Interviewverlauf

Variablenname	Erläuterung
kp6_browser	Gibt den Namen des Browsers (User Agent) so an, wie er an den EFS Survey-Server übertragen wurde. Beispiele finden Sie z.B. unter http://de.wikipedia.org/wiki/User_Agent
kp6_javascript	Enthält das Resultat der Überprüfung auf JavaScript, die optional bei Umfragebeginn ausgeführt werden kann (0=JavaScript abgeschaltet; sonst=Versionsnummer, z.B. 10 für JavaScript 1.0)
kp6_flash	Enthält das Resultat der Überprüfung auf das Adobe Flash-Plugin, die optional bei Umfragebeginn ausgeführt werden kann und insbesondere für die Flash-Fragetypen benötigt wird (0=Flash-Plugin nicht vorhanden; sonst=Versionsnummer, z.B. 800 für Flash-Plugin 8.0)
kp6_datetime	Datum und Uhrzeit des Umfragebeginns, d.h. des Aufrufs der ersten Fragebogenseite. Die Angaben zur Uhrzeit im Datensatz entsprechen der Greenwich Mean Time (GMT).
kp6_date_of_last_access	Datum und Uhrzeit des letzten Zugriffs auf die Umfrage. Die Angaben zur Uhrzeit im Datensatz entsprechen der Greenwich Mean Time (GMT).
kp6_lastpage	Gibt die letzte vom Teilnehmer abgeschickte Seite an. D.h. wenn der Teilnehmer die Umfrage abgebrochen hat, erfolgte der Abbruch auf der Folgeseite.
lfdn	Jeder Teilnehmer des Fragebogens bekommt eine laufende Nummer.
lfdn09	Laufende Nummer aus dem Wahlkampfpanel 2009, sofern daran teilgenommen wurde
lfdn13	Laufende Nummer aus dem Wahlkampfpanel 2013, sofern daran teilgenommen wurde
kp6_duration	Die Bearbeitungsdauer, d.h. die Zeit, die zwischen dem ersten Zugriff und dem letzten Zugriff des Probanden auf den Fragebogen abgelaufen ist. Unterbricht ein Pro-

band den Fragebogen und nimmt die Bearbeitung später wieder auf (Dispositions-codes 23, 32), so wird `kpX_duration= -99` gesetzt, da eine sinnvolle Berechnung nicht möglich ist.

4.3. Administrative Variablen

Tabelle 11: Administrative Variablen

Variablenname	Erläuterung
<code>kp6_dispcode</code>	Dispositionscode des Teilnehmers, der seinen Status im Feld angibt, also ob der Teilnehmer z.B. die Umfrage bereits begonnen oder bereits fertiggestellt hat.
<code>partner</code>	Befragter stammt aus dem GapFish-Panel
<code>kp6_dropout</code>	Vollständigkeit des Interviews (0=vollständig; 1=unvollständig)
<code>kp6_interrupt</code>	Unterbrechung der Befragung (0=keine Unterbrechung; 1= mit Unterbrechung)

4.4. Kontextvariablen: Zuordnung von Wahlkreisen

In den Kontrollquerschnitten sind Informationen zum Wohnsitz in Ost- oder Westdeutschland sowie Informationen zum jeweiligen Wahlkreis eines Befragten enthalten.

Die Zuordnung zu Ost- und Westdeutschland (`kp6_ostwest`) erfolgte auf der Grundlage der Informationen über das Bundesland, in dem die Befragten ihren Hauptwohnsitz haben. Da in der Umfrage keine Unterscheidung zwischen Ost- und Westberlin vorgenommen wurde, wurden die in Berlin lebenden Befragten der DDR zugeordnet.

Die Zuordnung der Befragten zu einem Wahlkreis erfolgte auf der Grundlage der von ihnen genannten Postleitzahl des Hauptwohnsitzes. Auf Grundlage der Korrespondenzlisten des Bundeswahlleiters nimmt das Stata-Ado „`plztowknr`“³ eine Zuordnung von Postleitzahlen zu Wahlkreisen (Gebietsstand 2017) vor.

Bei der Verwendung der Wahlkreisinformationen ist zu beachten, dass eine Postleitzahl nicht immer eine eindeutige Zuordnung zu einem Wahlkreis erlaubt. Vor allem in großstädtischen Gebieten kann einer Postleitzahl im Extremfall bis zu fünf Wahlkreisen zugeordnet werden. Sofern eine eindeutige Zuordnung möglich war – in 1.106 Fällen (94,3%) wurde der Wahlkreis in der Variable „`kp4_elecdist`“ gespeichert. Waren mehrere Einordnungen möglich – in 52 Fällen (4,4%) – so sind die bis zu vier möglichen Wahlkreise in den Variablen „`kp6_elecdist1`“-„`kp6_elecdist4`“ enthalten. Die restlichen 15 Befragten (1,3%) gaben keine gültige Postleitzahl an oder verweigerten die Angabe.

³ Glinitzer, K., Gummer, T., Kaukal, M., Roßmann, J. (2018): `plztowknr`: Stata module to translate German zip codes into electoral districts (Version: 1.0) [Computer Software]. Chestnut Hill, MA: Boston College.

4.5. Datenqualitätsvariablen

Tabelle 12: Datenqualitätsvariablen

Variablenname	Erläuterung
kp6_speederindex	Setzt die Beantwortungsdauer eines Befragten ins Verhältnis zu allen anderen Befragten. Siehe auch Abschnitt 6.2 „Schnelle Antworten“.
kp6_4240	Bewertung der Umfrage
kp6_4270s	Kommentarfeld, in das Befragte ihre Anmerkungen zur Umfrage eintragen konnten

4.6. Gewichtungsvariablen

Anhand von Querschnittsgewichten kann die Verteilung bestimmter Variablen eines Datensatzes an bekannte Verteilungen in der Grundgesamtheit angepasst werden. Dabei wird davon ausgegangen, dass zumindest ein schwacher Zusammenhang zwischen den Anpassungsvariablen und den inhaltlichen Untersuchungsvariablen besteht. Beim Einsatz von Gewichtungsvariablen ist Vorsicht geboten: Für jede spezifische Analyse muss die Forscherin oder der Forscher selbst entscheiden, ob eine der mitgelieferten Gewichtungsvariablen dem Untersuchungszweck dienlich ist oder nicht.

Zur Berechnung der Querschnittsgewichte der GLES wurden sowohl sozial- als auch regionalstrukturelle Merkmale ausgewählt. Bei der Berechnung der Querschnittsgewichte für die GLES wurde auf das iterative Anpassungsverfahren („iterative proportional fitting“, IPF) zurückgegriffen (Deming/Stephan 1940). Bei der IPF-Gewichtung wird die Ist-Verteilung der einzelnen Zellen schrittweise an die jeweilige Soll-Verteilung der Gewichtungsvariablen angepasst. Bei diesem schrittweisen Prozess der Anpassung (Iteration) bilden die nach jedem Anpassungsschritt berechneten Gewichtungsfaktoren die Ausgangsbasis für die Anpassung des nächsten Merkmals. Der Anpassungsprozess endet, wenn die Differenz zwischen der gewichteten Randverteilung aller Faktoren und der Soll-Verteilung das Abbruchkriterium von 0,05 unterschreitet.⁴ Um sehr große Gewichtungsfaktoren zu vermeiden, werden die Faktoren (nach jedem Iterationsschritt) auf den fünffachen Mittelwert der Gewichtungsvariable (also fünf) getrimmt.⁵

Tabelle 13: Übersicht Gewichtungsvariablen

Gewicht	Variablenname
Sozial- und regionalstrukturelles Gewicht (Anpassung MZ2016)	wei_mzz_kq6
Sozial- und regionalstrukturelles Gewicht (Anpassung MZ2016, ohne Zeitunterschreiter)	wei_mzoz_kq6
Sozial- und regionalstrukturelles Gewicht (Anpassung Onliner)	wei_onz_kq6
Sozial- und regionalstrukturelles Gewicht (Anpassung Onliner, ohne Zeitunterschreiter)	wei_onoz_kq6

Bei der Erstellung der Querschnittsgewichte für das Wahlkampfpanel wurde, basierend auf den Ist-Verteilungen der ersten Panelwelle, zum einen an die Randverteilungen des (N)Onliner Atlas 2014⁶ und zum anderen an die Verteilungen des Mikrozensus 2016 angepasst. Bei der Erstellung der Soll-

⁴ Die Querschnittsgewichte wurden mit dem Statistikprogramm Stata berechnet, wobei auf das Ado „ipfweight“ von Michael Bergmann zurückgegriffen wurde.

⁵ Dieses Vorgehen findet auch bei der Berechnung der Gewichte der American National Election Study (ANES) Anwendung (siehe dazu: DeBell et al. 2009).

⁶ Die Randverteilungen (für Personen ab 18 Jahren) aus dem (N)Onliner Atlas wurden für die Gewichtung freundlicherweise von TNS Infratest zur Verfügung gestellt.

verteilungen des Mikrozensus 2016 ging nur die wahlberechtigte Bevölkerung in Privathaushalten am Ort der Hauptwohnung ein.⁷

Angepasst wurde an die soziodemographischen und regionalstrukturellen Merkmale: Geschlecht, Alter, Bildung und alte bzw. neue Bundesländer (inkl. Berlin). Das Alter wurde in vier Gruppen kategorisiert: „18 bis unter 30 Jahre“, „30 bis unter 45 Jahre“, „45 bis unter 60 Jahre“ und „60 Jahre und älter“. Das Merkmal Bildung wurden in drei Gruppen unterteilt:

- niedrige Bildung: Schule beendet ohne Abschluss, Hauptschulabschluss, Volksschulabschluss, bin noch Schüler/in⁸
- mittlere Bildung: Realschulabschluss, Mittlere Reife, Fachschulreife oder Abschluss der polytechnischen Oberschule 10. Klasse
- hohe Bildung: Fachhochschulreife (Abschluss einer Fachoberschule etc.), Abitur bzw. erweiterte Oberschule mit Abschluss 12. Klasse (Hochschulreife)

Tabelle 14: Ist- und Soll-Verteilungen der zur Erstellung der Gewichte verwendeten Variablen nach Stichprobe

Merkmal	Ist (in Prozent)	Soll (in Prozent)	
	N=1.173	Mikrozensus 2016	(N)Onliner Atlas 2014
Geschlecht			
männlich	50,1	48,6	52,6
weiblich	49,9	51,4	47,4
Altersgruppe			
18 bis unter 30 Jahre	18,2	15,9	21,3
30 bis unter 45 Jahre	28,6	20,7	27,3
45 bis unter 60 Jahre	32,7	29,0	32,9
60 Jahre und älter	20,6	34,4	18,5
Bildung			
niedrig	33,1	37,6	34,9
mittel	33,3	30,4	31,3
hoch	33,7	32,0	33,8
Bundesland			
Alte Bundesländer	82,2	79,1	80,4
Neue Bundesländer (inkl. Berlin)	17,8	20,9	19,6

Die Berechnung der GewichtungsvARIABLEN nach Stichprobe wurde jeweils nach der dritten Iteration abgebrochen, da die Differenz zwischen den gewichteten Istwerten in der Stichprobe und den Sollwerten des Mikrozensus bzw. (N)Onliner Atlas das Abbruchkriterium von 0,05 unterschritt. In allen Fällen war keine Trimmung notwendig. Eine Übersicht über alle berechneten Gewichte bietet die nachfolgende Tabelle.

⁷ Bei den vom Mikrozensus ausgewiesenen Werten handelt es sich um absolute Personenzahlen nach gebundener Hochrechnung.

Tabelle 15: Übersicht über die Gewichtungsfaktoren

	N	Mean	Std.Dev	Min	Max	1.Q	Median	3.Q	Max./Min. ⁹
wei_mzz_kq6	1.173	1	0,37	0,61	2,07	0,77	0,85	1,03	3,39
wei_mzoz_kq6	1.055	1	0,33	0,65	1,94	0,76	0,86	1,11	2,99
wei_onz_kq6	1.173	1	0,13	0,76	1,35	0,90	0,97	1,10	1,78
wei_onoz_kq6	1.055	1	0,16	0,70	1,41	0,88	0,99	1,10	2,01

4.7. Inhaltliche Variablen

Tabelle 16: Inhaltliche Variablen

Variablenname	Erläuterung
kp6_screen1	Deutsche Staatsbürgerschaft
kp6_2280	Geschlecht
kp6_2291	Geburtsmonat
kp6_2290	Geburtsjahr
kp6_2320	Schulabschluss
kp6_2601	Bundesland
kp6_010	Politisches Interesse, allgemein
kp6_170	Wahlbeteiligungsabsicht
kp6_190aa	Beabsichtigte Stimmabgabe: Erststimme (Version A)
kp6_190ab	Beabsichtigte Stimmabgabe: Erststimme (Version B)
kp6_190ba	Beabsichtigte Stimmabgabe: Zweitstimme (Version A)
kp6_190bb	Beabsichtigte Stimmabgabe: Zweitstimmen (Version B)
kp6_191aa	Stimmabgabe Briefwähler: Erststimme (Version A)
kp6_191ab	Stimmabgabe Briefwähler: Erststimme (Version B)
kp6_191ba	Stimmabgabe Briefwähler: Zweitstimme (Version A)
kp6_191bb	Stimmabgabe Briefwähler: Zweitstimme (Version B)
kp6_260_c1-c3	Gründe für Wahlentscheidung (1. Codierung – 3. Codierung)
kp6_260s	Gründe für Wahlentscheidung, offen
kp6_2751a,c,g,i	Consideration Set am Tag der Briefwahl
kp6_330	Schwierigkeit der Wahlentscheidung
kp6_221a,c,g,i	Consideration Set
kp6_211 a,c,g,i	Consideration Set, hypothetisch
kp6_380	Wichtigkeit Wahlausgang
kp6_3490	Wahl-O-Mat, Nutzung
kp6_3500a-c	Wahl-O-Mat, Grund der Nutzung
kp6_3510	Wahl-O-Mat, Übereinstimmung
kp6_730	Skalometer Regierung
kp6_731a-c	Leistungen Regierungsparteien

⁹ Der Wert Max/Min gibt das Verhältnis zwischen dem höchsten und niedrigsten Gewichtungsfaktor an. Im Idealfall werden die Gewichte weder besonders groß noch sehr klein, so dass ein niedriger Wert positiv zu betrachten ist.

kp6_840_c1-c5	Wichtigstes Problem, codiert (1. Codierung - 5. Codierung)
kp6_840s	Wichtigstes Problem
kp6_850a	Lösungskompetenz: Wichtigstes Problem (Version A)
kp6_850b	Lösungskompetenz: Wichtigstes Problem (Version B)
kp6_860_c1-c5	Zweitwichtigstes Problem, codiert (1. Codierung - 5. Codierung)
kp6_860s	Zweitwichtigstes Problem
kp6_870a	Lösungskompetenz: Zweitwichtigstes Problem (Version A)
kp6_870b	Lösungskompetenz: Zweitwichtigstes Problem (Version B)
kp6_650a,z,t,c,s,w,y,p	Skalometer Politiker
kp6_671	Kanzlerpräferenz
kp6_662a,b	Ambivalenz Politiker, negative Gefühle
kp6_663a,b	Ambivalenz Politiker, positive Gefühle
kp6_1800	TV-Duell: Rezeption
kp6_1810; kp6_1820	TV-Duell: wahrgenommener Ausgang Merkel/Schulz
kp6_1870	TV-Duell: Rezeption des Medienurteils
kp6_1880, kp6_1890	TV-Duell: wahrgenommenes Medienurteil Merkel/Schulz
kp6_1900	TV-Duell: Rezeption des Urteils des sozialen Umfelds
kp6_1910, kp6_1920	TV-Duell: Wahrgenommenes Urteil des sozialen Umfelds zu Merkel/Schulz
kp6_2880ad,aa,w,k	Ego-Positionen I
kp6_430a-f,i	Skalometer Parteien
kp6_680a,b	Links-Rechts-Einstufung, Spitzenkandidaten
kp6_1500	Links-Rechts-Selbsteinstufung
kp6_1080a,b	Sozio-ökonomische Dimension, Kanzlerkandidaten
kp6_1090	Sozio-ökonomische Dimension, Ego
kp6_1120a,b	Zuzugsmöglichkeiten Ausländer, Kanzlerkandidaten
kp6_1130	Zuzugsmöglichkeiten Ausländer, Ego
kp6_1392a,b	Sicherheit und Privatsphäre, Kanzlerkandidaten
kp6_1411	Sicherheit und Privatsphäre, Ego
kp6_2880j,ae,y,ac	Ego-Positionen II
kp6_1230a-f,i	Europäische Integration, Parteien
kp6_1240a,b	Europäische Integration, Kanzlerkandidaten
kp6_1250	Europäische Integration, Ego
kp6_780	Eigene wirtschaftliche Lage, aktuell
kp6_770	Verantwortlichkeit eigene wirtschaftliche Lage
kp6_820	Allgemeine wirtschaftliche Lage, aktuell
kp6_810	Verantwortlichkeit allgemeine wirtschaftliche Lage
kp6_3430p,m,r	Politisches Wissen: Zuordnung Politiker/Parteien
kp6_1600	Internetnutzung, politisch aktuell
kp6_1681a-f	Fernsehnutzung, Nachrichten aktuell
kp6_1681flag	Fernsehnutzung, Nachrichten aktuell, Auskunftszeitpunkt
kp6_1661a-h	Printmediennutzung, politisch aktuell
kp6_1661flag	Printmediennutzung, politisch aktuell, Auskunftszeitpunkt
kp6_1701aa-dc	Nachrichtenmagazine, aktuell
kp6_1702	Nachrichtenmagazine, SPIEGEL Online
kp6_1933	Politische Gespräche, allgemein
kp6_1933flag	Politische Gespräche, allgemein, Auskunftszeitpunkt

kp6_1937a-c	Namensgenerator
kp6_1937	Anzahl der Gesprächspartner
kp6_1941a-c	Beziehung Gesprächspartner
kp6_1945a-c	Gesprächshäufigkeit
kp6_1945flag	Gesprächshäufigkeit, Auskunftszeitpunkt
kp6_1971aa-ca	Wahlabsicht Gesprächspartner (Version A)
kp6_1971ab-cb	Wahlabsicht Gesprächspartner (Version B)
kp6_1555a-e	Politische Motivation
kp6_2880x,z,ab	Ego-Positionen III
kp6_3430q,c,l	Politisches Wissen: Zuordnung Politiker/Parteien
kp6_910a-c,e-h	Koalitionsskalometer
kp6_962a-c,e-h	Koalitionserwartungen
kp6_911	Koalitionsbeteiligung AfD
kp6_1050	Wahrnehmung Meinungsumfragen
kp6_1051	Glaubwürdigkeit Meinungsumfragen
kp6_3010d,e,f,i	Einzugswahrscheinlichkeit kleine Parteien
kp6_390	Interesse am Wahlkampf
kp6_421ka-ei,y	Parteikontakte I
kp6_421flag1	Parteikontakte I, Auskunftszeitpunkt
kp6_421ha-ii,z	Parteikontakte II
kp6_421flag2	Parteikontakte II, Auskunftszeitpunkt
kp6_2250	Gerechte Gesellschaftsordnung
kp6_2270	Gerechtigkeit, Ego
kp6_090	Politisches Wissen: 5%-Klausel
kp6_090_v1	Politisches Wissen: 5%-Klausel (Dummy Richtig/Falsch)
kp6_110	Politisches Wissen: Erst-/Zweitstimme
kp6_130	Politisches Wissen: Wahlrecht D
kp6_2090a	Parteiidentifikation (Version A)
kp6_2090b	Parteiidentifikation (Version B)
kp6_2100	Parteiidentifikation, Stärke
kp6_2095a	Multiple Parteiidentifikation (Version A)
kp6_2095b	Multiple Parteiidentifikation (Version B)
kp6_2101	Multiple Parteiidentifikation, Stärke
kpx_2441	Haushaltsgröße
kpx_2450	Anzahl Haushaltsmitglieder unter 18 Jahre
kpx_2301	Familienstand
kpx_2311	Partner vorhanden
kpx_2312	Partner im Haushalt
kpx_2330	Berufliche Bildung
kpx_2340	Erwerbstätigkeit
kpx_2350	Frühere Erwerbstätigkeit
kpx_2380	Beruf
kpx_3610	Angestellter - Differenzierung
kpx_3620	Arbeiter - Differenzierung
kpx_3630	Akademisch freier Beruf, Selbständiger - Differenzierung
kpx_3640	Beamter - Differenzierung

kpx_3650	Beschäftigungssektor
kpx_3660	Wirtschaftssektor
kpx_2390	Früherer Beruf
kpx_3710	Früherer Beruf Angestellter - Differenzierung
kpx_3720	Früherer Beruf Arbeiter - Differenzierung
kpx_3730	Früherer Beruf Akademisch freier Beruf, Selbständiger - Differenzierung
kpx_3740	Früherer Beruf Beamter - Differenzierung
kpx_3750	Früherer Beruf - Beschäftigungssektor
kpx_3760	Früherer Beruf - Wirtschaftssektor
kpx_2590	Nettoeinkommen HH, mit Kategorien
kpx_2580	Schichtzugehörigkeit, subjektiv
kpx_2481	Religionszugehörigkeit
kpx_2491	Häufigkeit Gottesdienst
kpx_3910	Religiosität
kpx_2600	Wohnort

4.8. Unveröffentlichte Variablen

Bei der Durchführung von Online-Umfragen werden Variablen erhoben oder an GESIS übermittelt, die zwar für die Durchführung der Online-Umfrage relevant sind, aber keine inhaltliche Aussagekraft haben. Diese Variablen werden nicht veröffentlicht, können aber bei Bedarf von GESIS (gles@gesis.org) bezogen werden. Ebenso können aus datenschutzrechtlichen Gründen einige erhobenen Variablen nicht zum freien Download zur Verfügung gestellt werden (wie z.B. die Postleitzahl). Bei diesen Variablen ist ein Bezug unter bestimmten Bedingungen möglich.

Tabelle 17: Übersicht über die nicht veröffentlichten Variablen und die Bezugsmöglichkeit

Variablenname	Erläuterung	Bezugsmöglichkeit
Kp6_ats	Absoluter Zeitstempel	auf Anfrage erhältlich
Kp6_page_history	Seitenhistorie	auf Anfrage erhältlich
Kp6_2602	Postleitzahl	auf Anfrage erhältlich
Kp6_xxxx_org bzw. _c1-2	Sämtliche Originalvariablen und Nachfragevariablen, die bei der Zusammenfassung der Nachfragvariablen, und der Parteiversionen zur Sicherung der ursprünglichen Daten erstellt wurden bzw. nicht mehr benötigt werden	auf Anfrage erhältlich
Stammdaten		
s0	Geschlecht	auf Anfrage erhältlich
s1	Geburtstag	auf Anfrage erhältlich
s2	Geburtsmonat	auf Anfrage erhältlich
s3	Geburtsjahr	auf Anfrage erhältlich
s5	Schulische Ausbildung	auf Anfrage erhältlich

5 Datenaufbereitung

5.1. Vorbemerkung

Die Datenaufbereitung der Kontrollquerschnitte des Wahlkampfpanels erfolgte analog zur Aufbereitung der jeweiligen Welle des Panels. Alle Aufbereitungsschritte folgten der Richtlinie einer minimal-invasiven Bearbeitung. Die Aufbereitung wurde mittels Syntax durchgeführt und lässt sich somit in all ihren Schritten nachvollziehen.

5.2. Codierung der fehlenden Werte

Fehlende Werte wurden bei allen Variablen nach dem einheitlichen Codierschema der GLES vergeben, welches im SPSS-Format negative Werte im Bereich -71 bis -99 und in Stata-Codes von .a bis .p vorsieht. Den Datensätzen liegt ein Do-File bei, anhand dessen die SPSS Missing Codes im Stata-Datensatz in die vorgesehenen Stata-Codes transformiert werden können.

Tabelle 18: Kurzübersicht über die Missing-Value-Codes der GLES

SPSS-Code	Stata-Code	Label
-71	.p	nicht bekannt*
-72	.o	nicht einzuschätzen*
-84	.k	keine Erst-/Zweitstimme abgeben
-85	.j	nicht wählen
-92	.h	Fehler in Daten
-93	.g	Interview abgebrochen
-95	.e	nicht teilgenommen
-97	.c	trifft nicht zu
-98	.b	weiß nicht
-99	.a	keine Angabe

*Die Variablen elecdist – elecdist4 sind gelabelt als -71 „unbekannte PLZ und -72 „nicht-eindeutige Zuordnung PLZ“.

5.3. Codierung der Parteien

Zur besseren Vergleichbarkeit der GLES Komponenten untereinander erfolgt die Codierung der politischen Parteien bei allen Fragen, die eine geschlossene oder offene Abfrage zu politischen Parteien beinhalteten, nach einem einheitlichen Codierschema. Das Codierschema kann auf den Seiten der GLES bei GESIS abgerufen werden (<http://www.gesis.org/gles>). Alle Angaben zu politischen Parteien werden in zwei Variablen veröffentlicht, die mit Version A und Version B gekennzeichnet sind. Version A enthält nur die im Bundestag vertretenen Parteien der Legislaturperiode 2013 bis 2017 sowie die FDP, AfD und die Kategorie „andere Partei“. Version B unterscheidet sich von Version A in einer detaillierten Ausweisung auch kleinerer Parteien (Freie Wähler, NPD, oedp, Piraten, REP). Diese Parteien konnten von den Befragten im Rahmen einer Nachfrage nur dann ausgewählt werden, wenn sie zuvor auf die Frage nach einer Partei – zur Auswahl standen zunächst nur die im Bundestag vertretenen Parteien – auf „andere Partei“ klickten.

Einige wenige Befragte haben bei der Nachfrage, bei denen sie zwischen den kleineren Parteien auswählen sollten, keine Angabe gemacht oder die Umfrage abgebrochen. Diese Personen erhielten sowohl in Version A als auch in Version B den Code -99 (keine Angabe) bzw. -93 (Umfrage abgebro-

chen), obwohl in diesen Fällen eigentlich eine partielle inhaltliche Information vorlag: Sie hatten auf der ersten Seite eine „andere Partei“ angeklickt. Falls man diese ursprüngliche Information wieder herstellen möchte, sind im Folgenden die laufenden Nummern der Befragten im Datensatz dokumentiert (Variable: lfdn), bei denen die hier beschriebene Codierung vorgenommen wurde.

Da bei der Frage nach der Stärke der Parteiidentifikation („kp6_2100“) nur danach gefiltert wurde, ob sie bei der Parteiidentifikation („kp6_2090“) eine Angabe gemacht haben, tritt nun für die Befragten, bei denen der Wert in -99 umkodiert wurde, der Fall auf, dass sie Werte bei „kp6_2100“ aufweisen. Um die Filterung konsistent zu gestalten, werden die Angaben dieser Fälle in -97 „trifft nicht zu“ umkodiert.

5.4. Markierung der Wissensfragen

Bei der Fragebatterie kp6_3430p, m, r, q, c, l wurde das politische Wissen der Umfrage-Teilnehmer getestet, indem sie die Bilder von bekannten Politikern zu ihren jeweiligen Parteien zuordnen sollten. Dabei standen mehrere Möglichkeiten zur Auswahl. Im Zuge der Datenaufbereitung wurden die richtigen Antworten mit einem Sternchen im Wertelabel markiert.

5.5. Zusammenfassung nachgefragter Variablen

Im Wahlkampfpanel 2017 wurden mehrere Maßnahmen getroffen um die Qualität der Daten hoch zu halten. Dazu gehörte auch, dass die Befragten die Möglichkeit bekamen, auch dann noch Angaben zu ausgewählten Fragen zu machen, wenn sie ohne irgendwelche Angaben auf „weiter“ geklickt hatten. Dies betrifft jene Fragen, bei denen zur Beantwortung Checkboxen oder Schieberegler verwendet wurden, die einen voreingestellten Wert (meistens von null) aufwiesen. Eine Verweigerung der Antwort hätte in diesen Fällen zur Folge, dass sich diese in der Codierung nicht von möglichen substantziellen Antworten, wie z.B. „null Tage“, unterscheiden. Wenn die Software daher keine Veränderung in den Checkboxen oder Schieberegeln feststellte, wurden die Befragten darauf hingewiesen, dass auf der Umfrageseite keine Veränderung vorgenommen wurde, so dass sie entweder die Gelegenheit hatten substantzielle Antworten zu geben oder eine Aussage dazu machen mussten, warum sie keine Angaben gemacht hatten. Für den veröffentlichten Datensatz wurden die sofort getätigten und durch Nachfrage gewonnenen Antworten in der Stammvariable zusammengefasst. Ob die Antwort erst auf Nachfrage gegeben wurde, lässt sich über eine Flagvariable herausfinden, die aus dem Variablenstammmen und dem Suffix „flag“ besteht. Dies betrifft die Variablen kp6_1681a-f (Fernsehnutzung, Nachrichten aktuell), kp6_1661a-h (Printmediennutzung, politisch aktuell), kp6_1933 (Politische Gespräche, allgemein), kp6_1945a-c (Gesprächshäufigkeit) und kp6_421ka-ei,y sowie kp6_421ha-ii,z (Parteikontakte).

5.6. Hinweise zu den Zeitvariablen

Die Zeitmessung erfolgte serverseitig. Dies hat den Vorteil, dass allen Zeiten dieselbe Systemzeit zu Grunde liegt. Dies wäre bei einer clientseitigen Messung nicht der Fall. Es liegen Informationen zur gesamten Ausfüllzeit sowie zum Verbleib bei den einzelnen Fragen vor.

6 Hinzuspielen anderer Datensätze

6.1. Hinzuspielen von Zeitvariablen

Die zum Kontrollquerschnitt des Wahlkampfpanels zugehörigen Zeitvariablen werden in einem separaten Datensatz zum Herunterladen angeboten und können mittels der Variable „lfdn“ (Laufende Nummer) an den jeweiligen Kontrollquerschnitt herangespielt werden. Alternativ kann auch ein Do-File für Stata oder eine SPSS-Syntax heruntergeladen werden, welche die Zusammenführung der beiden Datensätze übernehmen.

Der Datensatz enthält zwei Arten von Zeitvariablen: Die erste Gruppe von Zeitvariablen gibt die Antwortzeit eines Befragten auf einer speziellen Umfrageseite wieder. Diese Variablen sind nach dem Schema T_Variablenname benannt (Beispiel: „T_kp6_010“ für das Politische Interesse).

Die zweite Gruppe von Zeitvariablen enthält die kumulierte Zeit eines Befragten, die benötigt wurde, um bis an die jeweilige Stelle der Umfrage zu gelangen. Diese kumulierten Zeitvariablen sind entsprechend der Reihenfolge der Fragen innerhalb des Fragebogens durchnummeriert, wobei die Zahl in der Bezeichnung nach dem T die Stelle innerhalb der Umfragesequenz angibt. Zum Beispiel wurde das politische Interesse auf dem 8. Bildschirm in der 1. Welle abgefragt und die kumulierte Zeitvariable wurde demnach mit „T8_kp6_010“ benannt.

6.2. Hinzuspielen des GLES-Wahlkampfpanels 2017

Die Kontrollquerschnitte, die zeitgleich zur vierten, sechsten und achten Befragung des Wahlkampfpanels 2017 erhoben werden, umfassen in etwa 1.000 Personen und werden mit einem fast identischen Fragebogen durchgeführt (ZA-Nr. 6805, 6806 und 6807). Das Wahlkampfpanel wurde ebenfalls bei GESIS veröffentlicht. Da die Variablen im Wahlkampfpanel und in den Kontrollquerschnitten dieselben Namen haben, können alle Datensätze problemlos aneinandergesetzt werden. Hierfür wurde von GESIS ein Stata-Do-File bereitgestellt.

7 Versionshistorie und Errata

7.1. Errata

Eine aktuelle Errataliste kann über den GESIS Datenbestandskatalog (www.gesis.org/dbk) abgerufen werden. Dort werden auch alle Änderungen zwischen den einzelnen Versionen der Datensätze dokumentiert.

7.3. Versionshistorie

Änderungen von Version 1.0.0 zu Version 2.0.0

Für die Zuordnung zu Wahlkreisen wurden die Variablen `elecdist`, `elecdist1`, `elecdist2`, `elecdist3` und `elecdist4` mit dem Stata-Ado „`plztowknr`“ neu erstellt. Die Variablen beziehen sich nun auf den Gebietsstand 2017.

Die Anpassungsgewichtung auf Grundlage des Mikrozensus (`wei_mzmz` `wei_mzoz`) wurden neu berechnet und den Randverteilungen des Jahres 2016 angepasst.

Offene Fragen (Fragen ohne vorgegebene Antwortmöglichkeiten) zu den wichtigsten Problemen sowie den Gründen für die Wahlentscheidung wurden durch das Umfrageinstitut BACES auf Grundlage der Codierschemata der GLES codiert und dem Datensatz zugespielt. Zudem wurden die Variablen mit offenen Angaben auf Datenschutz geprüft und in eine separaten `.csv`-Datei veröffentlicht.

Der Name der Variable `kp6_262s` „Gründe für Wahlentscheidung, offen“ wurde entsprechend der Benennung im Wahlkampf-Panel in `kp6_260s` geändert.

Der Name der Variable `kp6_650x` „Skalometer Politiker: Frauke Petry“ wurde entsprechend der Benennung im Wahlkampf-Panel in `kp6_650p` geändert.

Jeder Datensatz der GLES ist, je nach seiner Struktur, mit bis zu sieben Variablen angereichert, die den Datensatz identifizieren und in seinen Grundzügen beschreiben (Metadaten). Dieses Variablenschema wurde aktualisiert und die Änderungen in der aktuellen Version des Datensatz angepasst. Es ergeben sich daraus folgende Änderungen:

Alte Variable	Neue Variable	Erläuterungen
<code>study</code>	<code>study</code>	Die numerische Variable <code>study</code> beinhaltet als Wert die (ZA-) Studiennummer des Datensatzes im vierstelligen Format. Das Wertelabel gibt nun neben der ZA-Nummer zusätzlich den Titel der jeweiligen Studie an.
<code>survey</code>	<code>sample</code>	Die Variable „ <code>survey</code> “ enthält redundante Informationen zur aktualisierten Variable „ <code>study</code> “ und entfällt in Zukunft. Die numerische Variable „ <code>sample</code> “ dient der Identifikation der Stichprobe bzw. Erhebung, aus der ein Befragter stammt. Als studienübergreifende GESIS Archivvariable harmonisiert sie diese Angaben zwischen den GLES-Studien.
<code>field</code>	<code>field_end</code> , <code>field_start</code>	In der String-Variablen „ <code>field_start</code> “ ist das Datum des ersten Tages der Erhebung im Format 'JJJJ-MM-TT' gespeichert. In der String-Variablen „ <code>field_end</code> “ ist das Datum des letzten Tages der Erhebung im Format 'JJJJ-MM-TT' gespeichert. Sie ersetzen die Variable <code>field</code> , in der diese Informationen vorher zusammengefasst enthalten war.
<code>kp6_field</code>	entfällt	Die Variable <code>kp6_field</code> entfällt. Informationen zur Feldzeit der einzelnen Wellen können der Studienbeschreibung entnommen

year	entfällt	werden. Die Variable „year“ enthält redundante Informationen zur neu erstellten Variable „sample“ und entfällt in Zukunft. Aus der Information zur Stichprobe kann abgelesen werden, in welchem Jahr die Befragten mit der Panelumfrage begonnen haben.
------	----------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Englischsprachige Versionen des Datensatzes, der Zeitvariablen und des Fragebogens wurden veröffentlicht.

Literaturverzeichnis

- 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.
- Schmitt-Beck, Rüdiger/Rattinger, Hans/Roßteutscher, Sigrid und Bernhard Weißels (2010): Die deutsche Wahlforschung und die German Lognitudinal Election Study (GLES). In: Frank Faulbaum und Christof Wolf (Hrsg.). *Gesellschaftliche Entwicklungen im Spiegel der empirischen Sozialforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 141-172.
- Steinbrecher, Markus/ Roßmann, Joss und Michael Bergmann (2013): Das Wahlkampfpanel der German Longitudinal Election Study 2009. Konzeption, Durchführung, Aufbereitung und Archivierung, GESIS Technical Reports 2013/17.
- 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.

Erstellt wurde der Methodenbericht von der Universität Mannheim und GESIS – Leibniz-Institut für Sozialwissenschaften.

Mannheim/Köln, Dezember 2019